Predicting protest outcomes

By: Aaron Hume, Ignacio Real, Jenn Williamson, Julia Christensen
Aka: Boo-Leans

Problem Statement

"In the short term, protests can work to the degree that they can scare authorities into changing their behavior."
- The Atlantic

The aim of this project was:
- To better understand which protests were successful based on whether or not a protest received accommodations
- To create a predictive model for classifying a successful protest

Data from: https://massmobilization.github.io
World wide protests from 1990-2019
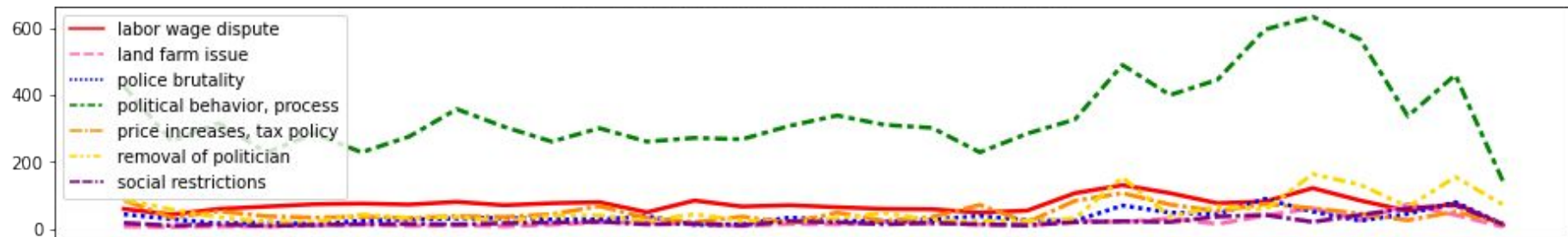
Methodology

- Analyze entire dataset

- Analyze regions

- Create models by region

- Use best models to look at entire dataset (Next Steps)
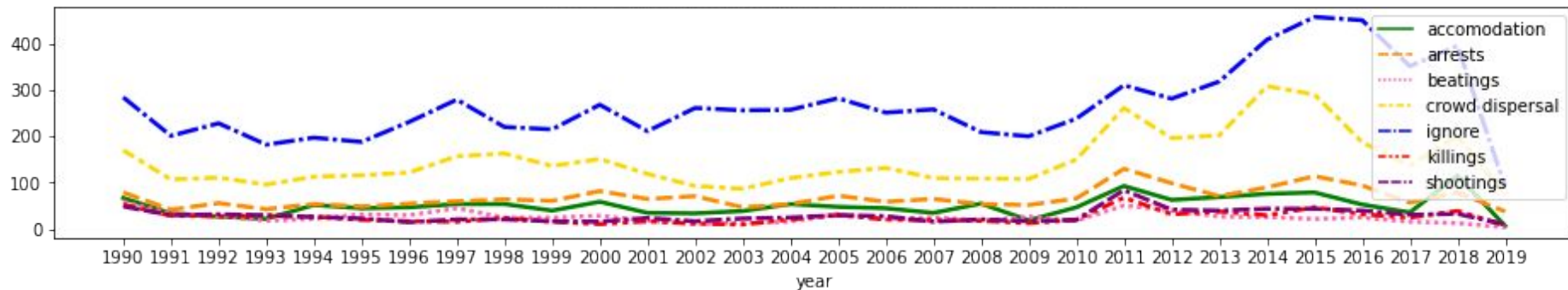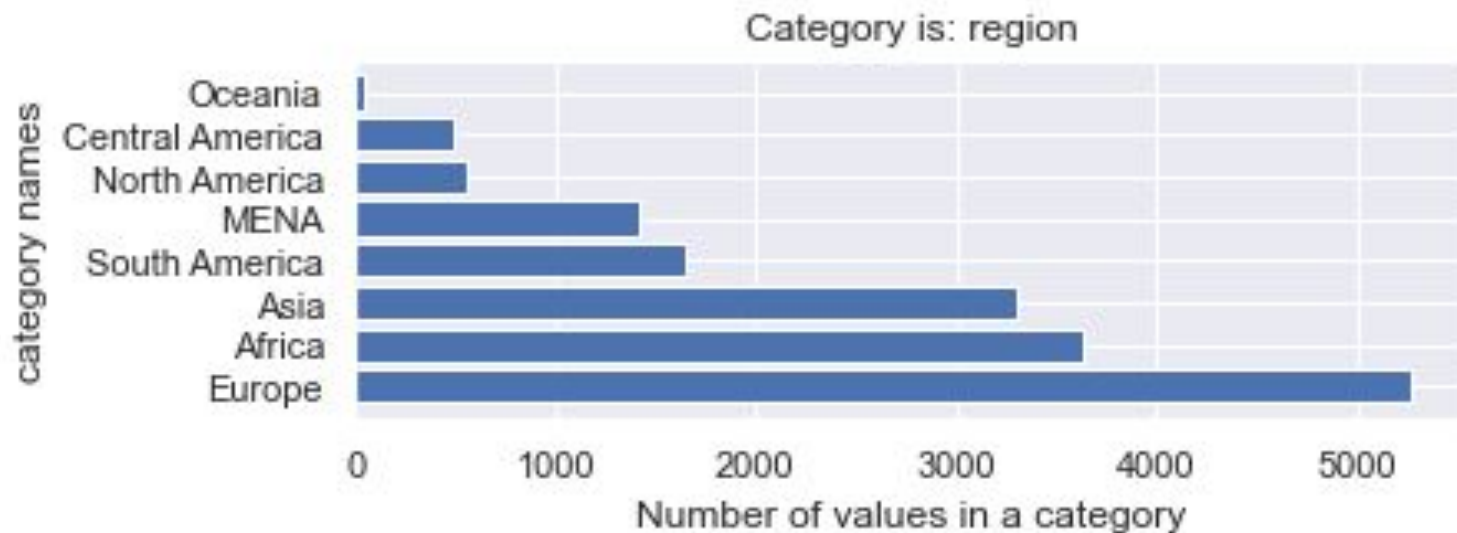
Worldwide Analysis

