

# IS IT ISIS?

A Presentation by Terror Inc.





# Defining Terrorism with Three Criteria

- 01 Aimed at attaining a political, economic, religious, or social goal
- 02 Intention to coerce, intimidate, or convey a message to a larger audience
- 03 Outside the context of legitimate warfare activities

# Data Sources



- 170,000 attacks
- Since 1970
- 139 potential features
- Clean!



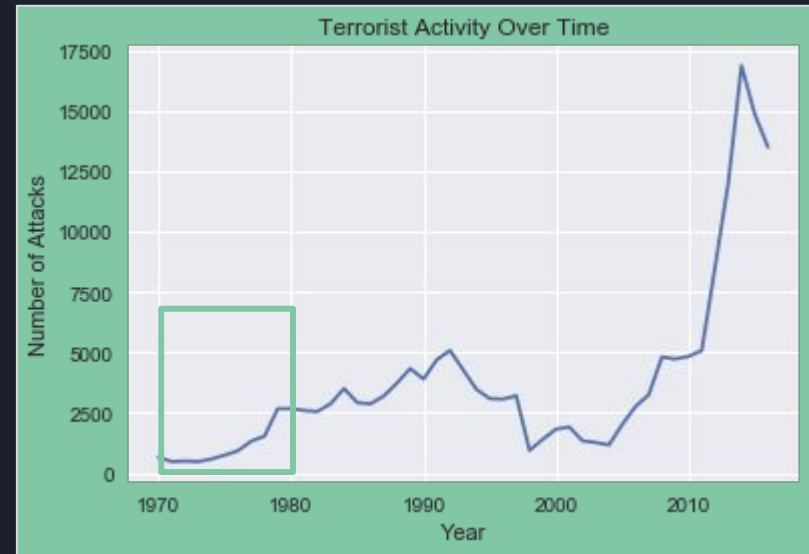
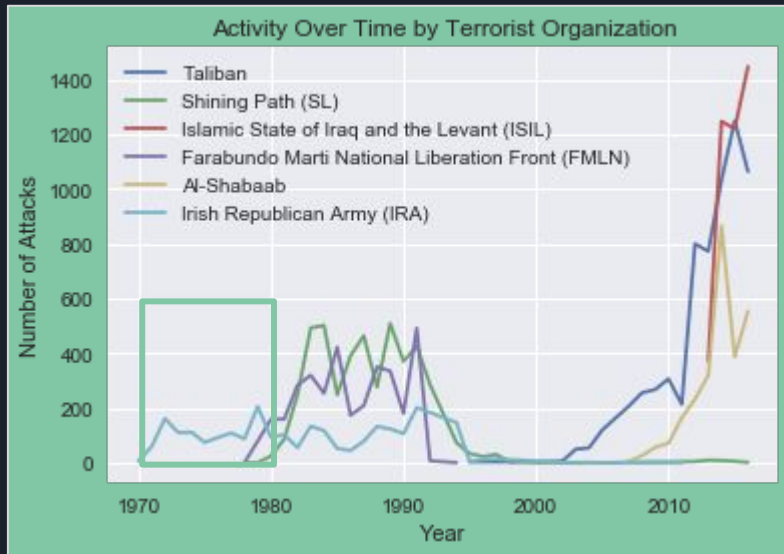
- Country level socio-economic data
- Monthly frequency



