



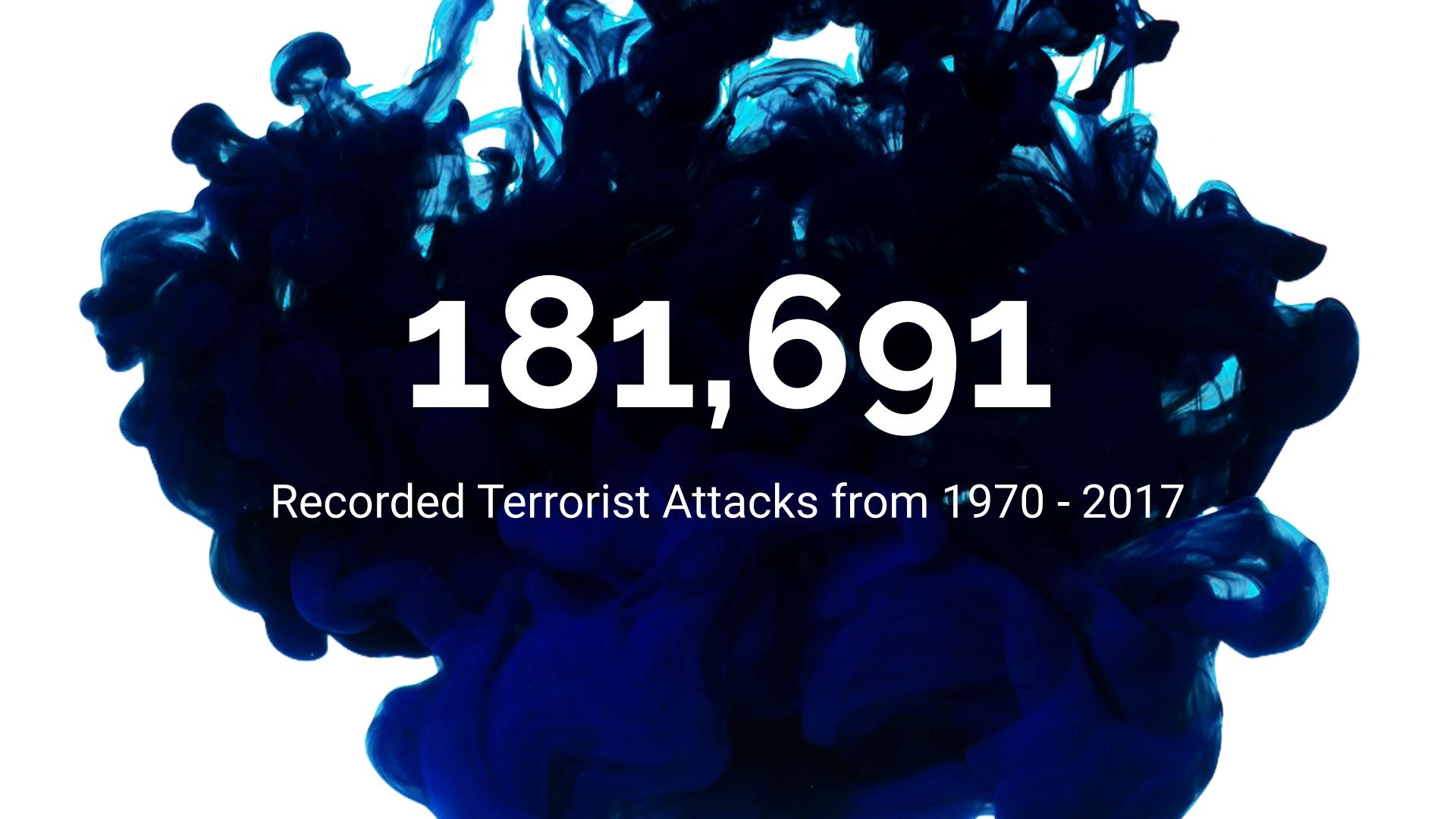
Global Terrorism: Predicting the Success of Terrorist Attacks

By Jeff Spagnola

INTRODUCTION

- Terrorism is a worldwide problem.
- Attacks have been recorded on every continent and 180+ countries.
- Many factors determine the success of a terrorist attack.
- We will determine specific factors that help determine the success of a terrorist attack.
- What is success? Whether or not any version of the attack took place.