Protests over Time – Worldwide
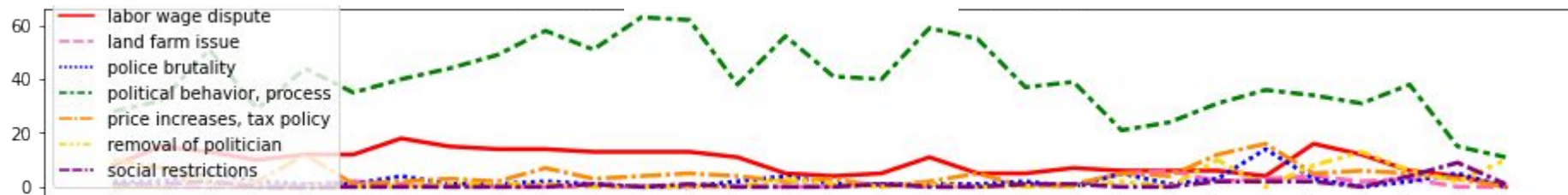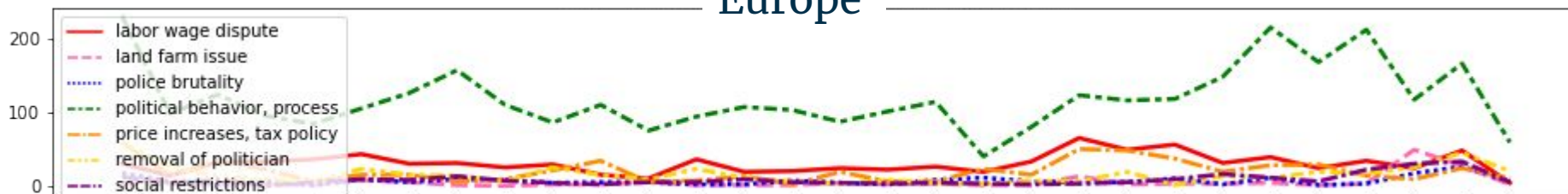
Protests per Region

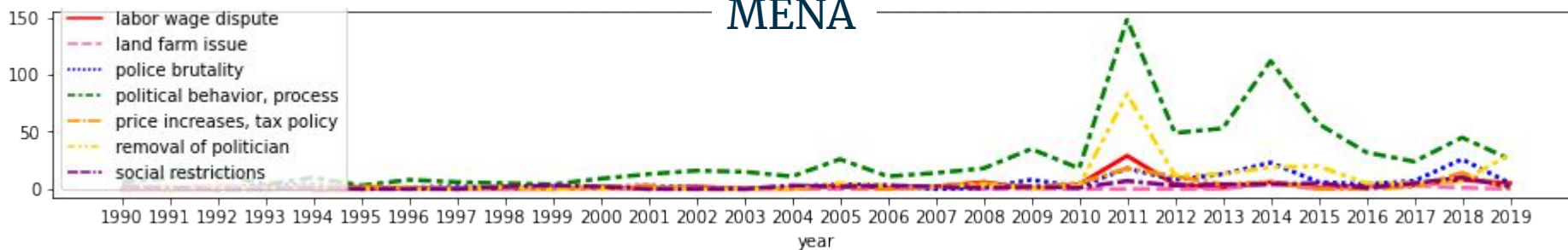# Regional Analysis
– General
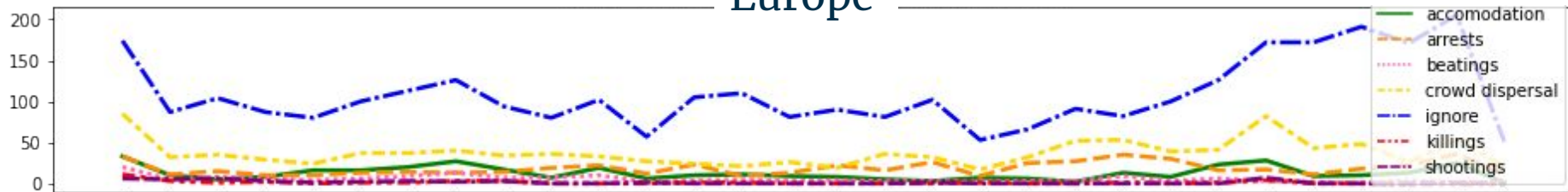
Demands by Region

