

Comparison of Gun Violence within the United States dependent on Political Party:

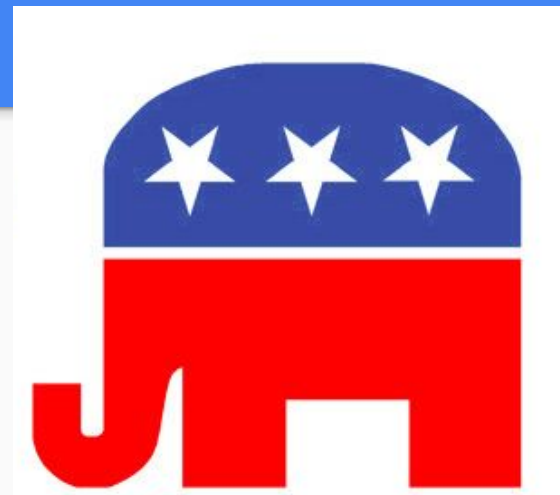
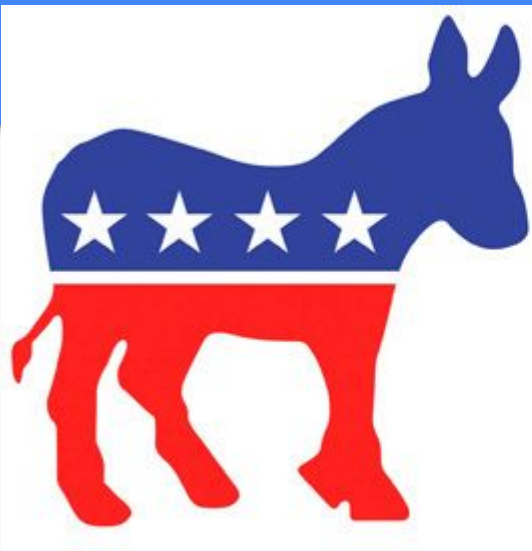
The Obama Administration vs. The Trump Administration

