**Project Title:**

Testing increase/decrease In FBI NICS Firearm Background Checks in Relation to Mass Shootings, 1998 - 2019

**Team Name:**

Project Black Ops

**Members:**

Misael Obregon, Michael Harper, Richard Harris

**Project Description/Outline**

Testing increase/decrease In FBI NICS Firearm Background Checks in Relation to Mass Shootings, 1998 - 2019

**Research Question to Answer**

Is there an increase in firearm checks after incidents of mass shootings?

Is there an increase in firearm checks after firearm/gun control remarks made by politicians?

Can we gauge the duration of spike, if any, after both, mass shooting incident and/or political remarks?

**Datasets to be Used**

FBI NICS Firearm background Check Data (November 1998 – June 2019

US Mass Shootings, 1982 – 2019: Data from Mother Jones’ Investigation

Twitter API

**Rough Breakdown of Tasks**

Team will download two csv files (see above) and will merge datasets. API data will be requested for Twitter feeds from political figures on subject of firearms and gun control for years 2006 – 2019 as part of the analysis.

Mass Shooting Incident (MSI) dates will be used to check for spikes in firearm background checks during and after the month of MSI. A second source of data will do the same (Twitter) after remarks made by political figures (2006 – 2019).

1. **A. Data cleaning note:** data in csv file for mass shooting was deleted before November 1998 (a total of 26 cases). Rationale, firearm permit data starts November 1998.

**B. Merged** and created groupby data frames from two csv files

1. A. Create counts of permit applications for each “month” entry; currently “month” permit counts are by state. We need to combine all states into one month for a national level count.

B. Same for mass shooting data if more than one mass shooting in a single month

C. Create Boolean variables for mass shooting happening in “month” (yes/no)

1. Create tables and graphs for descriptive statistics
2. Trend Line graph merging two data sets:
   * Firearm data
   * Mass shooting data
3. Histograms
4. Bar graphs
5. Map graph:

* Location of mass incidents
* States with permit applications (perhaps heat map of states with high concentration of permit applications.

1. T-test:
2. Mean of number of permit applications in months where there was no mass shooting(s)
3. Mean of number of permit applications during a month when there was a mass shooting(s)
4. Logistic regression:

* Indep. Variable: Likelihood of person applying for gun permit (application)
* Dependent Variables
  + prior to mass shooting incident
  + month of mass shooting incident
  + month after mass shooting incident
  + two months after mass shooting incident

Source for Logistic Regression and Stata comparison in Pandas:

<https://stackoverflow.com/questions/49086721/multinomial-logit-model-python-and-stata-different-results>

import plotly\_express as px

suicides\_gby\_Continent\_2007 = suicides\_gby\_Continent[suicides\_gby\_Continent['year']==2007]

px.scatter(suicides\_gby\_Continent\_2007,x = 'suicides/100k pop', y = 'gdp\_per\_capita ($)')

px.scatter(suicides\_gby\_Continent\_2007,x = 'suicides/100k pop', y = 'gdp\_per\_capita ($)',color='continent')

px.scatter(suicides\_gby\_Continent\_2007,x = 'suicides/100k pop', y = 'gdp\_per\_capita ($)',color='ContinentName',size ='suicides/100k pop')

px.scatter(suicides\_gby\_Continent,x = 'suicides/100k pop', y = 'gdp\_per\_capita ($)',color='continent',

size='suicides/100k pop',symbol='sex',animation\_frame='year', animation\_group='continent',range\_x = [0,0.6],

range\_y = [0,70000],text='continent')