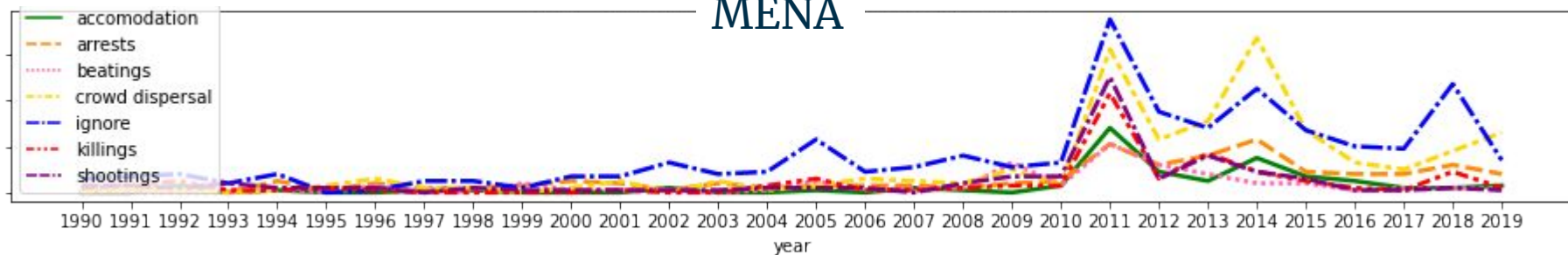South America

Europe

MENA

Responses by Region

South America

Europe

MENA

Regional Analysis – South America
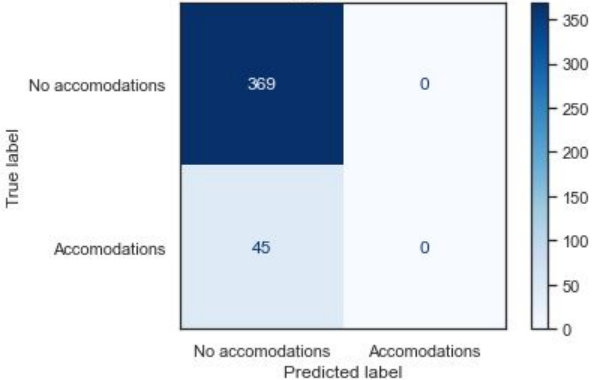
Protest Trends in South America

State Response Counts

(bar chart showing Count on y-axis)

Category	Count
ignore	~880
crowd dispersal	~405
accomodation	~120
arrests	~115
999	~55
killings	~35
shootings	~10
beatings	~5
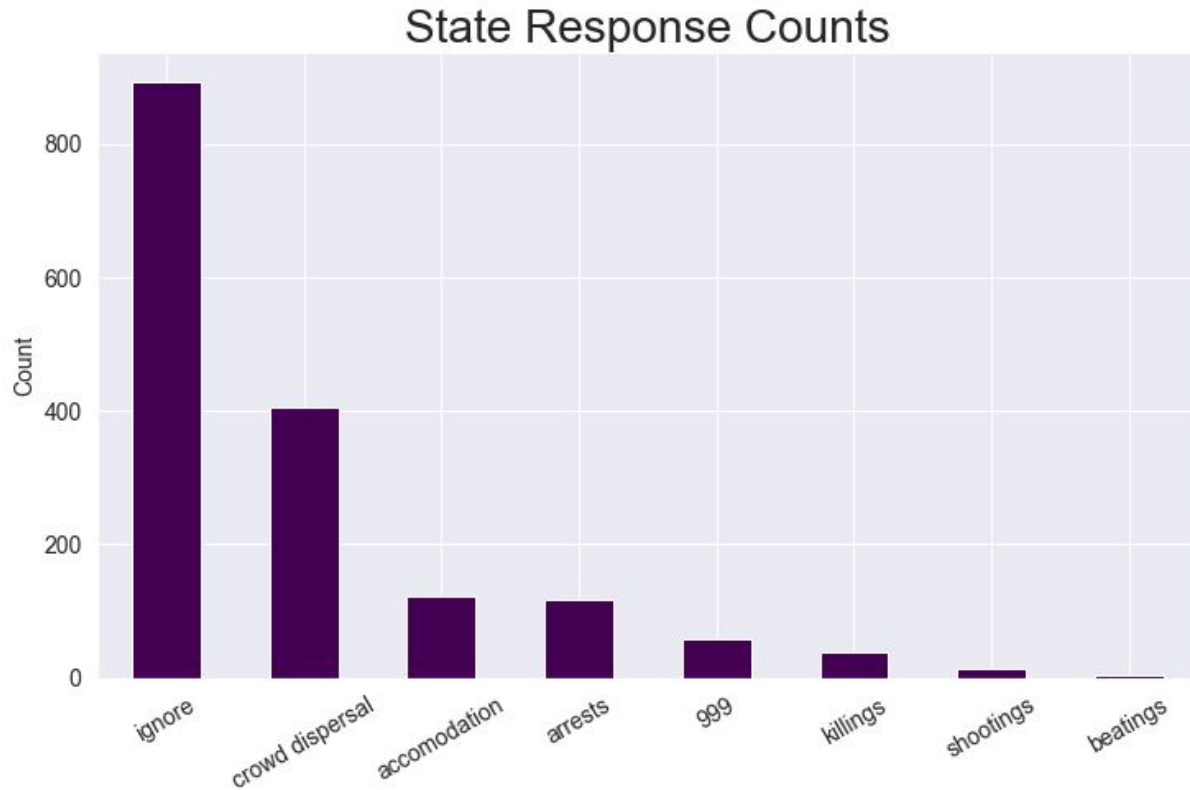
Protest Trends in South America

Support Vector Machine	Model training data	Model testing data
Accuracy	0.9015	0.8913

Grid Search: Support Vector Machine	Model training data	Model testing data
Accuracy	0.8918	0.8913

Confusion matrix of predicted accomodations vs actual in Support Vector Machine

Multiclass Logistic Regression	Model training data	Model testing data
Accuracy	0.7062	0.6594