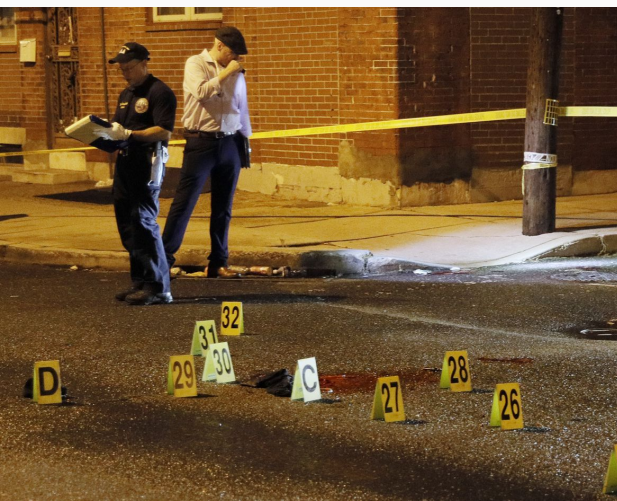
Group 8



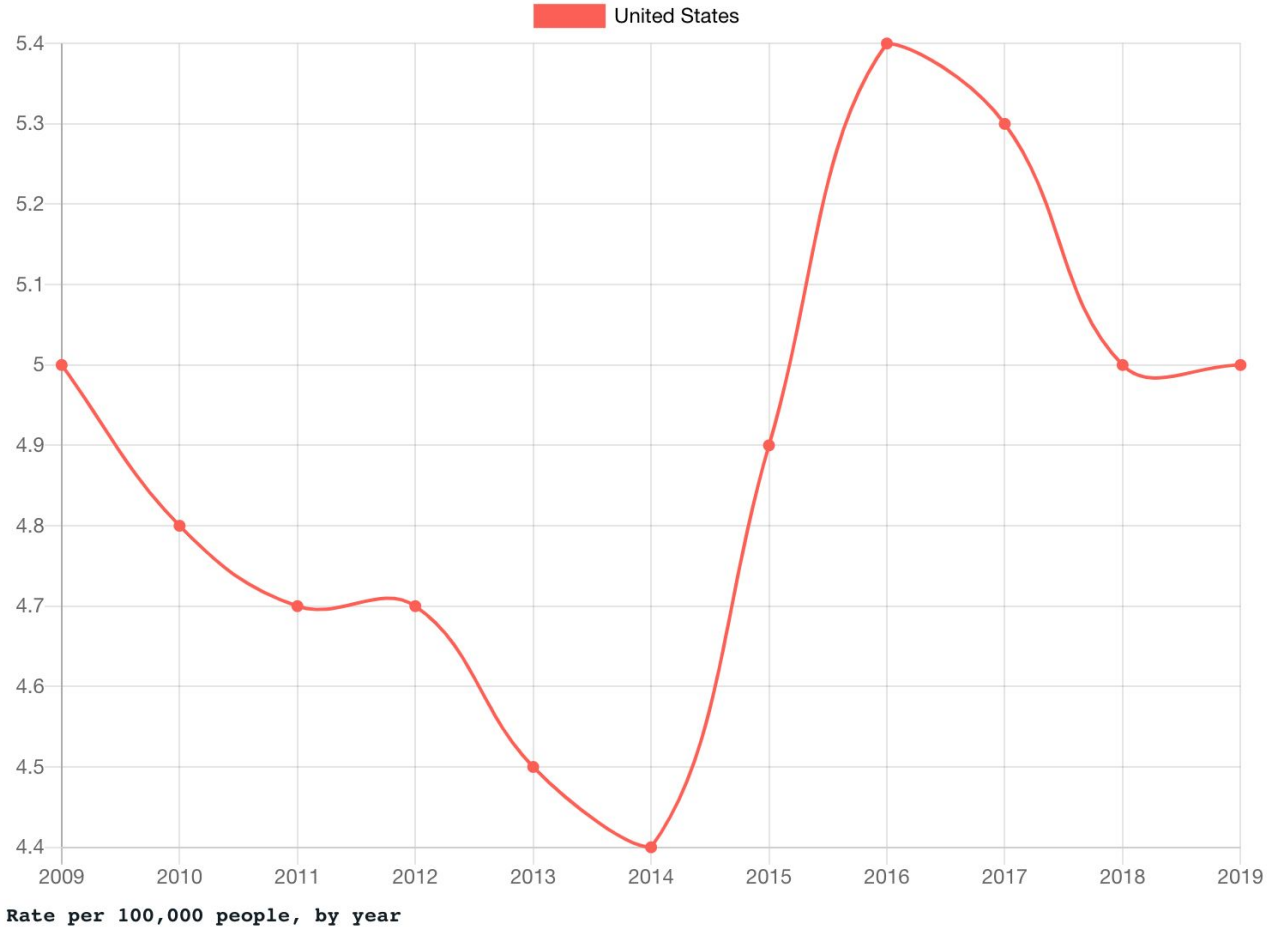
Group 8 Members

- Giang Cao
- Farhana Hossain
- Enoch Kwon
- Jeriel Tenorio
- Nikole Yeung





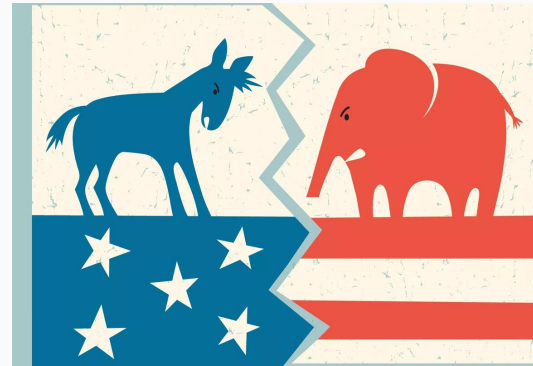
Rate of Homicide Offenses by Population



Hypothesis

Alternative Hypothesis: There are more murder victims by firearms during a **conservative** presidential term (**Trump**) than during a **liberal** presidential term (**Obama**).

Null Hypothesis: There is no relationship between the political party affiliation of the president and murder victims by firearms.



Research Questions

Hypothesis: There are more murder victims by firearms during a **conservative** presidential term (**Trump**) than during a **liberal** presidential term (**Obama**)

- Is this strongly differentiated by specific states?

Firearm/Gun “type”

The Firearms and Guns in question will be based on the following categories:

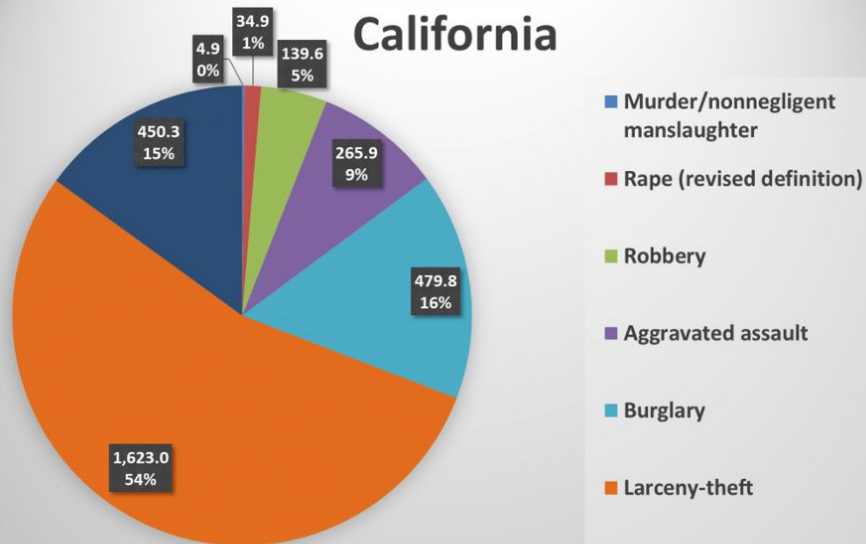
- Handguns
- Rifles
- Shotguns
- Other guns
- Firearms (type not stated)



FBI's Uniform Crime Report (UCR)

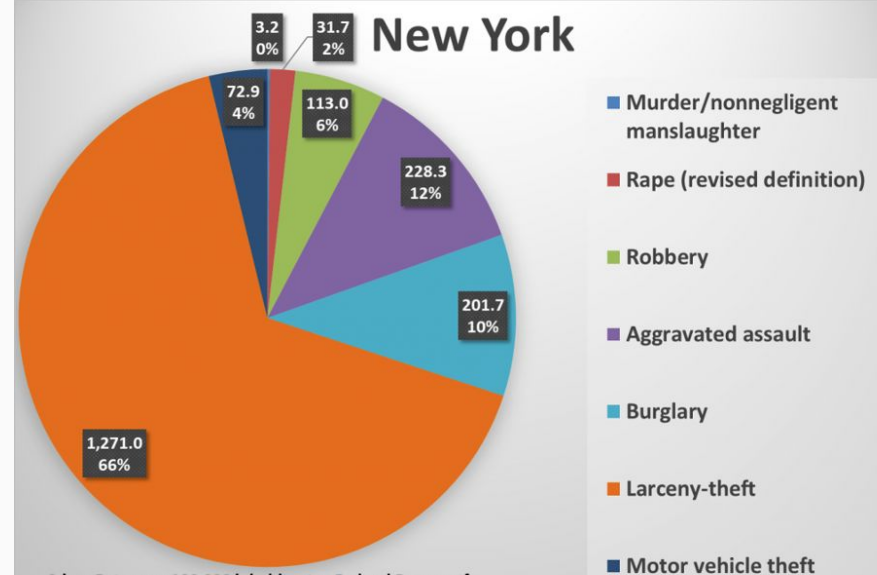


California



Crime Rates per 100,000 inhabitants - Federal Bureau of Investigation - Uniform Crime Reports - Crime in the United States 2016

