

Capstone Project abstract

CKM136XJ0

Student: Emil Ibrahim
Student Id: 500862456

Ryerson University
Date: 18-Jan-2019

The theme:

Classification, Regression and Predictive Analytics

Datasets link:

https://www.kaggle.com/START-UMD/gtd#globalterrorismdb_0718dist.csv

Context:

Information on more than 180,000 Terrorist Attacks

The Global Terrorism Database (GTD) is an open-source database including information on terrorist attacks around the world from 1970 through 2017. The GTD includes systematic data on domestic as well as international terrorist incidents that have occurred during this time period and now includes more than 180,000 attacks.

Content:

Geography: Worldwide

Time period: 1970-2017, *except 1993*

Variables: >100 variables on location, tactics, perpetrators, targets, and outcomes ,....etc

Research questions:

- What is the trends for Terrorism and can it be predicted (gone up or down over time and with population growth, with internet/technology growth)?
- What is the distribution of attacks (time, location, population,...ect)
- Is there difference in casualties per (terrorist group, type of weapon, country/location)?
- What is the signature for each terrorist group (targets, weapon types, motive,etc)?
- Can number of future attacks be predicted?
- Data visualization (location, groups,etc)
- Deeper analysis for two or three countries.

Tools and Programming language for coding:

Python and different libraries: Pandas, matplotlib, seaborn, numpy, scipy,...etc

I may use AWS (Amazon Web services) if my laptop processing is slow during building / training the models.