Regional Analysis – Middle East & North Africa

Protest Trends in Middle East & North Africa

Average Protest Amount by Country

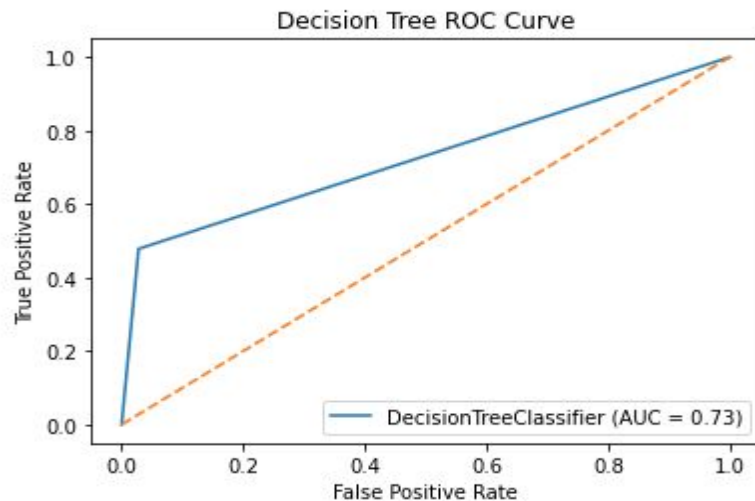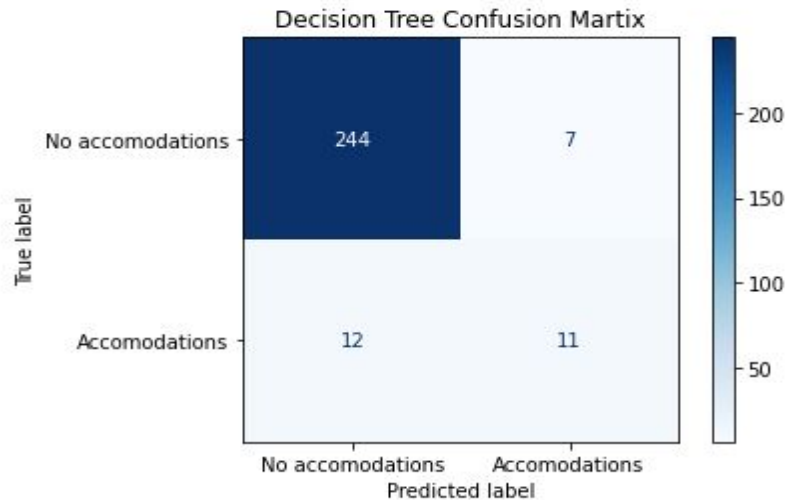**Average Protest Duration by Country**
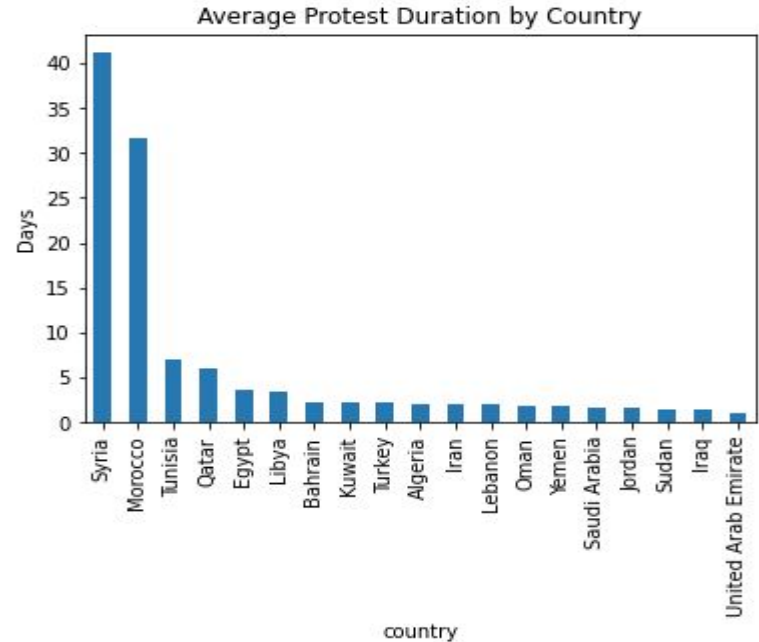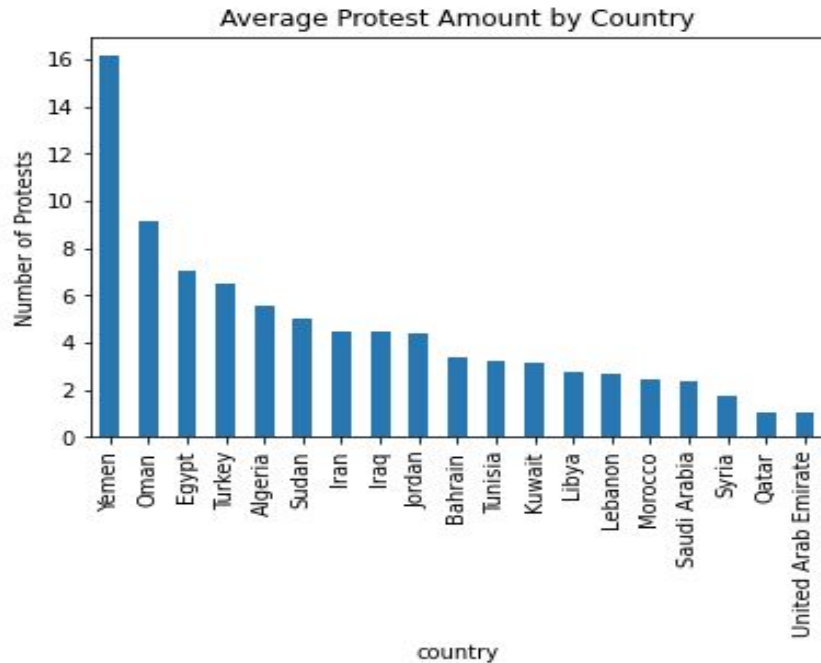
- Nations affected by the Arab Spring stand out the most but in different ways