New York



Crime Rates per 100,000 inhabitants - Federal Bureau of Investigation - Uniform Crime Reports - Crime in the United States 2016

Pros and Cons of the UCR



Pros:

- It looks at crime per 100,000 inhabitants per state, allowing for comparison between states and regions
- Clean and clear Data Set from a respectable department
- Well organized and publically available

Cons:

- Typically a larger population needs a larger sample, however, they are all 100,000.
 - Larger states might not be accurately represented.
- Some of the descriptions are vague, especially with the available API Data.

Navigating through the CSV/Excel data

Obama vs. Trump Administration Gun Violence Combined Data

Firearm Murders by year

```
[3]: import warnings
warnings.filterwarnings('ignore')
```

```
[4]: # Import Dependencies
%matplotlib inline
from matplotlib import pyplot as plt
import numpy as np
import scipy.stats as stats
import pandas as pd
```

```
[5]: # Store filepath in a variable
crime_df = pd.read_excel('HomicideData.xlsx')
crime_df
```

```
t[5]:
```

	Weapons	2014	2015	2016	2017	2018
0	Total	12278	13780	15318	15195	14123
1	Total_firearms	7803	9103	10372	11006	10265

Finding the Means

```
# find means
```

```
crime_df.rename(columns = {2014: "Year2014", 2015: "Year2015", 2016: "Year2016", 2017: "Year2017", 2018: "Year2018"})
```

```
crime_df.head()
```

	Weapons	Year2014	Year2015	Year2016	Year2017	Year2018
0	Total	12278	13780	15318	15195	14123
1	Total_firearms	7803	9103	10372	11006	10265
2	Handguns	5342	6176	6762	7051	6603
3	Rifles	235	215	300	390	297
4	Shotguns	238	247	247	264	235

```
Obama_average_15_16 = ((crime_df["Year2015"] + crime_df["Year2016"])/2).round().astype(int)
```

```
Obama_average_15_16
```

Combined Data (Firearm Murders by Year and State)

```
crime_firearms_df = crime_df.set_index("Weapons")  
crime_firearms_df.head()
```

	Year2014	Year2015	Year2016	Year2017	Year2018	obama_mean	trump_mean
Weapons							
Total	12278	13780	15318	15195	14123	14549	14659
Total_firearms	7803	9103	10372	11006	10265	9738	10636
Handguns	5342	6176	6762	7051	6603	6469	6827
Rifles	235	215	300	390	297	258	344
Shotguns	238	247	247	264	235	247	250

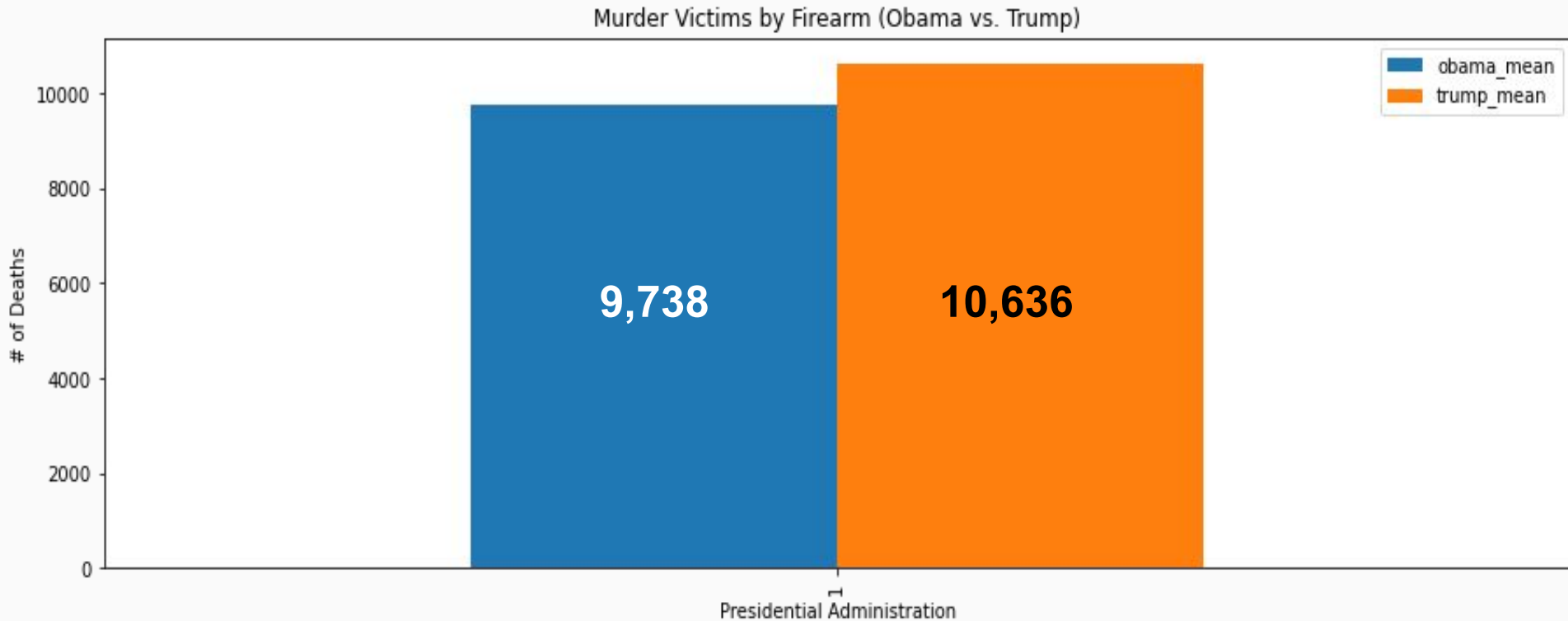
Creating the Bar Graphs...

```
means = averages.loc[averages["Weapons"] == "Total_firearms"]  
means
```