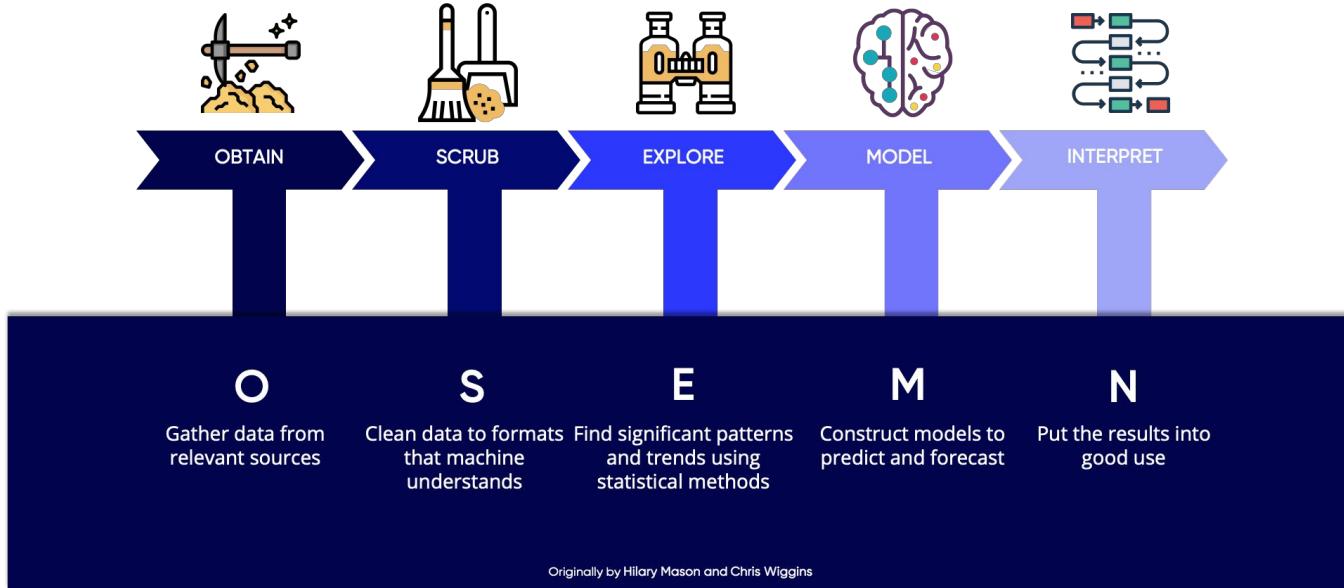
181,691

Recorded Terrorist Attacks from 1970 - 2017

PROCESS

PROCESS

Data Science Process



DATA

THE DATA



GLOBAL TERRORISM DATABASE

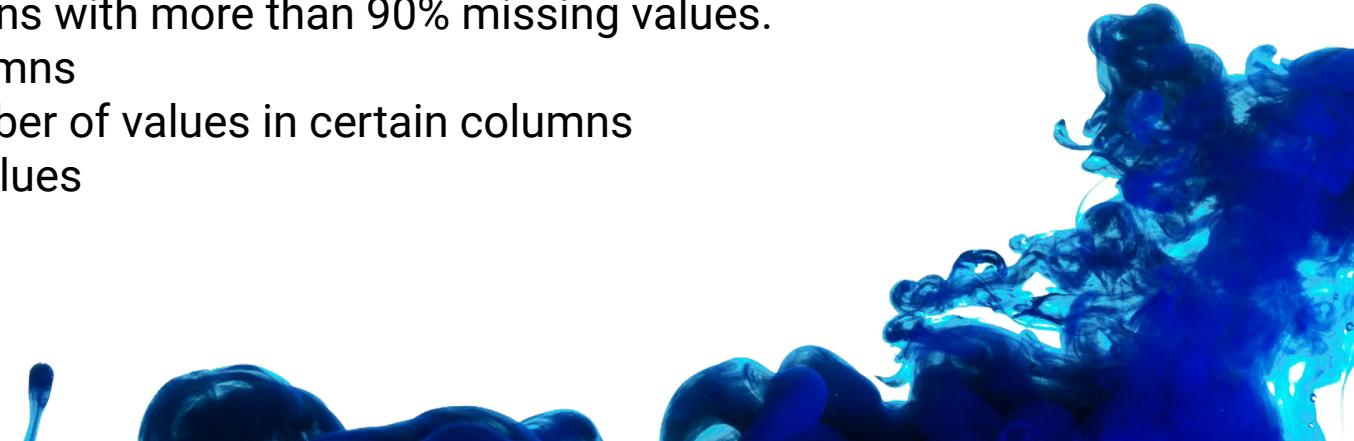
- The data is from the Global Terrorism Database.
- Recorded terrorist attacks between 1970 - 2017.
- Compiled by the National Consortium for the Study of Terrorism and Response to Terrorism.
- The dataset contains 181,691 data points and 135 features.



