

---

**Arianna Lang Wang**

Springboard Jan 2nd, 2018 Cohort

# Global Terrorism Project Proposal

Jan 16th, 2018

## DATA

The Global Terrorism Database (GTD) is an open-source database including information on terrorist attacks around the world from 1970 through 2016. The GTD includes systematic data on domestic as well as international terrorist incidents that have occurred during this time period and includes more than 170,000 cases. For each GTD incident, information is available on the date and location of the incident, the weapons used and nature of the target, the number of casualties, and when identifiable, the group or individual responsible. There is also a significant number of missing information. For example, there are 4606 missing longitude values and 4606 missing latitude values.

This dataset is available on the Kaggle website: <https://www.kaggle.com/START-UMD/gtd/data>

## APPROACH

There are many questions I can ask. For instance, which countries have the highest number of terrorist attacks? Which regions are the safest to live in, i.e. have the least number of terrorist attacks? What type of terrorism attacks is most common in each region? Conversely, can I predict which region will have a particular type of terrorism next? Additionally, I can focus my analyses on terrorism in the United States. For instance, I can map out the number of terrorist attacks by state from 1970 to 2016.

My “imaginary” client is The Central Intelligence Agency (CIA), which is interested in understanding more about terrorist attacks that had happened around the world and in the United States. My plan is to use data wrangling and data mining to clean and explore this large and ‘messy’ dataset and make concise and beautiful graphs and charts so that the agency can visualize and make sense of the data.

---

## **PROJECT DELIVERY**

The final project deliveries will include my Python code, a paper and/or a PowerPoint slide deck.