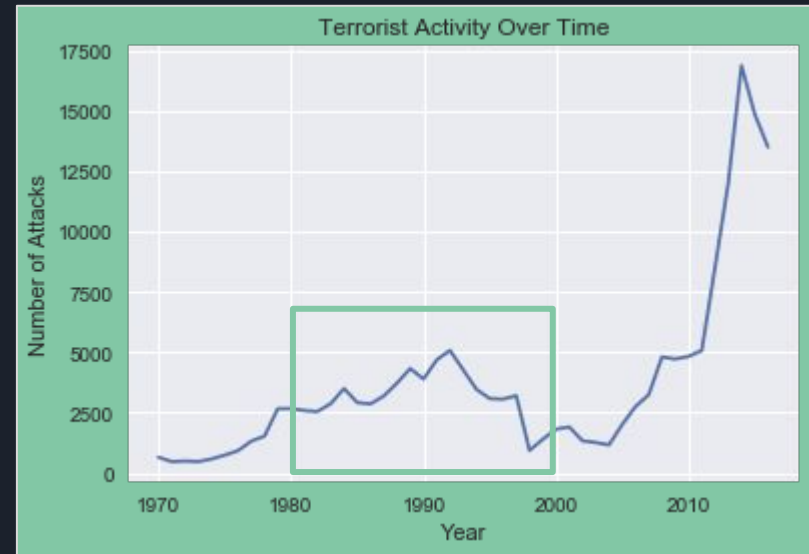
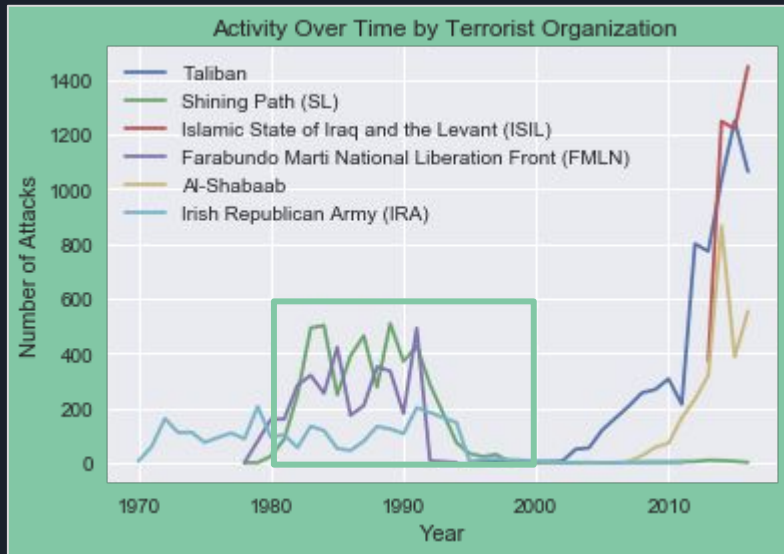
# First Question

Given the parameters available shortly after an attack can we predict the name of the responsible group?

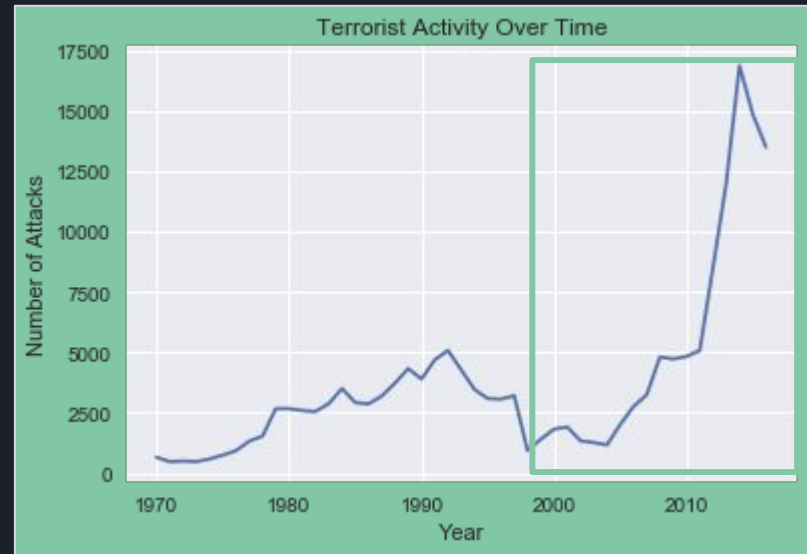
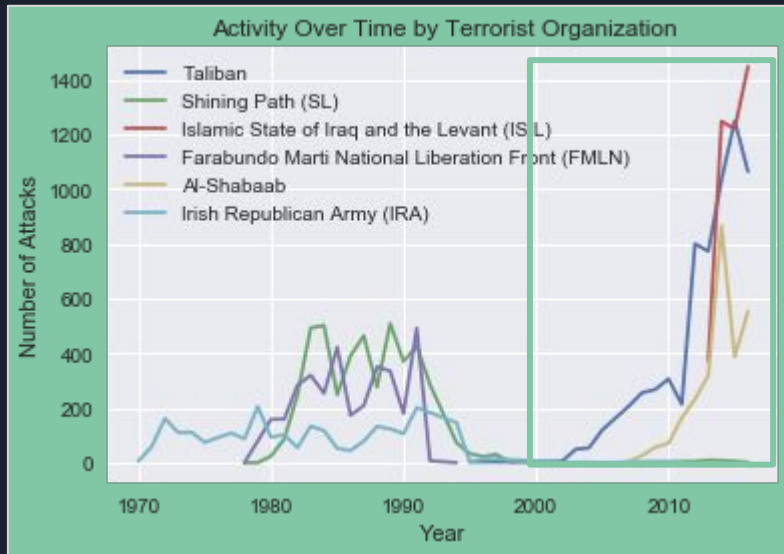
# Historical Context - Segmenting the Problem