Scrubbing or Cleaning Data in Data Science



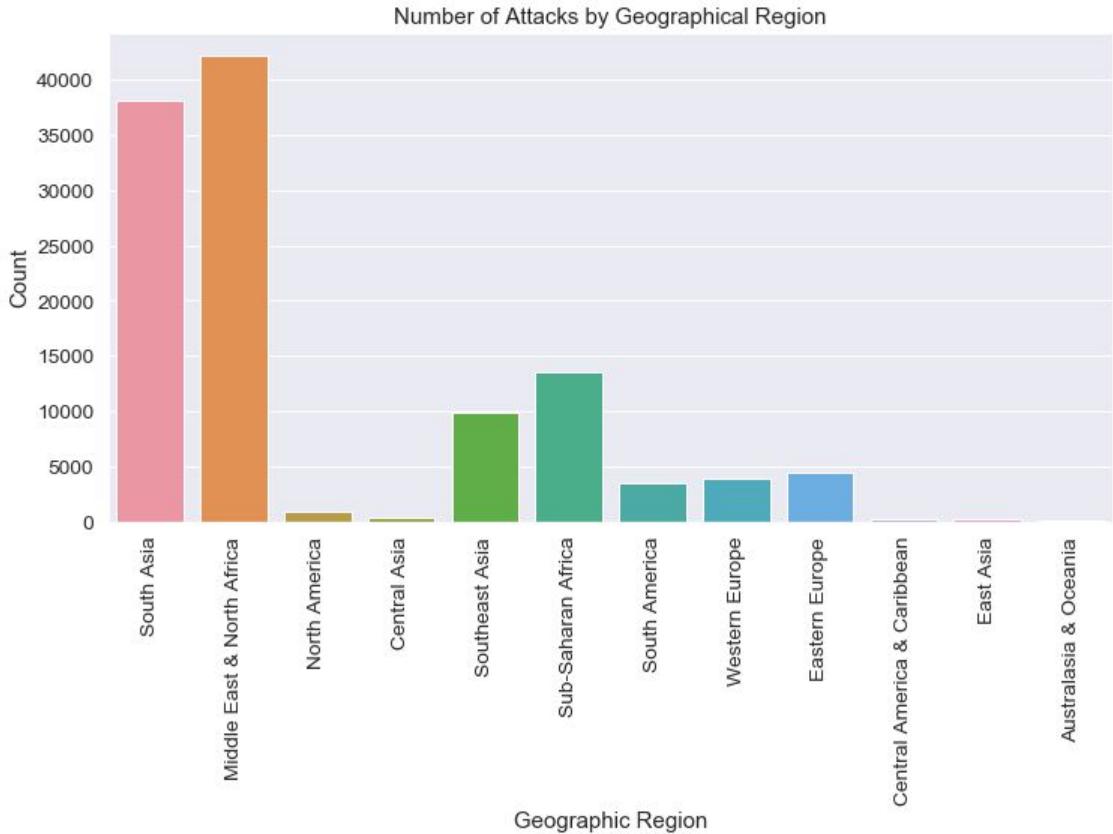
- Focused on most recent 20 years - 1997 to 2017
- Removed Irrelevant Columns
- Removed Columns with more than 90% missing values.
- Condensed columns
- Condensed number of values in certain columns
- Filled missing values
- Scaled the data