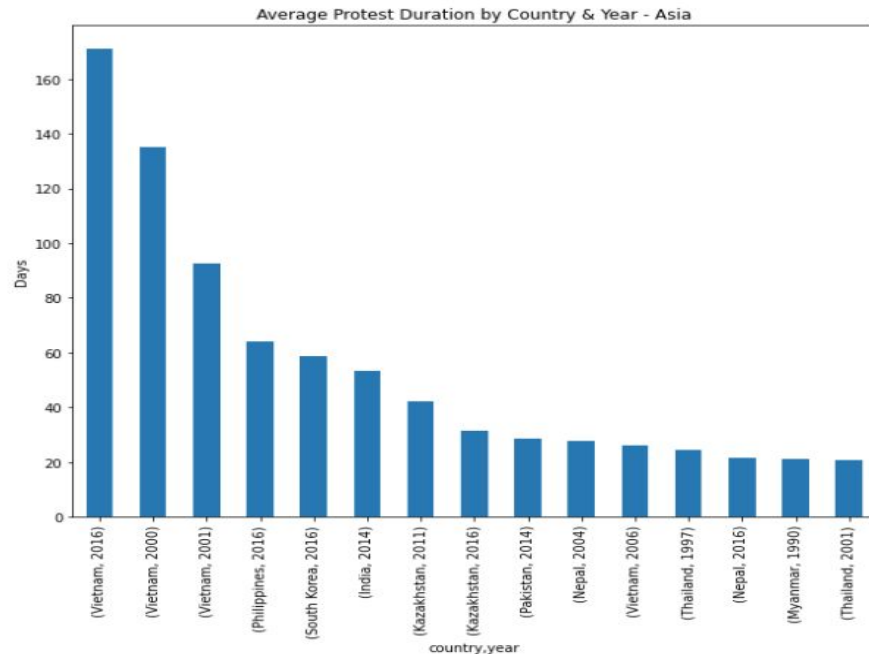# Models & Metrics: Middle East & North Africa

- Our decision tree was slightly overfit: Train 97% – Test 93% Baseline of 91.6%
- In spite of this, it had a decent performance at distinguishing extremely unbalanced classes

Decision Tree Confusion Martix

Decision Tree ROC Curve

Regional Analysis – Asia

Protest Trends in Asia: Year by Year

Average Protest Amount by Country & Year - Asia

Average Protest Duration by Country & Year - Asia
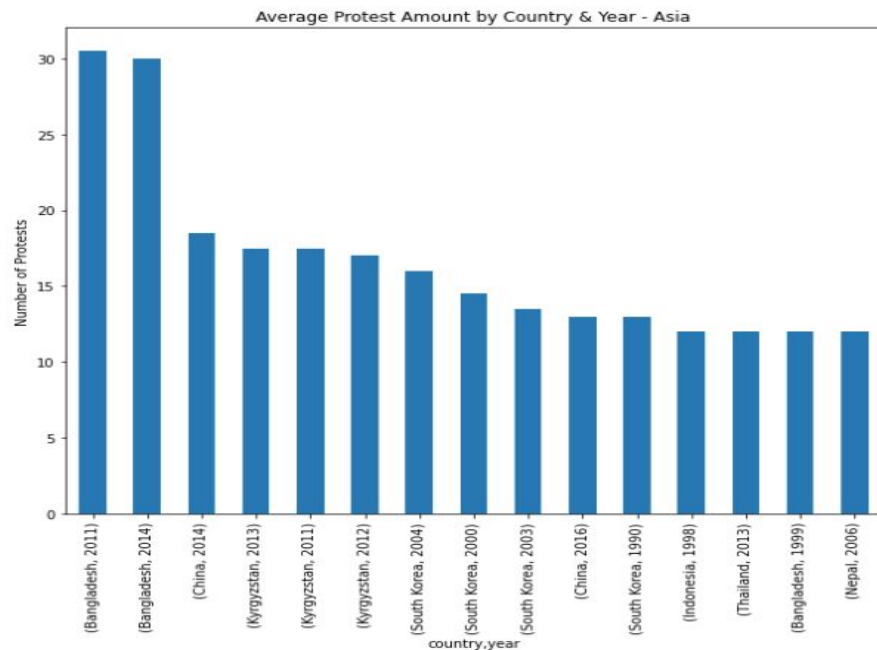
- Grouping by year reveals another pattern when considering Asia