# Historical Context - Segmenting the Problem



# Historical Context - Segmenting the Problem





# Top 10 Feature Importance

## Continuous Features:

- Year
- Month
- # Dead
- # Dead US
- # Dead Terrorists
- # Wounded
- # Wounded Terrorists

## Dummy Features

- Multiple Attack
- Success
- Suicide
- Hostages
- Property Damage

## Categorical Features

- Country
- City
- Attack Type
- Weapon Type
- Target Type
- Victim Nationality



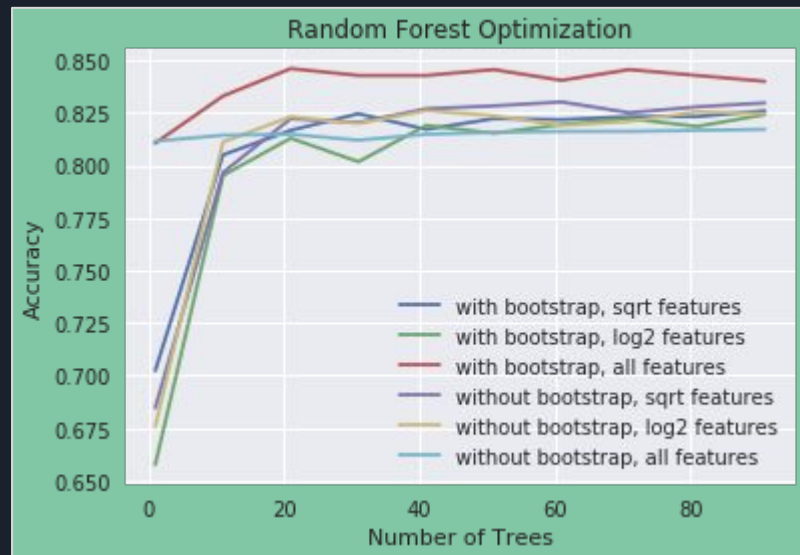
# Model Selection and Tuning

	Random Forest	Logistic Regression	Naive Bayes
Accuracy	84%	78%	57%
Precision	83%	77%	37%

Dataset: 8761 Rows x 4437 Columns (w/ dummies)

117 Labels (very unbalanced)

Baseline (always Taliban): 26%





# Top 10 Feature Importance

## Continuous Features:

- Year
- Month
- # Dead
- # Dead US
- # Dead Terrorists
- # Wounded
- # Wounded Terrorists

