# CSE485 - Data Mining
## Final Project

# "Terrorism and Ideology Analysis"

Date: 10.05.2019

Instructor = F. Boray TEK

Student = Ş. Taha Demirer / 216CS2007

## Abstract:

There are two datasets, "Global Terrorsim Database" and ""Government Composition". The links between political tendencies of ruling parties, governments and terrorism will be researched. There are 5 main countries with terrorism and government ideology data.

- US, UK, Germany, France, Spain

## Introduction and Problem:

Global Terrorism Database is one of the most analyzed databases around Kaggle platform. It includes all terrorist attacks between 1970-2017, terrorist attack's country, date, attack type, number of wounded persons, terrorist organisation identifiers etc. It's useful for exploring places of attacks.

Government Composition database holds ideology, distribution of parliament chairs, dates etc. That dataset is helpful for describing a government with it's ruling years and ideology.

Main hypothesis is **If the ruler parties are tend to be "right" ideology, terrorist events are more than the other times.**

## Datasets:

**-** Global Terrorism Database, CSV, 155MB , https://www.kaggle.com/START-UMD/gtd

Government Composition Dataset, XLSX, 1.54MB, Armingeon, Klaus, Virginia Wenger, Fiona Wiedemeier, Christian Isler, Laura Knöpfel and David Weisstanner. 2018. Supplement to the Comparative Political Data Set – Government Composition 1960-2016, Bern: Institute of Political Science, University of Berne.

## Code Sources:

https://www.kaggle.com/raghuchaudhary/global-terrorism

https://www.kaggle.com/ash316/terrorism-around-the-world

# Explatory Data Analysis:

- Firstly without doing anything on datasets, I imported them and described their mean, std, count etc.

| | Year | Month | Day | Country | Region | city | latitude | longitude | AttackType | Killed | Wounded | Target | Summary | Group |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1970 | 7 | 2 | Dominican Republic | Central America & Caribbean | Santo Domingo | 18.456792 | -69.951164 | Assassination | 1.0 | 0.0 | Julio Guzman | NaN | MANO-D |
| 1 | 1970 | 0 | 0 | Mexico | North America | Mexico city | 19.371887 | -99.086624 | Hostage Taking (Kidnapping) | 0.0 | 0.0 | Nadine Chaval, daughter | NaN | 23rd of September Communist League |
| 2 | 1970 | 1 | 0 | Philippines | Southeast Asia | Unknown | 15.478598 | 120.599741 | Assassination | 1.0 | 0.0 | Employee | NaN | Unknown |
| 3 | 1970 | 1 | 0 | Greece | Western Europe | Athens | 37.997490 | 23.762728 | Bombing/Explosion | NaN | NaN | U.S. Embassy | NaN | Unknown |
| 4 | 1970 | 1 | 0 | Japan | East Asia | Fukouka | 33.580412 | 130.396361 | Facility/Infrastructure Attack | NaN | NaN | U.S. Consulate | NaN | Unknown |
| 5 | 1970 | 1 | 1 | United States | North America | Cairo | 37.005105 | -89.176269 | Armed Assault | 0.0 | 0.0 | Cairo Police Headquarters | 1/1/1970: Unknown African American assailants ... | Black Nationalists |
| 6 | 1970 | 1 | 2 | Uruguay | South America | Montevideo | -34.891151 | -56.187214 | Assassination | 0.0 | 0.0 | Juan Maria de Lucah/Chief of Directorate of in... | NaN | Tupamaros (Uruguay) |
| 7 | 1970 | 1 | 2 | United | North | Oakland | 37.791927 | 122.225906 | Bombing/Explosion | 0.0 | 0.0 | Edes | 1/2/1970: Unknown perpetrators | Unknown |

**Fig. 1.0**

The type and first 8 rows of global terrorism data can be clearly seen in **Fig. 1.0.**

I will use only Year and Country. Because every row shows a different attack with different features.