Models & Metrics: Asia

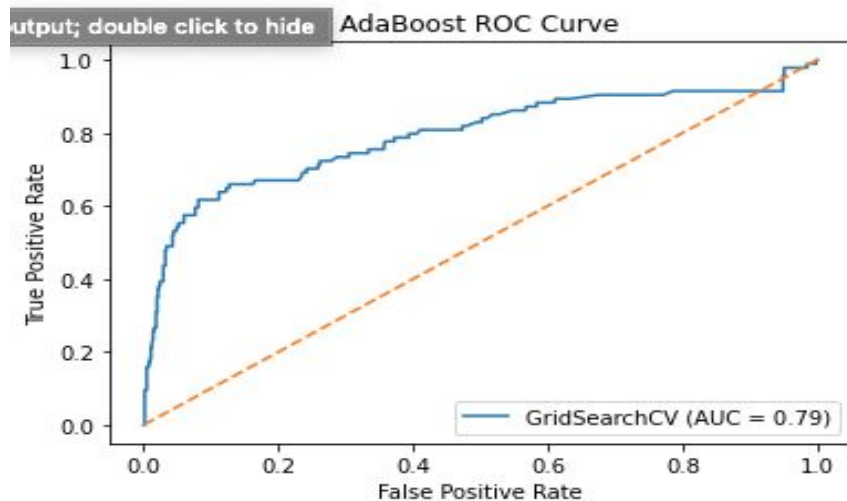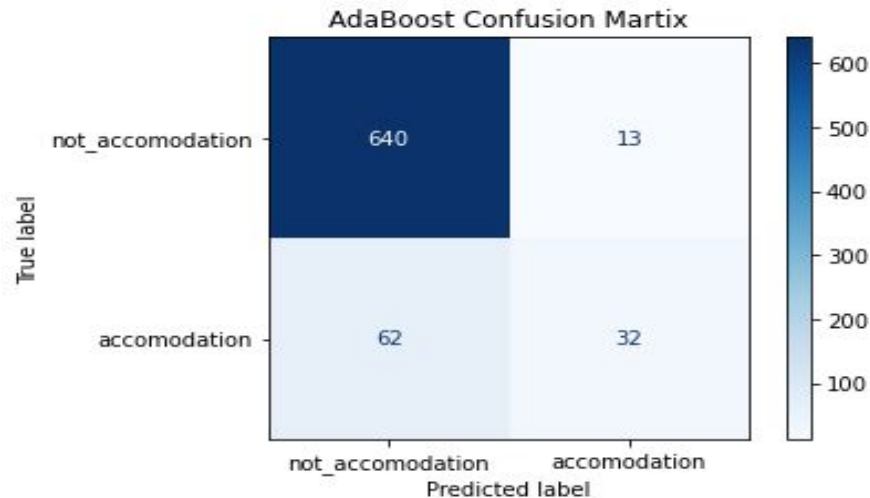- AdaBoost delivered a decent AUC score
- This model, like all the others, struggled to distinguish between classes where each model scored either at, or just above baseline accuracy

AdaBoost Confusion Martix
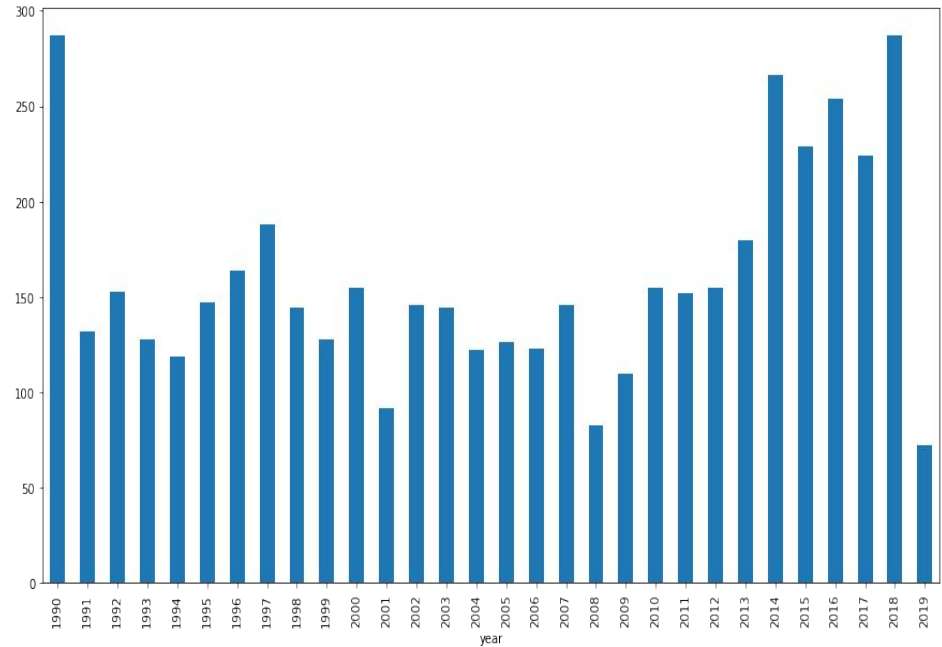

AdaBoost ROC Curve
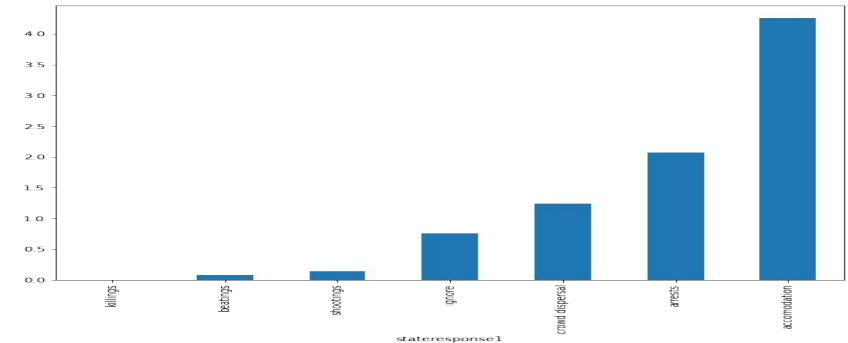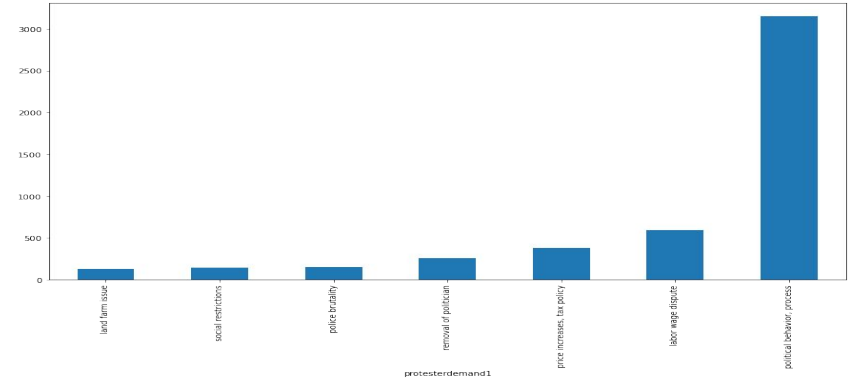
Regional Analysis
– Europe