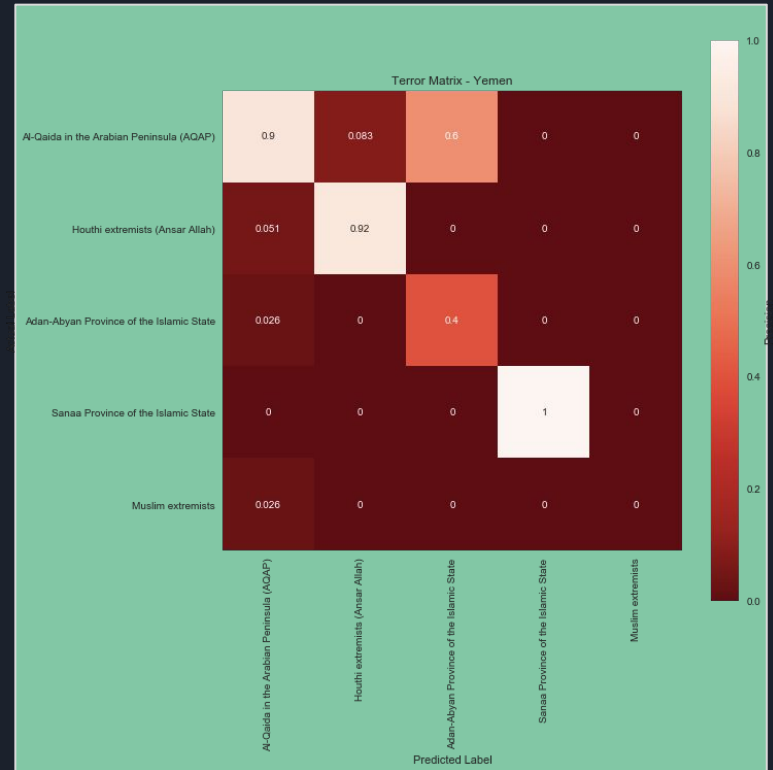
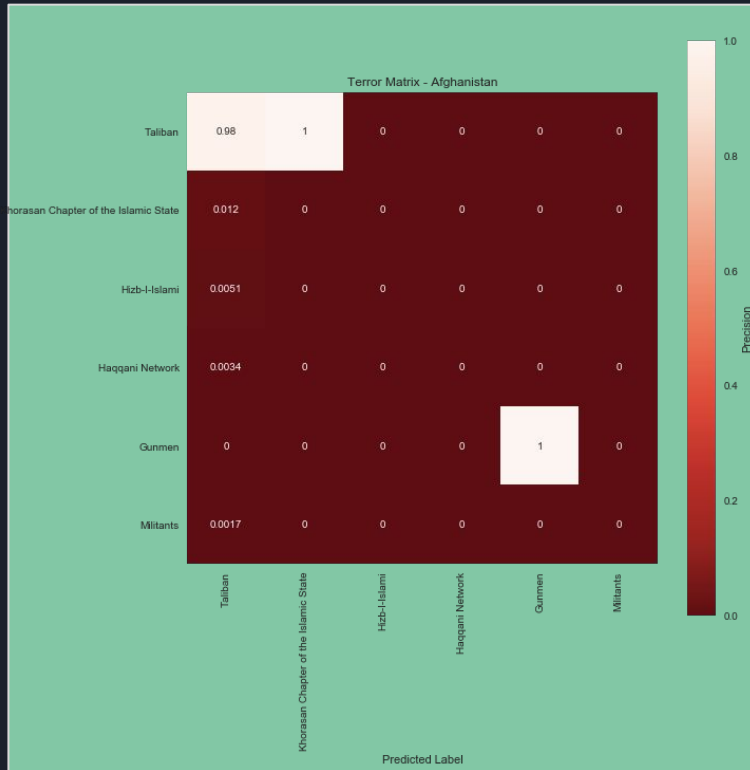
## Dummy Features

- Multiple Attack
- Success
- Suicide
- Hostages
- Property Damage

## Categorical Features

- Country
- City
- Attack Type
- Weapon Type
- Target Type
- Victim Nationality

# Terror Matrices



# 2017 - Validation

	<b>gname</b>	<b>country_txt</b>	<b>city</b>	<b>nkill</b>	<b>attacktype1_txt</b>	<b>Prediction</b>
<b>0</b>	ISIL	Syria	Damascus	114	Bombing/Explosion	Islamic State of Iraq and the Levant (ISIL)
<b>1</b>	Taliban	Afghanistan	Mazar-i-Sharif	140	Armed Assault	Taliban
<b>2</b>	Misrata Militants	Libya	Brak al-Shati	141	Armed Assault	Tripoli Province of the Islamic State
<b>3</b>	ISIL (suspected)	Egypt	Al-Rawda	311	Armed Assault	Sinai Province of the Islamic State
<b>4</b>	Al-Shabab (suspected)	Somalia	Mogadishu	512	Bombing/Explosion	Al-Shabaab
<b>5</b>	Unknown	United States	New York City	50	Bombing/Explosion	Boko Haram
<b>6</b>	Unknown	France	Paris	50	Bombing/Explosion	Islamic State of Iraq and the Levant (ISIL)

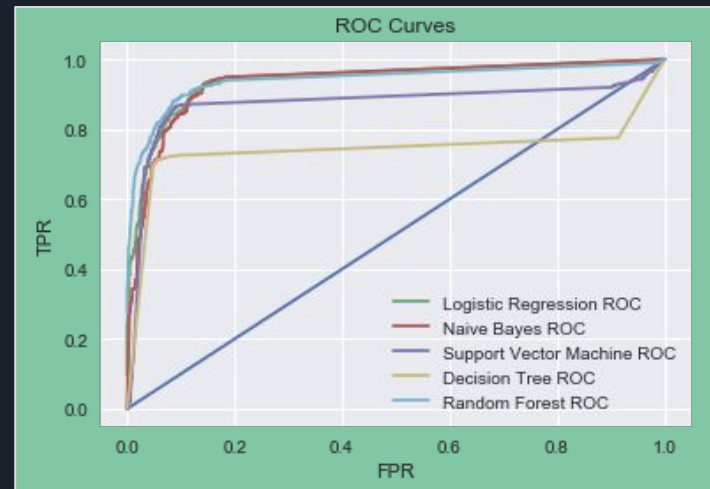


## Second Question

Given a country's history can we predict an attack in the next 30 days?