	Weapons	obama_mean	trump_mean
1	Total_firearms	9738	10636

```
means.plot(kind="bar", figsize=(15,5))  
  
plt.title("Murder Victims by Firearm (Obama vs. Trump)")  
plt.xlabel("Presidential Administration")  
plt.ylabel("# of Deaths")  
  
plt.show()  
plt.tight_layout()
```

Murder Victims by Firearm (Obama vs. Trump)



Firearm Murders By “STATE”

```
In [3]: # Import Dependencies
%matplotlib inline
from matplotlib import pyplot as plt
import numpy as np
import scipy.stats as stats
import pandas as pd
```

```
In [4]: # Store filepath in a variable 2015
states2015_df = pd.read_excel('Murder2015.xlsx')
states2015_df.head()
```

```
Out[4]:
```

	State	Total_murders	Total_firearms	Hand_guns	Rifles	Shotguns	Firearms(typeunknown)	Knives orcuttinginstruments	Other_weapons	Hands_fists
0	Alabama	3	3	1	0	1	1	0	0	0
1	Alaska	57	39	12	2	1	24	7	8	3
2	Arizona	278	171	128	4	3	36	42	55	10
3	Arkansas	164	110	51	10	4	45	18	30	6
4	California	1861	1275	855	34	33	353	263	233	90

```
In [5]: #Rename columns 2015
states2015_df.rename(columns = {"Total_murders": "Total_murders2015", "Total_firearms": "Total_firearms2015"}, inplace = True)
states2015_df.head()
```

```
Out[5]:
```

	State	Total_murders2015	Total_firearms2015	Hand_guns	Rifles	Shotguns	Firearms(typeunknown)	Knives orcuttinginstruments	Other_weapons	Hands_fists
0	Alabama	3	3	1	0	1	1	0	0	0
1	Alaska	57	39	12	2	1	24	7	8	3
2	Arizona	278	171	128	4	3	36	42	55	10
3	Arkansas	164	110	51	10	4	45	18	30	6
4	California	1861	1275	855	34	33	353	263	233	90

```
In [6]: # sift columns 2015

data2015 = states2015_df [["State", "Total_murders2015", "Total_firearms2015"]]
data2015.head()
```

```
Out[6]:
```

	State	Total_murders2015	Total_firearms2015
0	Alabama	3	3
1	Alaska	57	39
2	Arizona	278	171
3	Arkansas	164	110
4	California	1861	1275

Combining Year Data + Finding Means

```
: # calculate obama total firearms for years 2015 and 2016
```

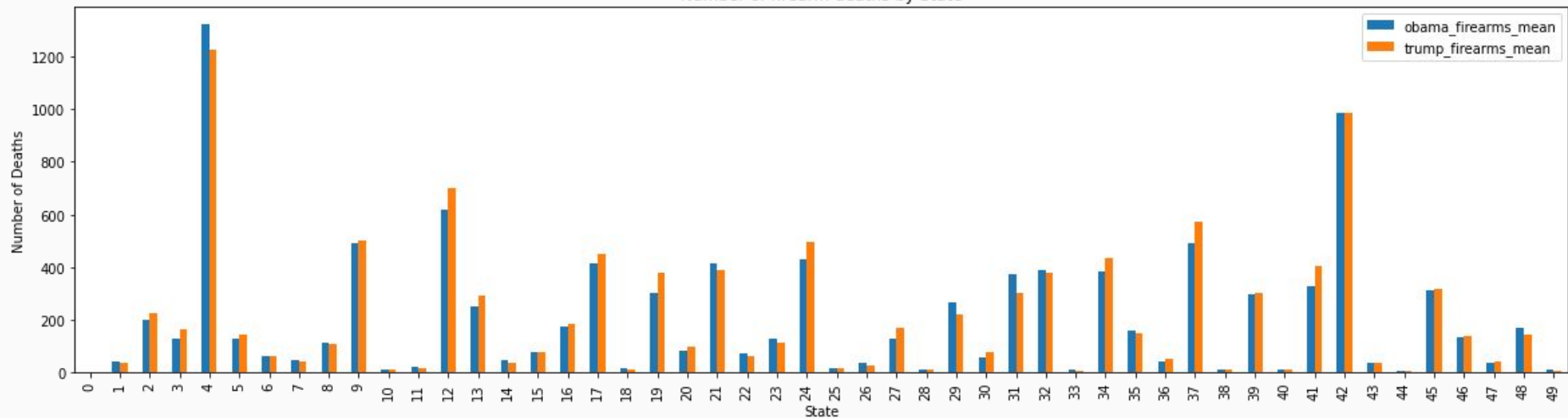
```
Obama_total_firearms_15_16 = ((obama_merge_df["Total_firearms2015"] + obama_merge_df["Total_firearms2016"])/2).round().astype(int)
Obama_total_firearms_15_16.head()
```

```
<
0      2
1     41
2    199
3    130
4   1322
dtype: int64
```

```
: obama_merge_df["obama_murder_mean"] = Obama_total_murders_15_16
```

```
: obama_merge_df.head()
```