EXPLORE

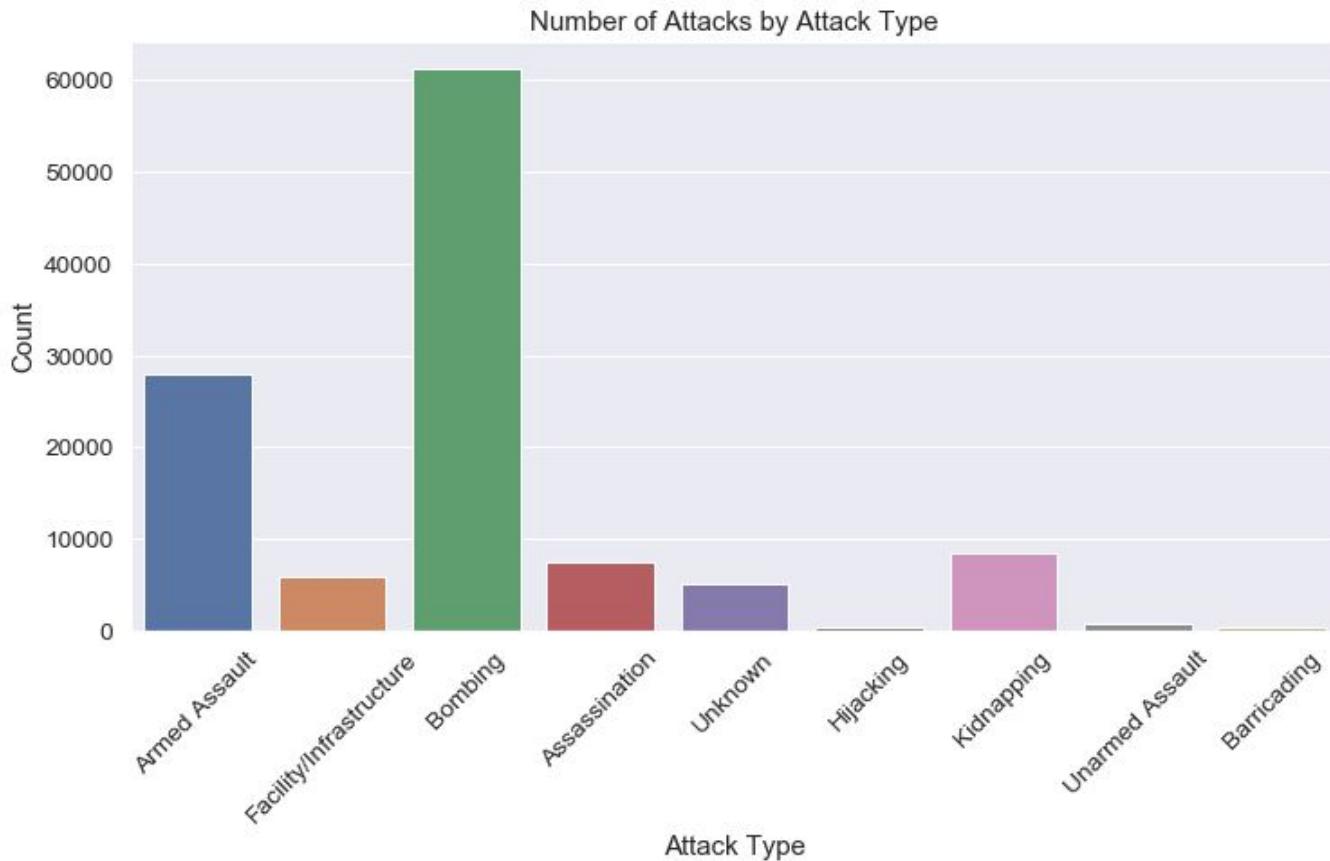
GEOGRAPHIC REGION

A majority of attacks occur in the Middle East/North Africa and South Asia.



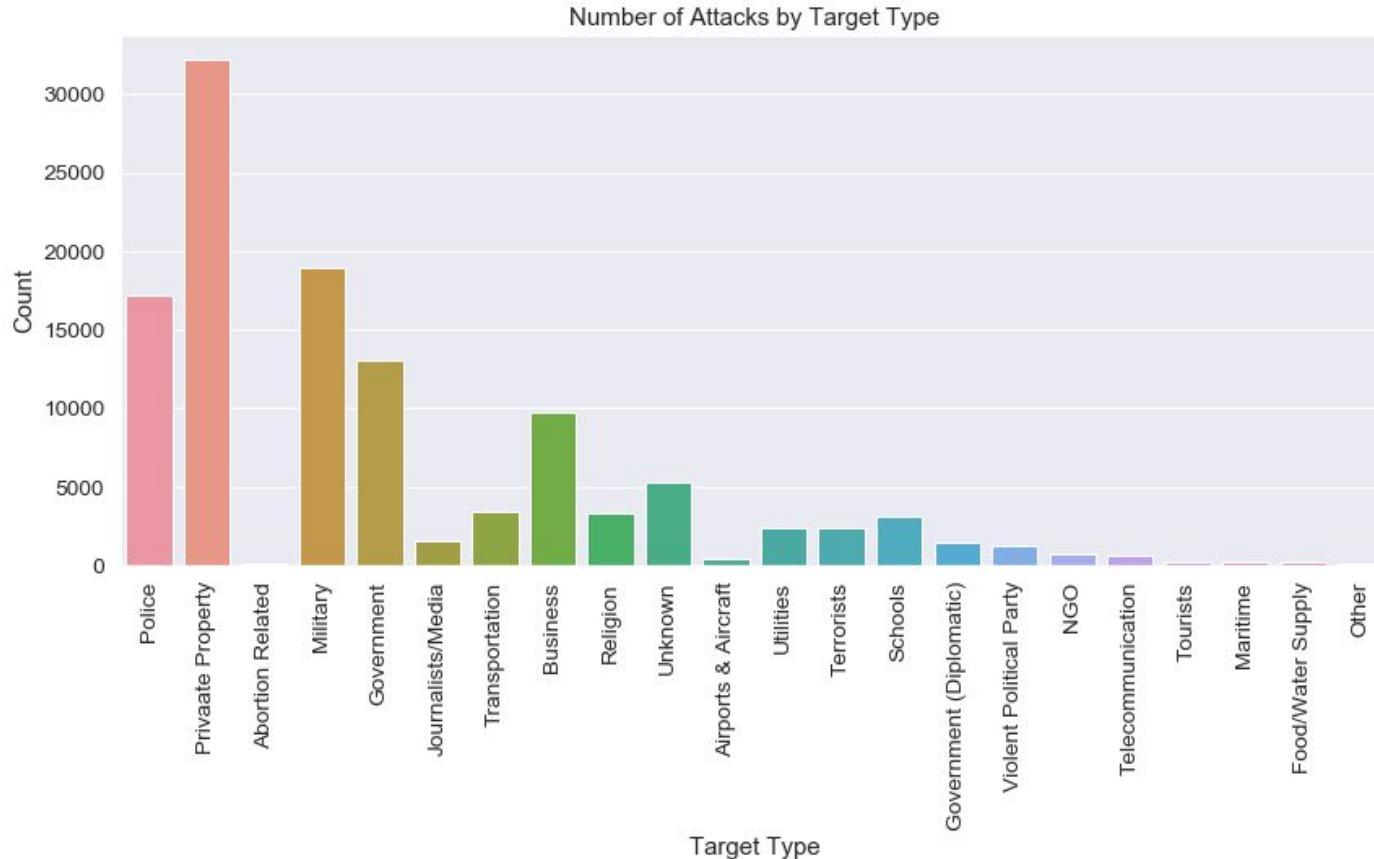
ATTACK TYPE

A majority of attacks are either bombings or armed assaults.