# Model Selection

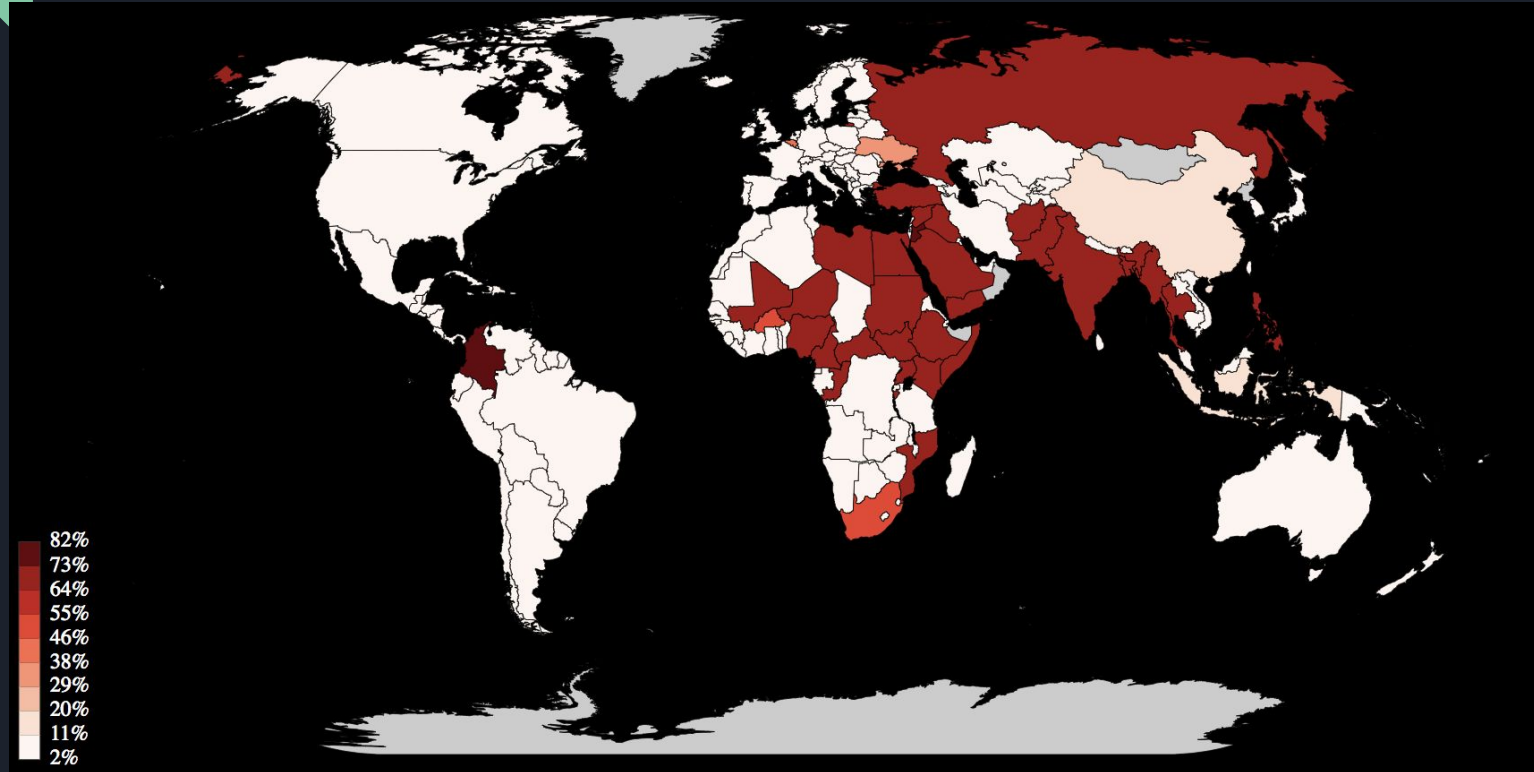
	Support Vector Machine	Random Forest	Logistic Regression	Naive Bayes
Accuracy	94%	93%	90%	88%
Yes-Recall	80%	74%	47%	35%