Europe

- Total protests in europe were on the rise after 2013
- Protests between 2014 - 2018 were the highest in nearly 20 years
- 2014's rise in protests were led by a Ukraine's Euromaidan revolution, which accounted for 91 protests
- Following years saw protests rise in Ireland, Germany, France, UK

Europe

- Political behavior and process are the leading protestor demands
- Labor Wages are second

- Protests that are Accommodated have a duration of 4 days, slightly higher than other responses
- However, Accomodations only account for 5% of state responses

Europe
Models & Metrics:

- Logistic Regression model, was in line with baseline, not predicting higher results
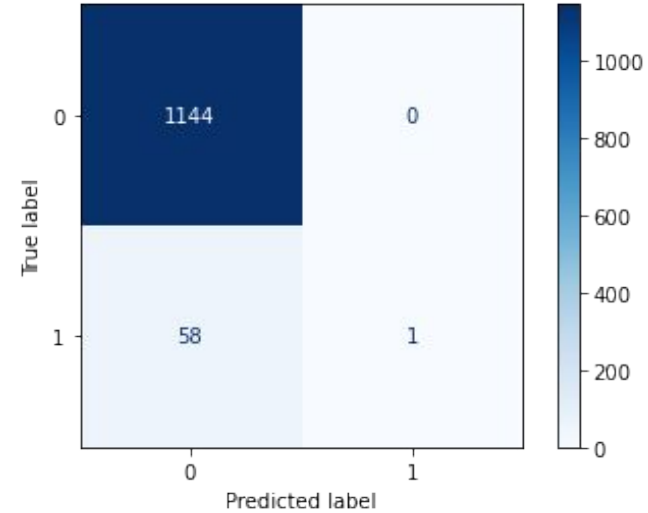
Logistic Regression	Model training data
Accuracy	95%
Specificity	1
Precision	1
Sensitivity	2%

