GLOBAL TERRORISM DATABASE: MODERN ANALYSIS AND VISUALIZATION

About GTD & Project Objectives

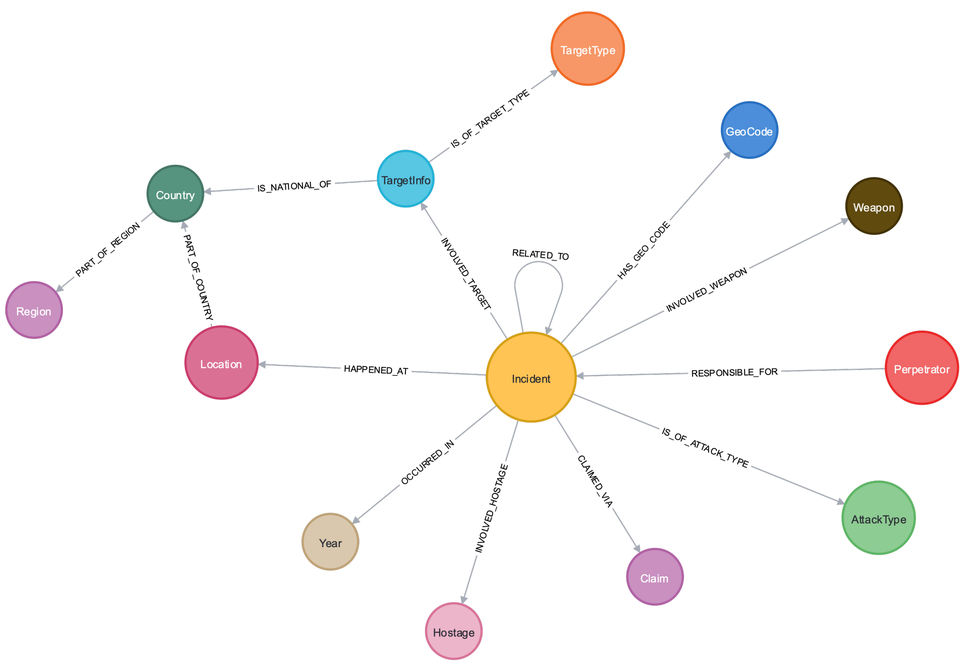
1970-2018
Years covered

200k
Terrorist Attacks

135
Variables per attack

- The Global Terrorism Database (GTD) is an open-source database of global terrorist attacks from 1970 to 2018. It includes almost 200,000 cases and 135 features detailing the type of attack, location, weapons used, casualties, etc.
- The GTD does not represent the data in an interesting and visually appealing manner and users could more effectively understand trends or combat the effects of terrorism, if there was a better way to visualize and mine the data.
- The project objective is to transform the GTD into a more effective database and create a visualization that help citizens, researchers, and government agencies explore terrorism trends and fight terrorism.

Building a Graph Database



- The current GTD database design has all entities and attributes of a terrorism incident in one row. This is an inefficient database design and fails to explicitly show the relationships between entities.
- Our solution is to transform the tabular dataset into a network dataset and build the first network graph based on the GTD, which is stored in a Neo4j Database.
- The advantages of a graph database are that it:
  - stores relationship information as a first-class entity and are adept at working with evolving relationship networks.
  - allows for the addition of new nodes and relationships without compromising the existing network or expensively migrating data.
  - is centered around data relationships and is highly efficient when it comes to query performance, even for deep and complex queries.

Machine Learning to Understand & Predict Fatality Rates

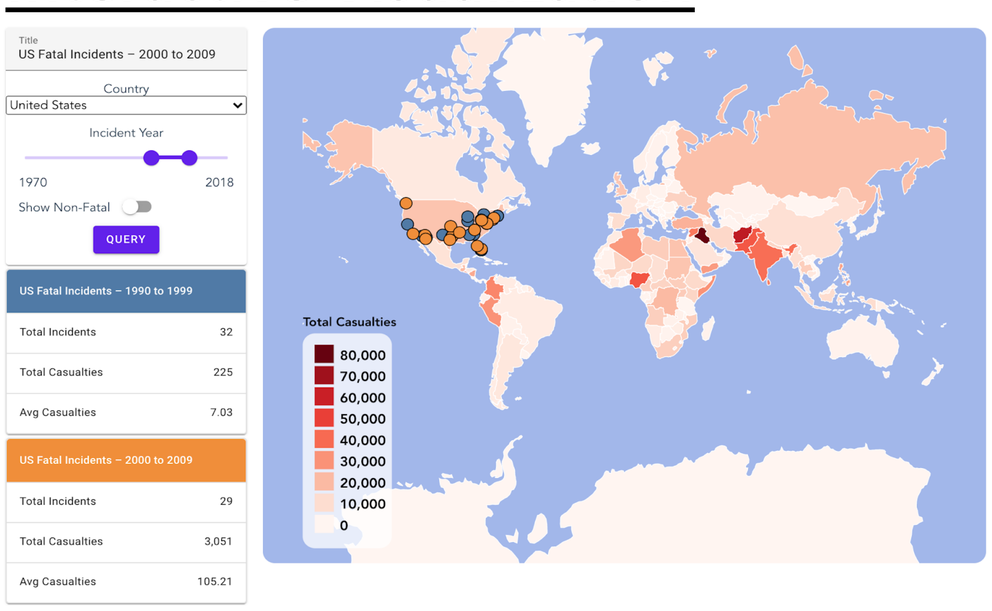
7
Variables selected

4
Models tested

.373
R-squared

- To help users better understand terrorism trends, elastic net, a combination of LASSO and ridge regression, was used to identify event features most closely linked to an attacks fatality rate.
- Elastic net performs variable selection and regularization and identified 7 features that played a large role in an attacks fatality rate. Those features are: # wounded, suicide, weapon type, success, target type, weapon subtype, and terrorism certainty.
- These 7 features are incorporated in the interactive visualization below and were used to build a regression model to predict fatalities.

Interactive Visualization



- A choropleth map of countries was created for users to interact with and explore terrorism events. The map is colored according to the number of casualties suffered by terrorism.
- Using the form in the left panel, users can create multiple queries to explore the GTD. Summarized results appear below the form and points appear on the map where individual events occurred. Only the first 50 points are shown for each query.
- Users get a closer look by zooming in (click or pinch to zoom; drag the map to pan). Hovering over an incident will show details such as the perpetrator, target, actual causalities, predicted causalities and 7 features described above as the strongest predictors of the number of casualties.