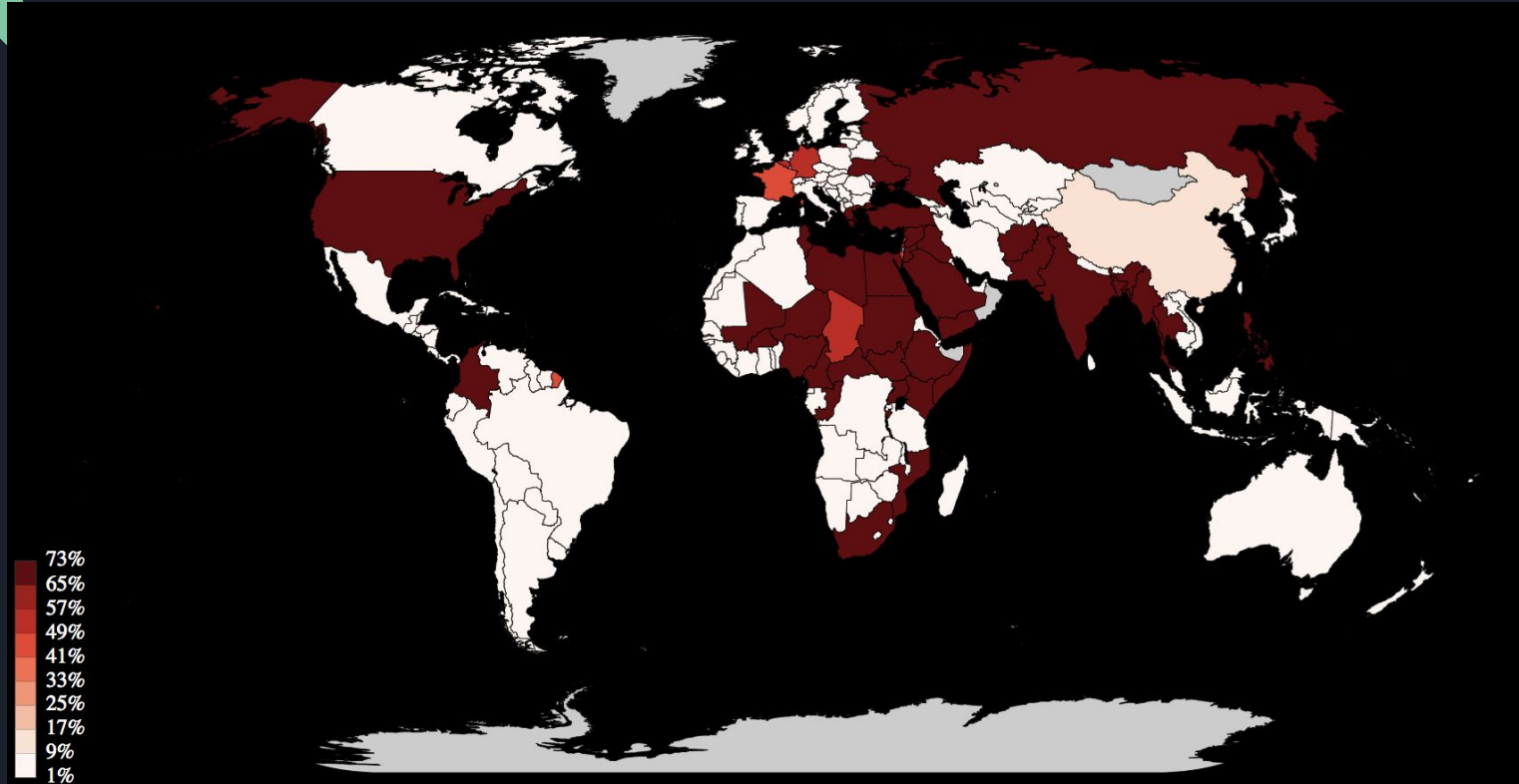
Features: 12 lags of Number of Deaths

Baseline (always No): 83%

# Predicted Probabilities Map



# Predicted Probabilities Map - (w/ Individuals)







# Conclusions

- Sadly, terrorism is relatively predictable in countries it occurs the most (deadly vicious cycle)
- Focusing on precision, we sacrifice recall on smaller classes
- Very hard to predict sporadic attacks in the West
- Very hard to predict turnarounds
- In the US, much higher likelihood of an attack by unaffiliated individuals than organized terrorist organizations