## Describing of GTDB

| | Year | Month | Day | latitude | longitude | Killed | Wounded | casualities |
|---|---|---|---|---|---|---|---|---|
| count | 181691.000000 | 181691.000000 | 181691.000000 | 177135.000000 | 1.771340e+05 | 171378.000000 | 165380.000000 | 164817.000000 |
| mean | 2002.638997 | 6.467277 | 15.505644 | 23.498343 | -4.586957e+02 | 2.403272 | 3.167668 | 5.296128 |
| std | 13.259430 | 3.388303 | 8.814045 | 18.569242 | 2.047790e+05 | 11.545741 | 35.949392 | 42.069023 |
| min | 1970.000000 | 0.000000 | 0.000000 | -53.154613 | -8.618590e+07 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 1991.000000 | 4.000000 | 8.000000 | 11.510046 | 4.545640e+00 | 0.000000 | 0.000000 | 0.000000 |
| 50% | 2009.000000 | 6.000000 | 15.000000 | 31.467463 | 4.324651e+01 | 0.000000 | 0.000000 | 1.000000 |
| 75% | 2014.000000 | 9.000000 | 23.000000 | 34.685087 | 6.871033e+01 | 2.000000 | 2.000000 | 4.000000 |
| max | 2017.000000 | 12.000000 | 31.000000 | 74.633553 | 1.793667e+02 | 1570.000000 | 8191.000000 | 9574.000000 |

**Fig 1.1**

The first impression on untouched data is quite interesting.

What can we say from looking to Fig 1.1?

- There were 181.691 Terrorist attacks happened between 1970-2017
- Attacks were happened around June, 15<sup>th</sup> day.
- 2.4 People died, 3.16 people wounded per attack.
- Maximum kills in a single attack is 1570, max woundeds 8191.


- And then, let's describe governments data.

| | Year | Country | Ideology |
|---|---|---|---|
| 0 | 1959.000000 | USA | right |
| 189 | 1992.000000 | UK | right |
| 175 | 1985.000000 | UK | right |
| 176 | 1986.000000 | UK | right |
| 177 | 1986.000000 | UK | right |
| 178 | 1987.000000 | UK | right |
| 179 | 1987.000000 | UK | right |
| 180 | 1988.000000 | UK | right |
| 181 | 1988.000000 | UK | right |
| 182 | 1989.000000 | UK | right |
| 183 | 1989.000000 | UK | right |
| 184 | 1990.000000 | UK | right |

Fig 1.3

As it seen, only year, country and ideology of government data used in Government Composition Database (GCDB).

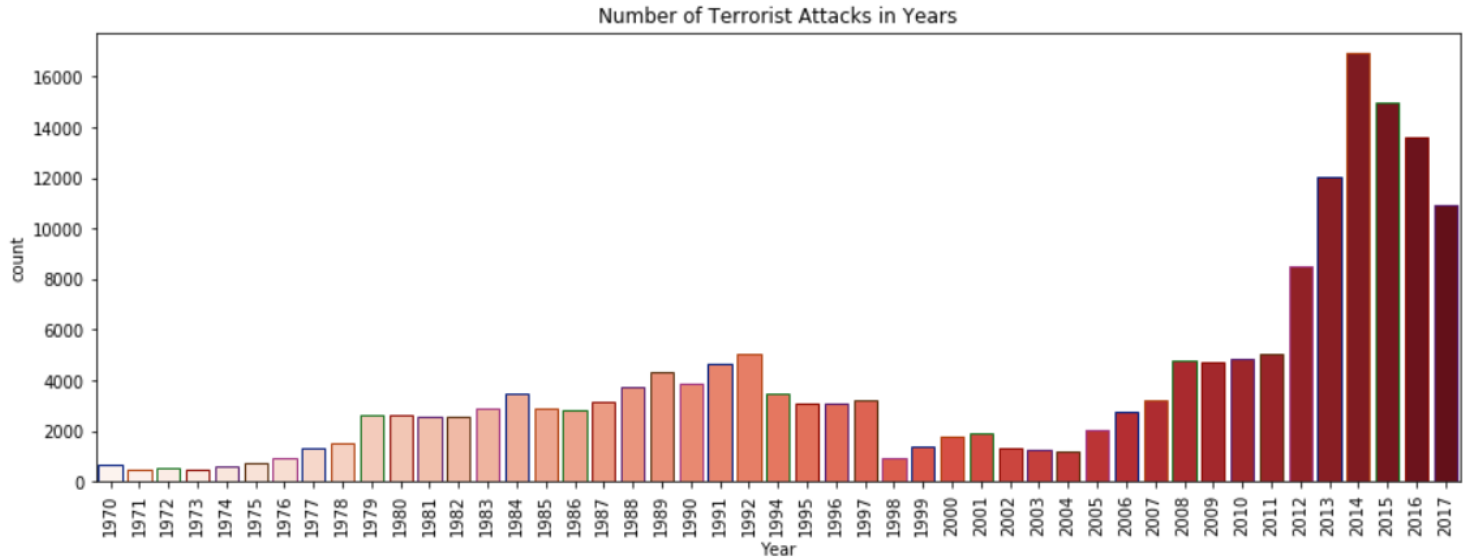- Exploring terrorist attacks depends on years.