TARGET TYPE

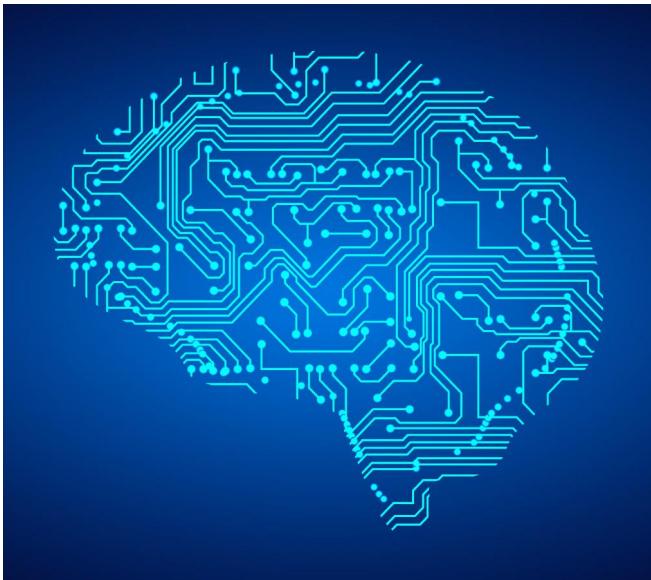
Terrorists tend to target private property, Police, Military, Government and Business.



MODELING

MODELING

We performed a variety of classifiers in order to analyze the data for the most important features.



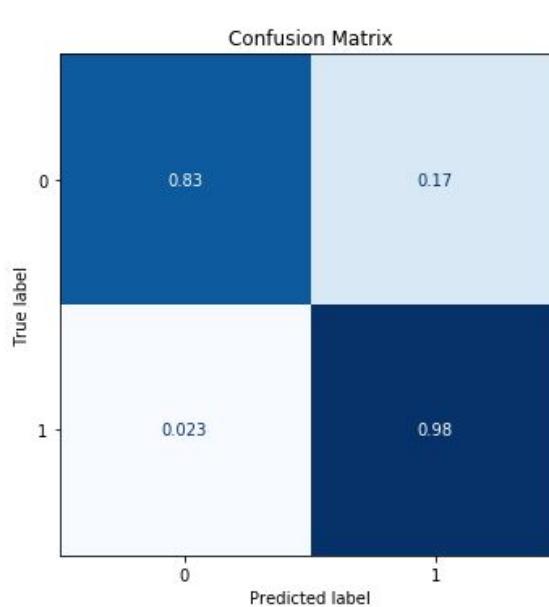
Machine Learning Algorithms Used:

- Logistic Regression
- Decision Tree
- Random Forest
- XGBoost
- Stacking Classifier

MODELING

For the issue of terrorist attacks, it is important to limit False Negatives.

Recall Score: Number of True Positives divided by the total number of positives.