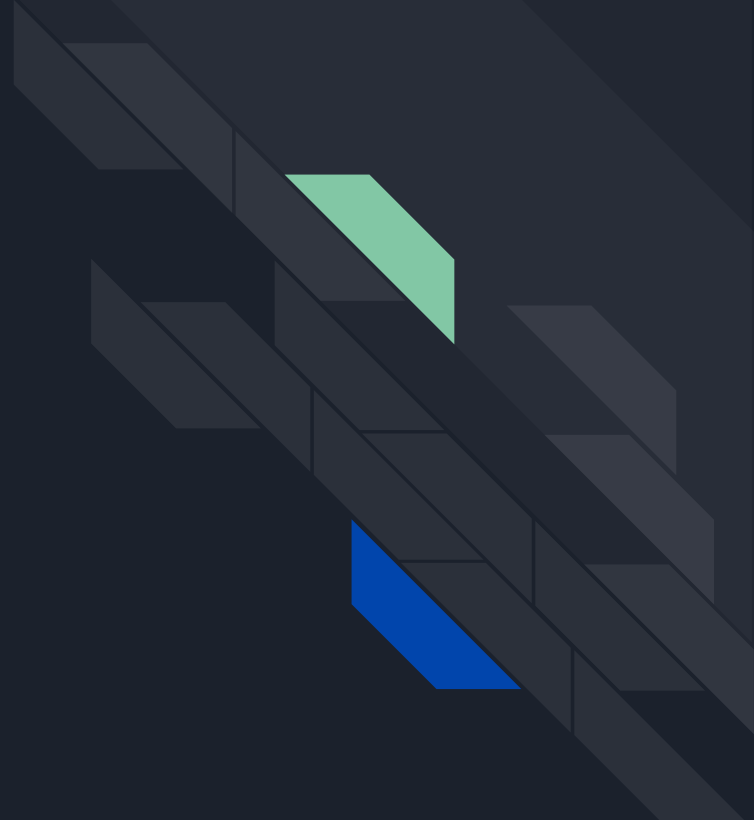


## Next Steps

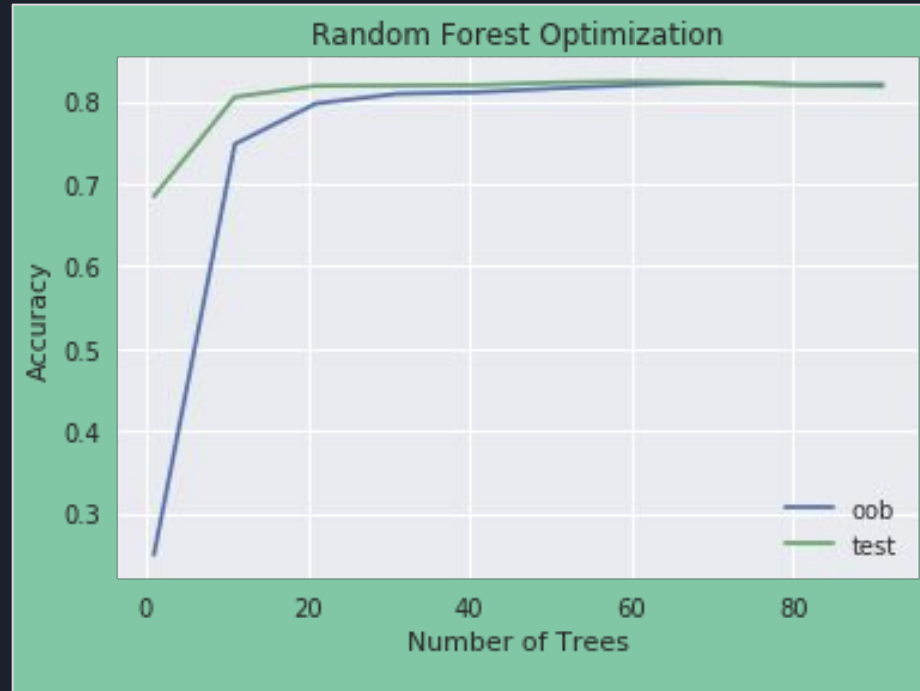
- Get more complete country-level socio-economic data
- Consider my first question from a time series perspective
- Look at the link between foreign intervention and terrorism
- Obtain 2017 data!