Fig 2.0

The year which has most terrorist attacks happened is 2014. Figure 2.0 shows attacks increased after 2004.

- Governments analysis.

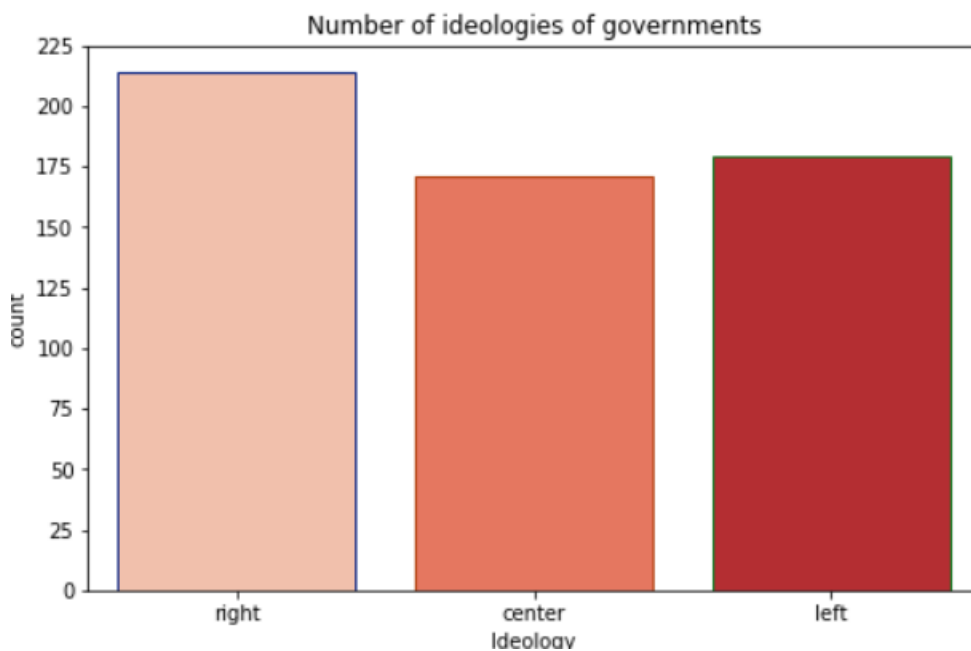We should look at the number of ideologies governments has support.



Fig 2.1

Fig. 2.1 shows us the number of ideologies that governments have.
Between 1960-2016, Right Ideology is the most supported.

- Now, get these two results together. If these two datas gets together depends on their "Year" and "Country" column with pandas.concat(), all the terrorist attack in our five country will be listed.

| | Country | Ideology | Year |
|---|---|---|---|
| 0 | Dominican Republic | NaN | 1970.0 |
| 1 | Mexico | NaN | 1970.0 |
| 2 | Philippines | NaN | 1970.0 |
| 3 | Greece | NaN | 1970.0 |
| 4 | Japan | NaN | 1970.0 |
| 5 | United States | NaN | 1970.0 |
| 6 | Uruguay | NaN | 1970.0 |
| 7 | United States | NaN | 1970.0 |
| 8 | United States | NaN | 1970.0 |
| 9 | United States | NaN | 1970.0 |
| 10 | United States | NaN | 1970.0 |
| 11 | United States | NaN | 1970.0 |
| 12 | Italy | NaN | 1970.0 |
| 13 | United States | NaN | 1970.0 |
| 14 | United States | NaN | 1970.0 |
| 15 | East Germany (GDR) | NaN | 1970.0 |
| 16 | Ethiopia | NaN | 1970.0 |
| 17 | United States | NaN | 1970.0 |
| 18 | United States | NaN | 1970.0 |

Fig 3.0

After concatenating our two lists, there are so many Countries which are not useful for this data, and null fields.

All of them have to be cleaned.

| | | | |
|---|---|---|---|
| 19 | 1972 | France | right |
| 20 | 1972 | France | right |
| 21 | 1972 | France | right |
| 22 | 1973 | France | right |
| 23 | 1973 | France | right |
| 24 | 1973 | France | right |
| 25 | 1973 | France | right |
| 26 | 1973 | France | right |
| 27 | 1973 | France | right |
| 28 | 1973 | France | right |
| 29 | 1973 | France | right |
| ... | ... | ... | ... |
| 12936 | 2016 | Germany | center |
| 12937 | 2016 | Germany | center |
| 12938 | 2016 | Germany | center |
| 12939 | 2016 | Germany | center |
| 12940 | 2016 | Germany | center |
| 12941 | 2016 | Germany | center |
| 12942 | 2016 | Germany | center |

Remaining dataset is so clear and useful to undestand.