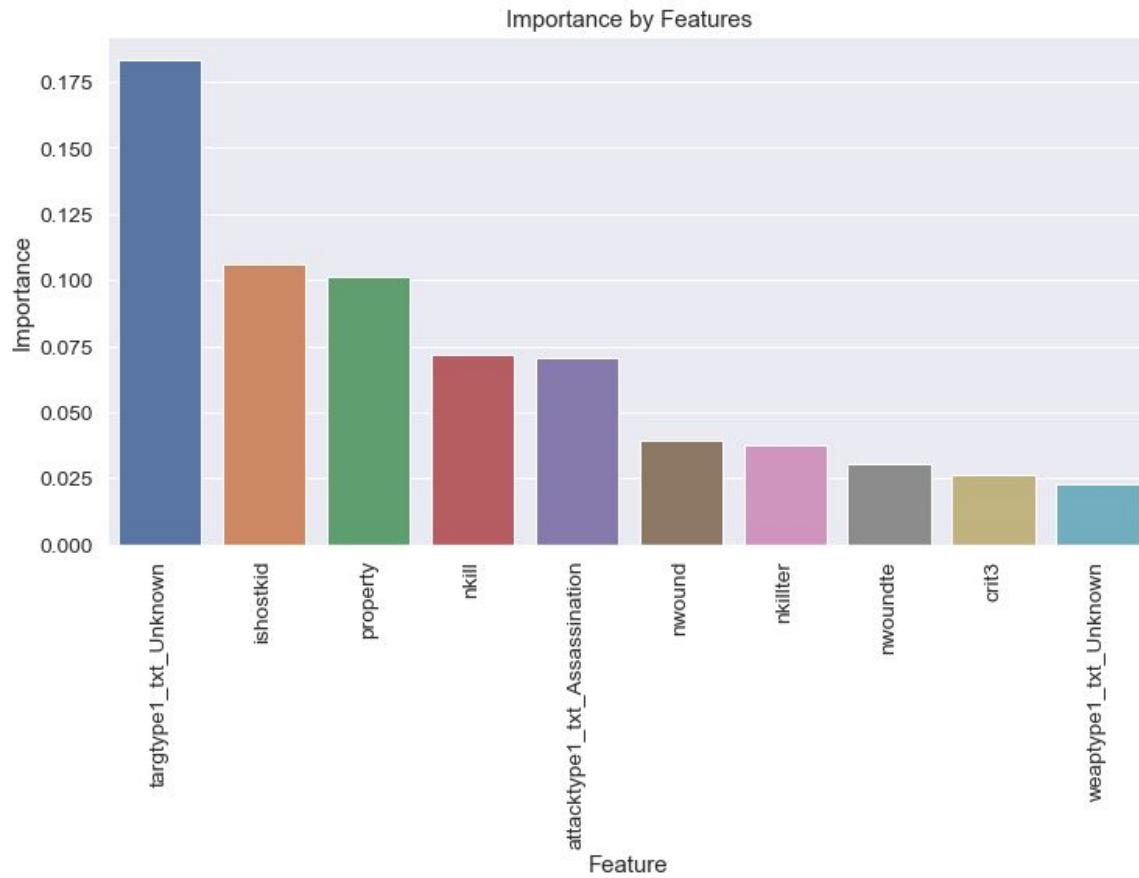


Classification Report

	precision	recall	f1-score	support
0	0.84	0.83	0.83	3628
1	0.98	0.98	0.98	25718
accuracy			0.96	29346
macro avg	0.91	0.90	0.90	29346
weighted avg	0.96	0.96	0.96	29346

RESULTS

RESULTS



Most Important Features from Model:

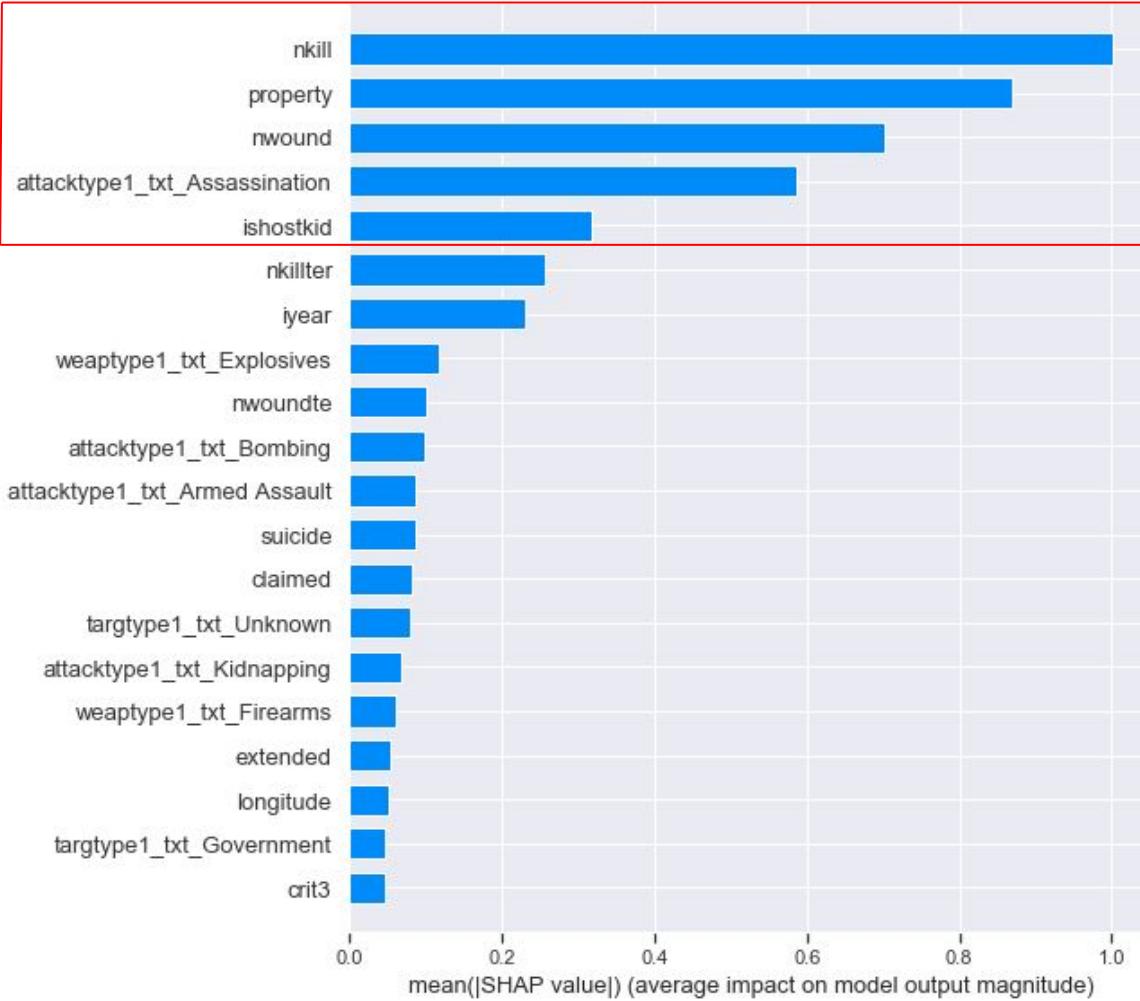
1. Unknown Target
2. Hostage/Kidnapping Victims
3. Property Damage Evidence
4. Total Number of Fatalities
5. Assassination Attempts