Number of firearm deaths by State



0 - Alabama

1 - Alaska

2 - Arizona

3 - Arkansas

4 - California

5 - Colorado

6 - Connecticut

7 - Delaware

8 - District of Columbia

9 - Georgia

10 - Hawaii

11 - Idaho

12 - Illinois

13 - Indiana

14 - Iowa

15 - Kansas

16 - Kentucky

17 - Louisiana

18 - Maine

19 - Maryland

20 - Massachusetts

21 - Michigan

22 - Minnesota

23 - Mississippi

24 - Missouri

25 - Montana

26 - Nebraska

27 - Nevada

28 - New Hampshire

29 - New Jersey

30 - New Mexico

31 - New York

32 - North Carolina

33 - North Dakota

34 - Ohio

35 - Oklahoma

36 - Oregon

37 - Pennsylvania

38 - Rhode Island

39 - South Carolina

40 - South Dakota

41 - Tennessee

42 - Texas

43 - Utah

44 - Vermont

45 - Virginia

46 - Washington

47 - West Virginia

48 - Wisconsin

49 - Wyoming

Data Cleanup and Analysis (1)

Tar file

```
1 # read tar g.zip file - tar.zip file can load to main repo
2 df = pd.read_csv('Crime_DataFiles/DATA_01-2013_03-2018.tar.gz', compression='gzip', header=0, sep=',', error_bad
< >
```

```
1 # check column values
2 crime_df = pd.DataFrame(df)
3 print(crime_df.columns.values)
```

```
['stage3.csv' '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']
```

```
1 # create crime dataframe and rename columns
2 crime_df = pd.DataFrame(df, columns = ['stage3.csv', 'date', 'state', 'city_or_county', 'n_killed', 'n_injured',
3 crime_df.rename(columns = {'stage3.csv': 'Incident', 'date': 'Date', 'state': 'State', 'city_or_county': 'City or Countr
4 'n_killed': 'Number Killed', 'n_injured': 'Number Injured', 'participant_gender': 'Gender
5 'participant_age_group': 'Age Group', 'latitude': 'Latitude', 'longitude': 'Longitude'}, i
6 crime_df.head()
< >
```

6]:

	Incident	Date	State	City or Country	Number Killed	Number Injured	Gender Involved	Age Group	Latitude	Longitude
0	461105.0	2013-01-01	Pennsylvania	Mckeesport	0.0	4.0	0::Male 1::Male 3::Male 4::Female	0::Adult 18+ 1::Adult 18+ 2::Adult 18+ 3::A...	40.3467	-79.8559
1	460726.0	2013-01-01	California	Hawthorne	1.0	3.0	0::Male	0::Adult 18+ 1::Adult 18+ 2::Adult 18+ 3::A...	33.9090	-118.3330

Data Cleanup and Analysis (2)

```
1 # Try to 2 year of data for Obama and Trump Presidency, but we find out there was no data after 3/31/2018.
2 ObamaTrump_df = crime_df.loc[(crime_df['Date']>='2015-01-01') & (crime_df['Date']<'20190101')]
```

```
1 Add Column Name - Presidency and fill in base on 15 months of Obama and Trump was President
2 Fill in the column Presidency - Obama or Trump based on Date column.
3 amaTrump_df.loc[(ObamaTrump_df['Date']>='2015-01-01') & (ObamaTrump_df['Date']<'2016-04-01'), 'Presidency'] = 'Ob
4 amaTrump_df.loc[(ObamaTrump_df['Date']>='2017-01-01') & (ObamaTrump_df['Date']<'2018-04-01'), 'Presidency'] = 'Tr
5 amaTrump_df
```

	Incident	Date	State	City or Country	Number Killed	Number Injured	Gender Involved	Age Group	Latitude	Longitude	Presidency
52132	274168.0	2015-01-01	Oklahoma	Tulsa	0.0	2.0	0::Female 1::Male	0::Adult 18+ 1::Adult 18+	36.0934	-95.8870	Obama
52133	276211.0	2015-01-01	Louisiana	Labadieville	1.0	0.0	0::Male	0::Adult 18+	29.8312	-90.9609	Obama
52134	272302.0	2015-01-01	Mississippi	Hattiesburg	0.0	0.0	0::Male	0::Adult 18+	31.3271	-89.2903	Obama
52135	272482.0	2015-01-01	Alabama	Bessemer	0.0	2.0	0::Male 1::Male	NaN	33.4423	-86.9322	Obama
52136	272487.0	2015-01-01	Illinois	Chicago	0.0	1.0	0::Male	NaN	41.9202	-87.7857	Obama
...
239672	1083142.0	2018-03-31	Louisiana	Rayne	0.0	0.0	0::Female	0::Adult 18+	NaN	NaN	Trump

Data Cleanup and Analysis (3)

Parse Gender Involved and Age Group

```
1 # Parse Gender Involved and Age Group to its own columns columns and add Year Month column to roll up
2 # counts by months
3 ObamaTrump_df['Male'] = ObamaTrump_df['Gender Involved'].str.count('Male')
4 ObamaTrump_df['Female'] = ObamaTrump_df['Gender Involved'].str.count('Female')
5 ObamaTrump_df['Child 0-11'] = ObamaTrump_df['Age Group'].str.count('Child 0-11')
6 ObamaTrump_df['Teen 12-17'] = ObamaTrump_df['Age Group'].str.count('Teen 12-17')
7 ObamaTrump_df['Adult 18+'] = ObamaTrump_df['Age Group'].str.count('Adult 18+')
8 ObamaTrump_df['Year Month'] = pd.DatetimeIndex(ObamaTrump_df['Date']).to_period('M')
9 #ObamaTrump_df['Year Month'] = pd.to_datetime(ObamaTrump_df['Date']).dt.to_period('M')
10 #ObamaTrump_df['Year Month'] = [''.join(x.split('-')[0:2]) for x in ObamaTrump_df.Date]
11 ObamaTrump_df.head()
12
13 #output_file = "Crime_DataFiles/test_analysis.csv"
14 #ObamaTrump_df.to_csv(output_file, index = False)
```