Fig 3.1

- Finally, if we visualize that data depends on It's Ideology column, the main hypothesis can be tested.
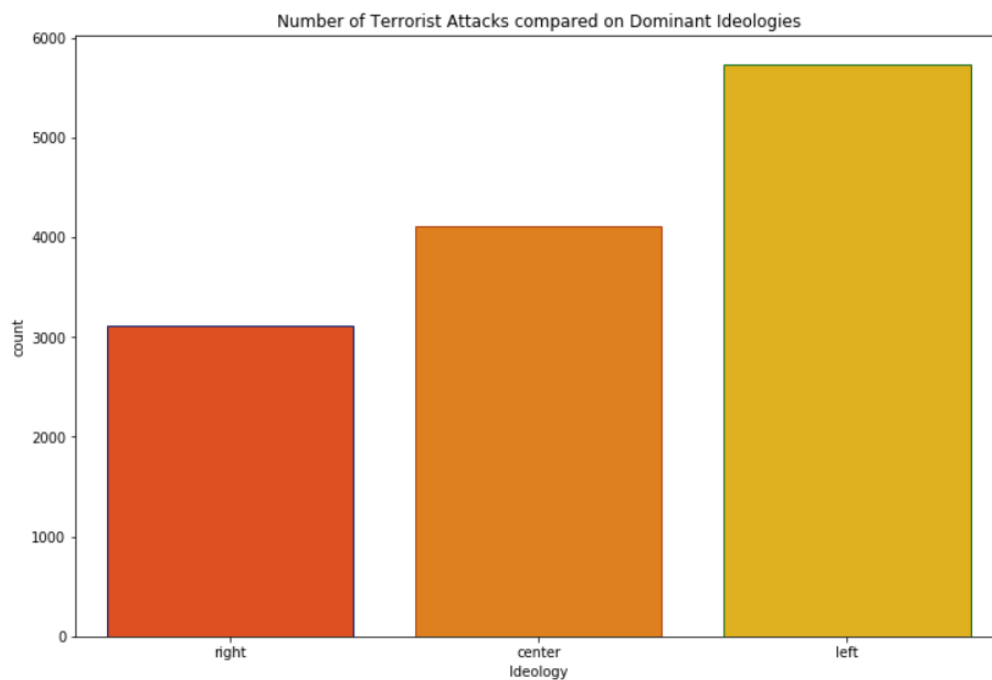


Fig 4.0

- Fig 4.0 shows that hypothesis is totally opposite for 5 countries. Terrorist attacks happening while left ideology is on rule more than right ideology.
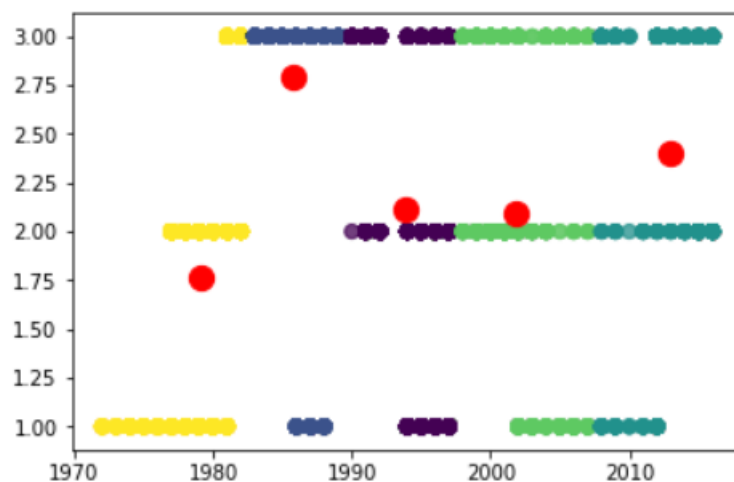
- 5 Clusters K-Means test on final data.



Fig 4.1

**Conclusion:**

There are several final results in this analysis.

1- Governments which has left ideology are weak on counterterrorism.
2- Governments which has left ideology may be founding terrorist organisations.
3- Governments which has Right ideology are better than left ones on counterterrorism.
4- Governments which has on center ideology are not successful against terrorist organisations.

After all these results, It is cleary open that main hypothesis for that analysis is false.
The opposite of the main hypothesis has been proven.