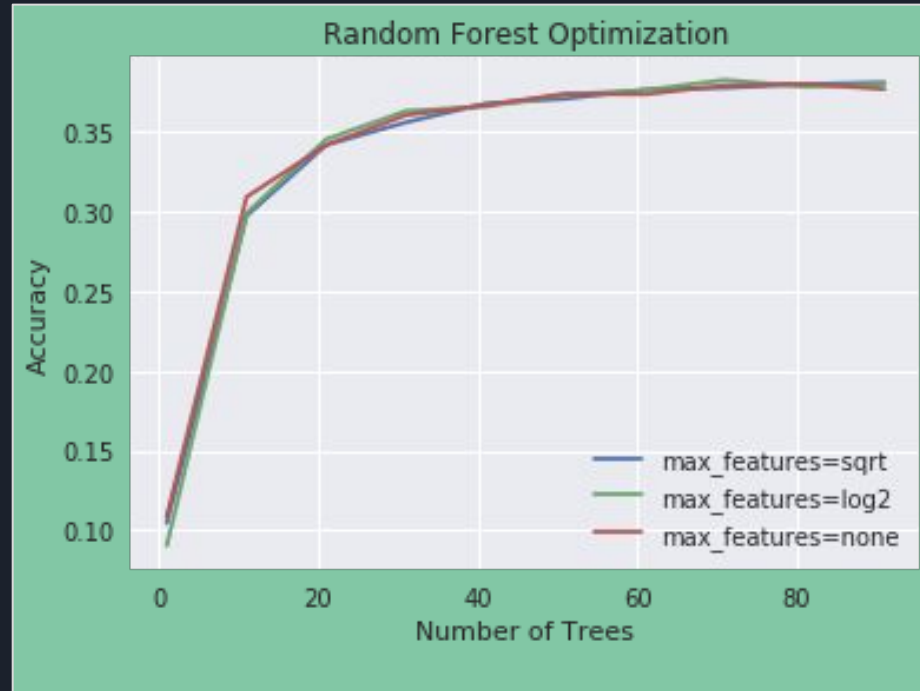
Thank you!  
Q&A



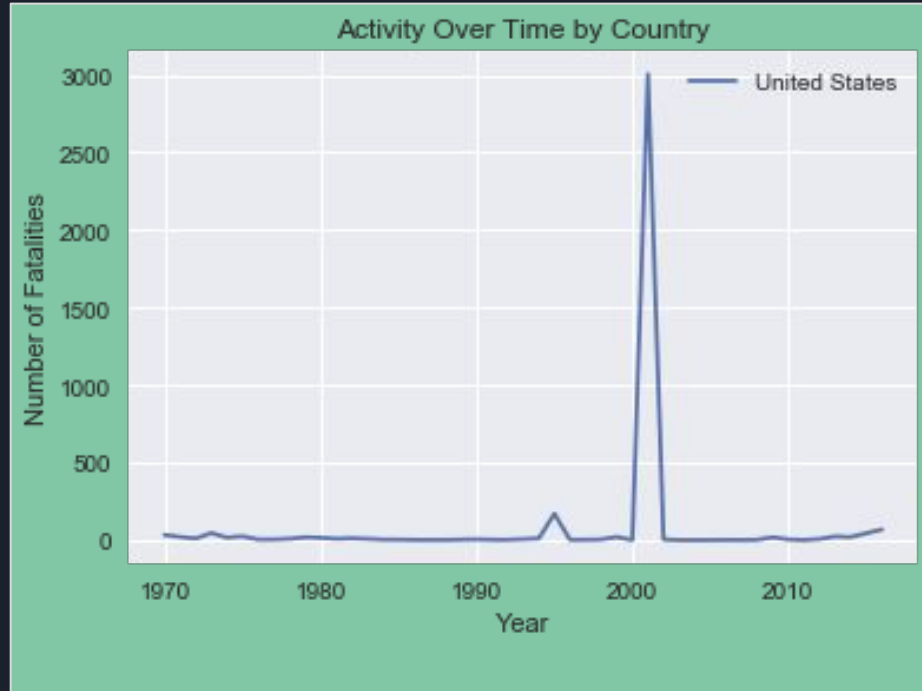
# Appendix 1 - Testing Out-Of-Bag Observations



## Appendix 4 - RF Excluding Categoricals



## Appendix 3 - US Terrorism History



# Appendix 4 - Resampled Heatmap