01]:

	Incident	Date	State	City or Country	Number Killed	Number Injured	Gender Involved	Age Group	Latitude	Longitude	Presidency	Male	Female	Child 0-11
52132	274168.0	2015-01-01	Oklahoma	Tulsa	0.0	2.0	0::Female 1::Male	0::Adult 18+ 1::Adult 18+	36.0934	-95.8870	Obama	1.0	1.0	0.0
52133	276211.0	2015-01-01	Louisiana	Labadieville	1.0	0.0	0::Male	0::Adult 18+	29.8312	-90.9609	Obama	1.0	0.0	0.0
52134	272302.0	2015-01-01	Mississippi	Hattiesburg	0.0	0.0	0::Male	0::Adult 18+	31.3271	-89.2903	Obama	1.0	0.0	0.0
52135	272482.0	2015-01-01	Alabama	Bessemer	0.0	2.0	0::Male 1::Male	NaN	33.4423	-86.9322	Obama	2.0	0.0	NaN
52136	272487.0	2015-01-01	Illinois	Chicago	0.0	1.0	0::Male	NaN	41.9202	-87.7857	Obama	1.0	0.0	NaN

Data Cleanup and Analysis (4)

```
1 # Data cleaning - Drop null/NaN values on Presidency fields
2 ObamaTrump_df.dropna(subset=['Presidency'],inplace=True)
3 #ObamaTrump_df.dropna(subset=['Presidency','Male','Female','Child 0-11','Teen 12-17','Adult 18+'],inplace=True)
4 #ObamaTrump_df.dropna(subset=['Presidency','Male','Female'],inplace=True)
```

```
1 # Export Data result to CSV file - work with a smaller subset of data
2 output_file = "Crime_DataFiles/crime_analysis.csv"
3 ObamaTrump_df.to_csv(output_file, index = False)
```

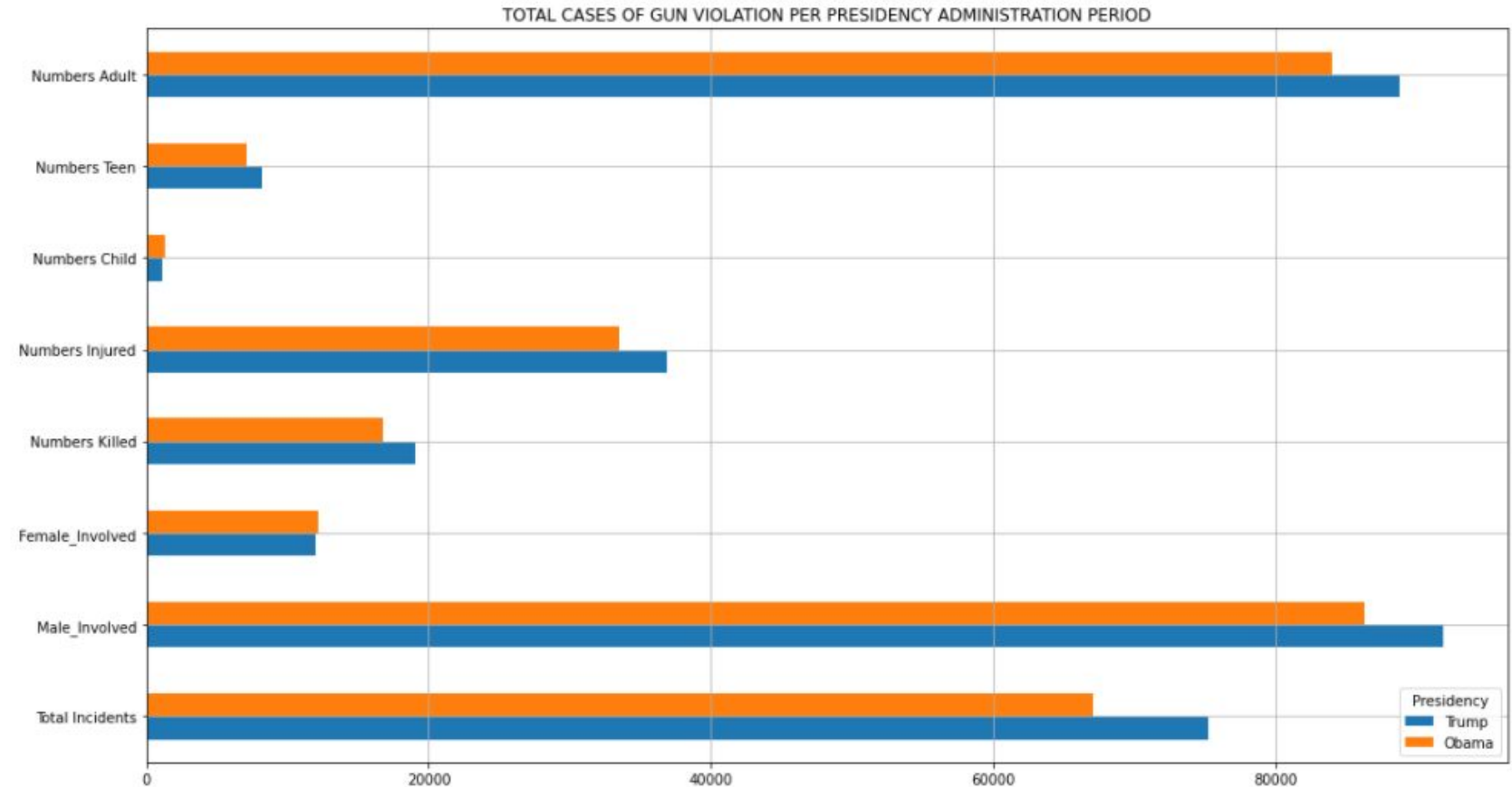
Run Data From Export CSV

```
1 # Run using a subset of data exported out for analysis
2 # read csv
3 crime_file= "Crime_DataFiles/crime_analysis.csv"
4 re_Crime =pd.read_csv(crime_file)
5 re_Crime.head()
```