RESULTS

Most Important Features from SHAP:

1. Total Number of Fatalities
2. Property Damage Evidence
3. Number of Injured
4. Assassination Attempt
5. Hostage/Kidnapping Victims

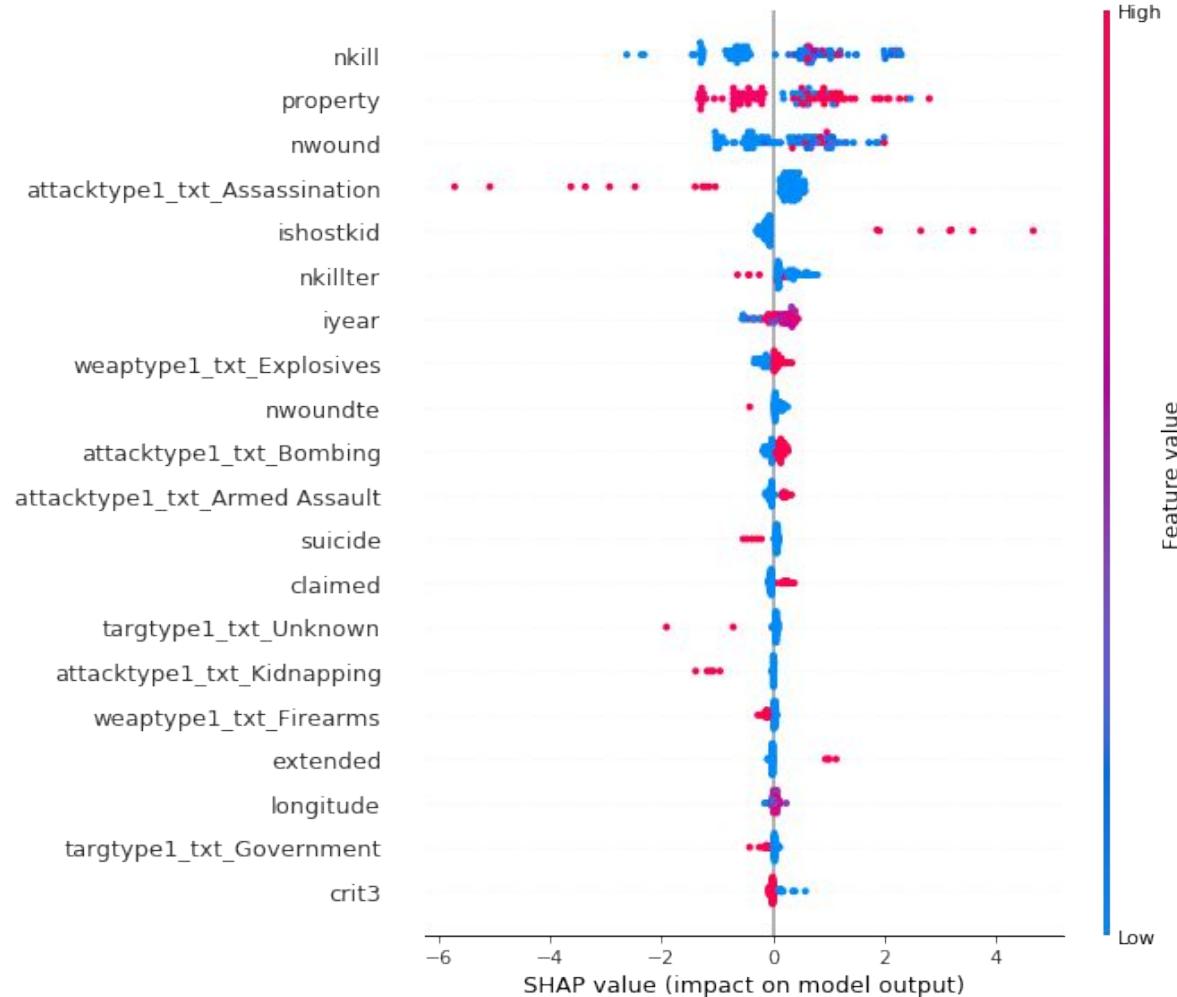


RESULTS

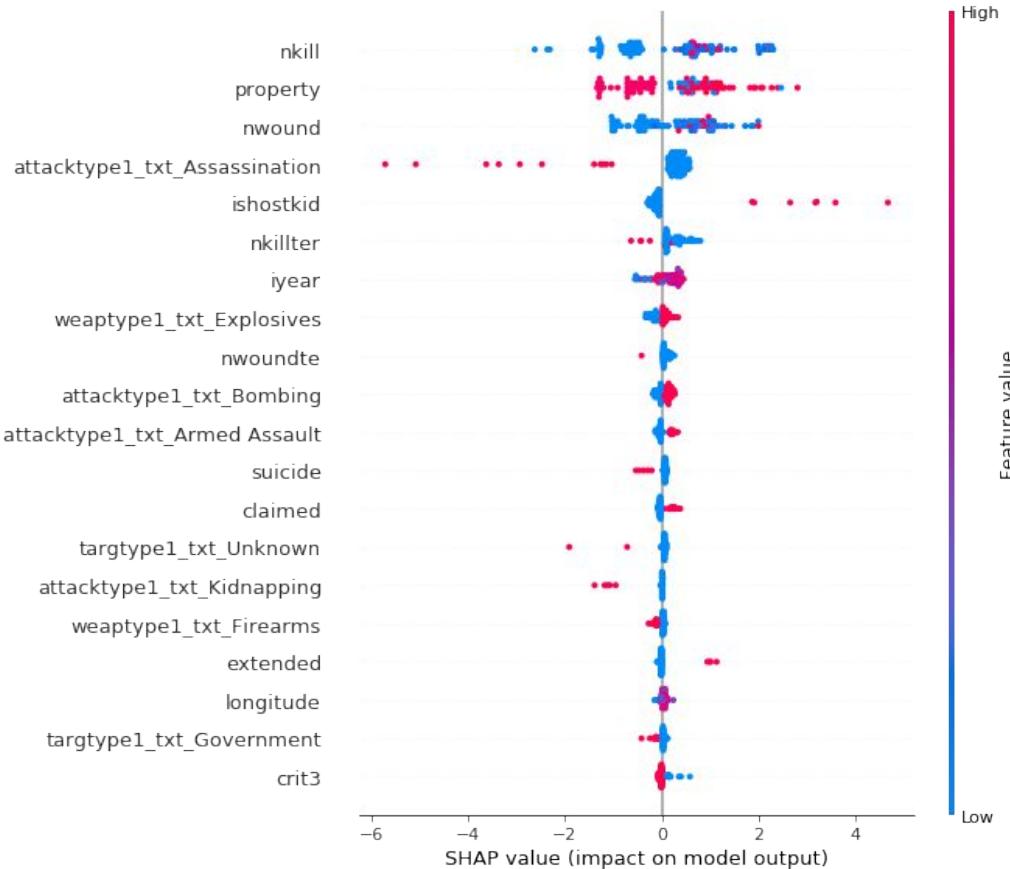
This plot indicates whether a **feature has a positive or negative effect** on the success of a terrorist attack.

Red indicates a **high value** and **blue** indicates a **low value**.

Left of center indicates **negative effect**, **Right** of center indicates **positive effect**.



RESULTS



Most Important Features from Modeling:

1. **Total Number of Fatalities**: High counts have positive effect, low counts have negative effect.
2. **Hostage/Kidnapping** - Very positive effect on success
3. **Assassination Attempt** - Very negative effect on success.
4. **Property Damage** - Evidence of damage positively affects success.

RESULTS

Most Important Insights from Exploration:

1. Majority of attacks occur in Middle East/North Africa and South Asia.
2. Majority of terrorist attacks are either bombings or armed assaults.



RECOMMENDATIONS

RECOMMENDATIONS

Based on this analysis, we can make the following recommendations:

- Maintain high levels of intel and security in the Middle East and South Asia.
- Develop better methods and/or technology for bomb detection and disarmament.
- Focus intel on target areas that have the highest concentration of people.
- Increase security for high level targets for potential kidnapping or hostage situations.

FUTURE WORK

With more time, we can gain even more insight into what can make a terrorist attack successful.

Time

We can increase the range of years of the data in our analysis. For example, 1970-2017.

Models

We can increase the size and complexity of our models to increase accuracy.

Data

We can research and compile additional data from other sources for a more well rounded dataset.



THANK YOU

Do you have any questions?

jeff.spags@gmail.com

<https://github.com/spags093>

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, infographics & images by Freepik