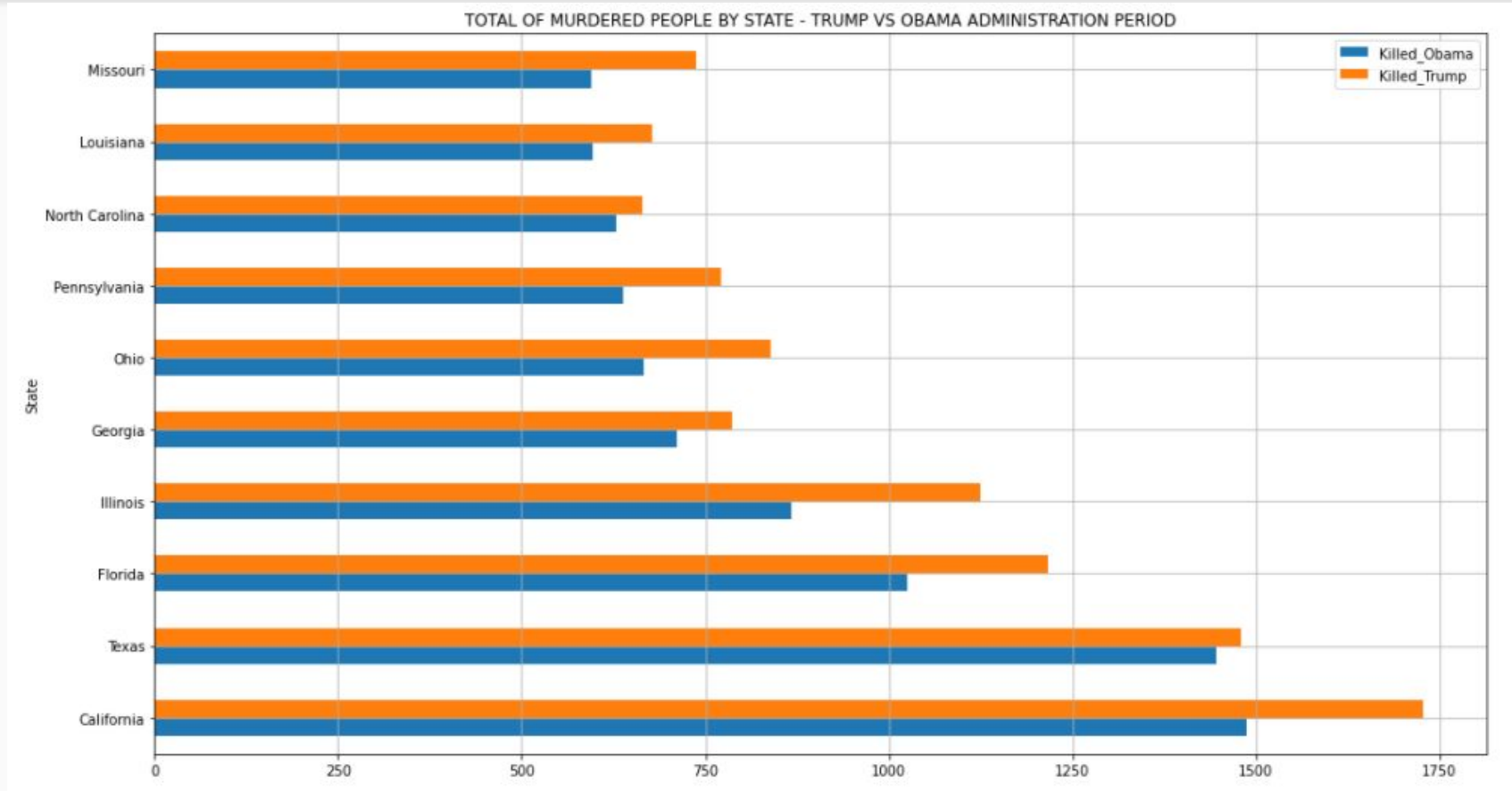
3]:

	Incident	Date	State	City or Country	Number Killed	Number Injured	Gender Involved	Age Group	Latitude	Longitude	Presidency	Male	Female	Child 0-11	Teen 12-17
0	274168.0	2015-01-01	Oklahoma	Tulsa	0.0	2.0	0::Female 1::Male	0::Adult 18+ 1::Adult 18+	36.0934	-95.8870	Obama	1.0	1.0	0.0	0.0
1	276211.0	2015-01-01	Louisiana	Labadieville	1.0	0.0	0::Male	0::Adult 18+	29.8312	-90.9609	Obama	1.0	0.0	0.0	0.0
2	272302.0	2015-01-01	Mississippi	Hattiesburg	0.0	0.0	0::Male	0::Adult 18+	31.3271	-89.2003	Obama	1.0	0.0	0.0	0.0

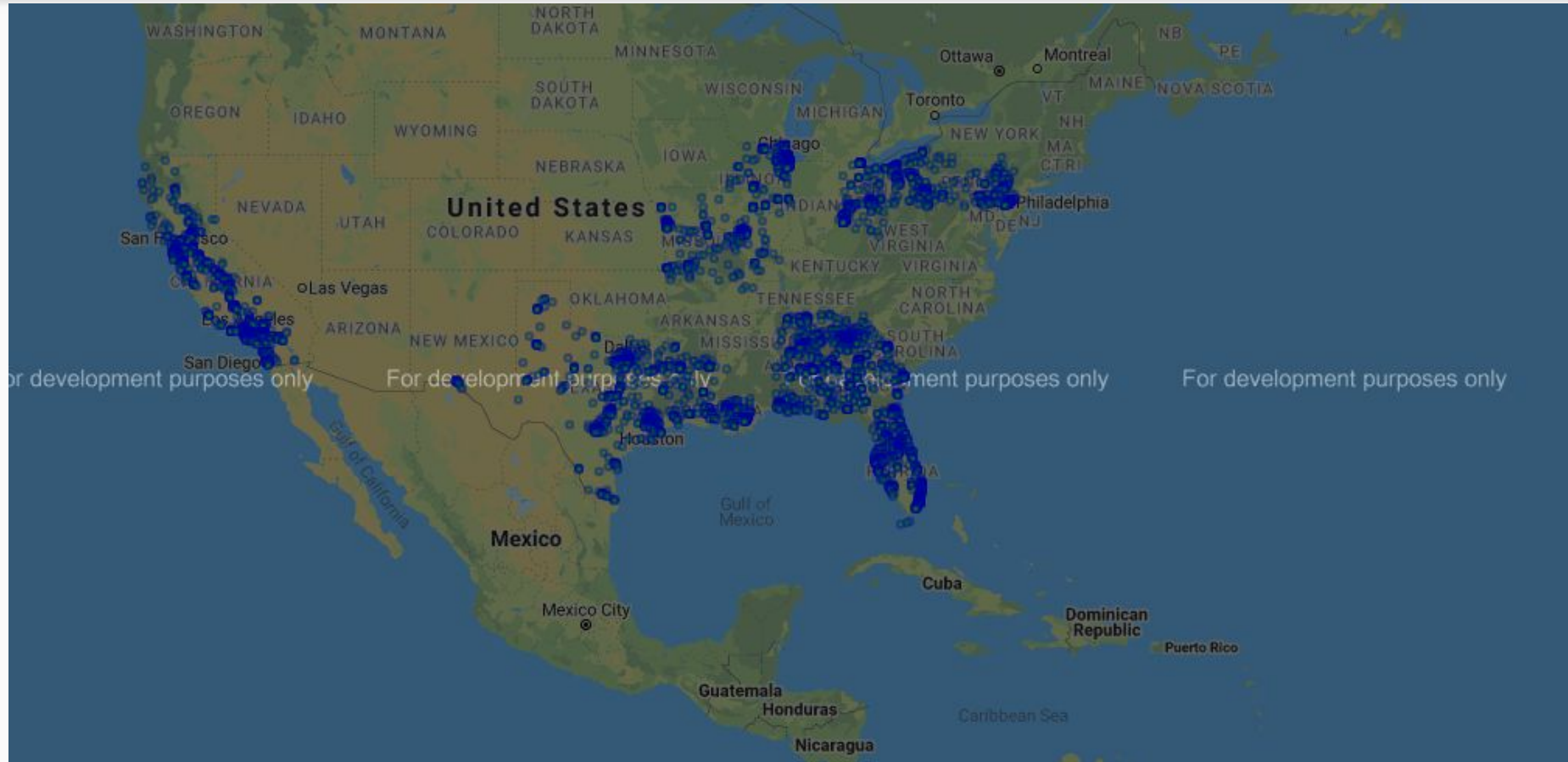
Demographic Visualization of Gun Violence by Presidency Administration



Total number of people murdered (By State)



Top 5 States for Firearm Murders(Jan. 2017 - March 31, 2018)



Statistics for Number of Deaths by Presidency

```
1 # Demographic Statistic by Presidency
2 combine_df = [summary_Trump_killed, summary_Obama_killed]
3 pres_summary_killed = pd.concat(combine_df)
4 pres_summary_killed
```

	Presidency admin	killed Mean	killed Median	killed Variance	killed Std. Dev	killed Std. Err
0	Trump	1270	1304.0	9659	98	25
0	Obama	1119	1133.0	11566	108	28

Independent Samples T-Test

t-Test to statistically significant as $P < 0.05$ and statistically highly significant as $P < 0.001$ (less than one in a thousand chance of being wrong) Prove our Hypothesis/Null Hypothesis

```
1 # T Testing for Killed Counts per month
2
3 population1 = Obama_months["Killed"]
4 population2 = trump_months["Killed"]
5
6 stats.ttest_ind(population1, population2, equal_var=False)
7
8
```

```
5]: Ttest_indResult(statistic=-3.998166236234966, pvalue=0.0004271047689859947)
```

```
1 print(f"HIGHLY statistical confirmation that Killed Counts per month in two presidency administrations are diffe
2
< >
```

HIGHLY statistical confirmation that Killed Counts per month in two presidency administrations are different

Results + Discussion

$t = -3.99$, $p < .001$

Alternative Hypothesis: There are more murder victims by firearms during a **conservative** presidential term (**Trump**) than during a **liberal** presidential term (**Obama**).

Null Hypothesis: ~~There is no relationship between the political party affiliation of the president and murder victims by firearms.~~

Future Directions

- **Stricter gun laws in Liberal states**
- **Would specific state GDP affect these statistics as well?**
- **Does the political affiliation of the State Governor affect these statistics?**
- **Do they differ by US region?**

Limitations

- **Multiple types of data**
- **Comparing Obama's 2nd term with Trump's First term**
- **Limited Info and vague descriptions.**