Regional Analysis – Africa

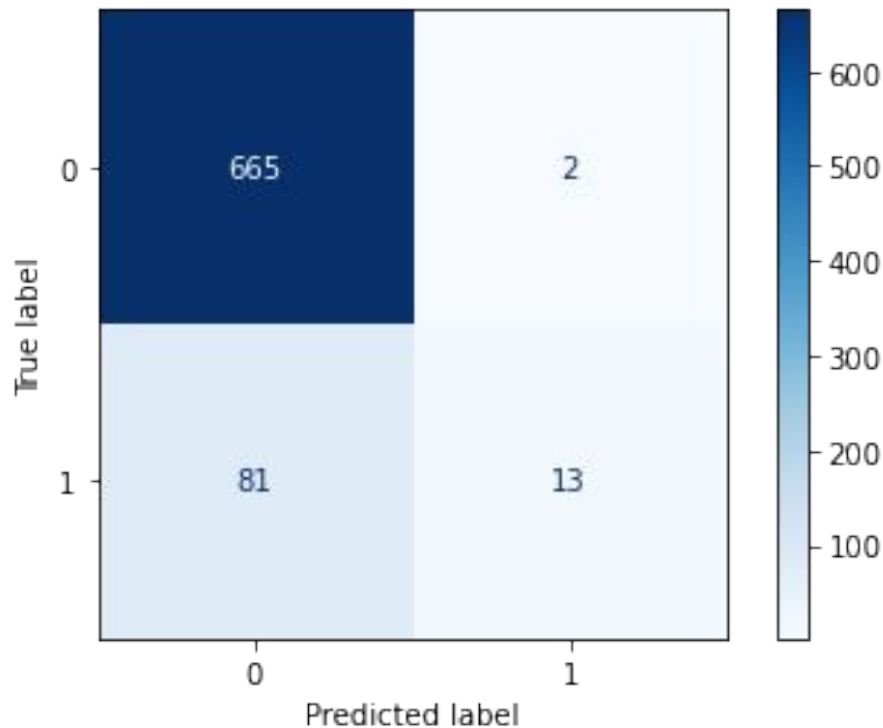# Protest Trends in Africa

State Response Counts
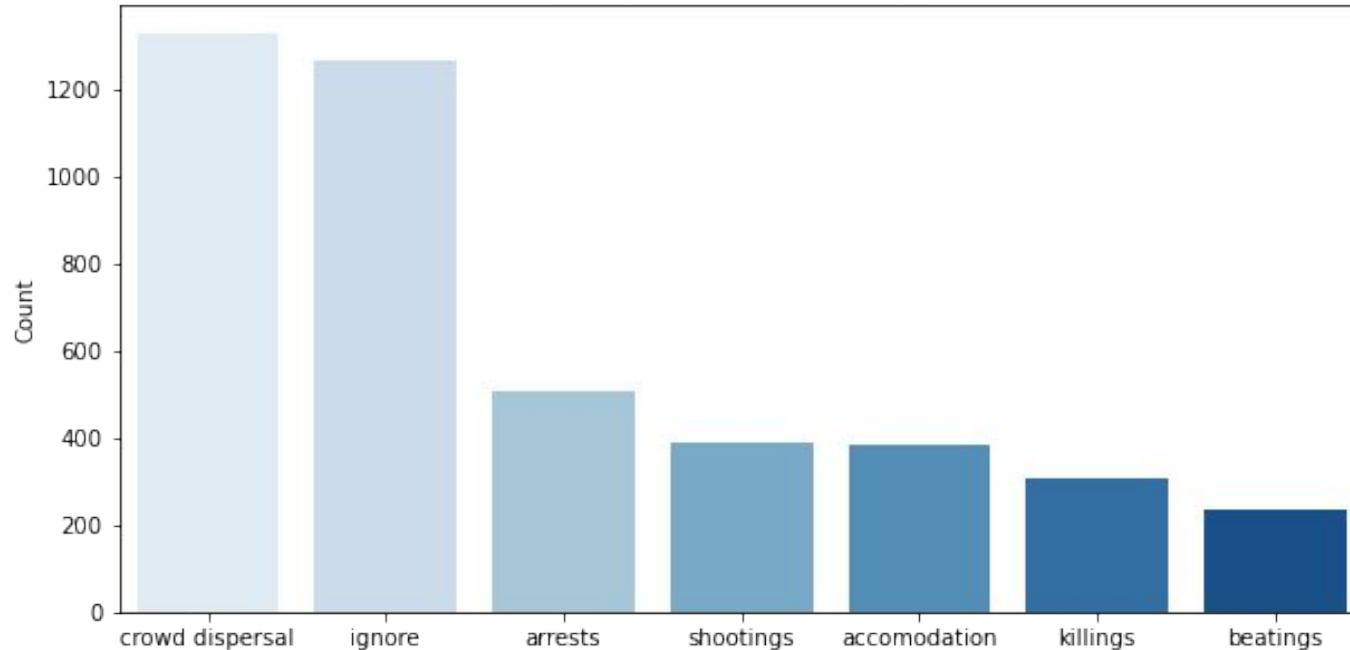
Crowd Dispersal and Ignore are the typical state responses in Africa

Models: Africa

- Random Forest Classifier with PCA
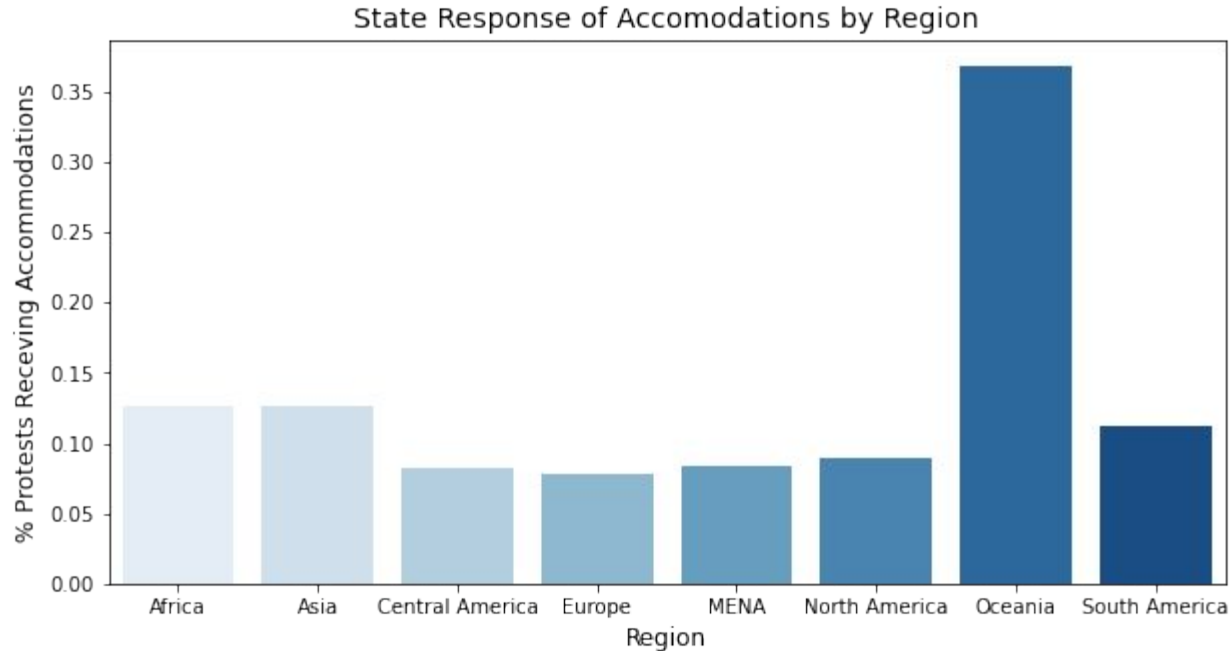
- Baseline Score:
  - 87.6% are not accommodated
- Accuracy:
  - Cross-Val: 88.8%
  - Train: 88.9%
  - Test: 89.1%

- Sensitivity: 99.7%
- Specificity: 13.8%

Return to Worldwide Analysis

Accommodations by region

State Response of Accomodations by Region

The best indicator of whether a protest will receive accommodations is what region it is in.

Note - Oceania is one country.

Conclusions & next steps

Conclusions

- We were not able to gain an insight on the features that predict a successful protest
- None of our modeling was able to significantly predict protest accomodations over baseline

Next Steps

- Combine our work on regional data across the entire data set
- Time series analysis
- Combining these data with data on the rise of different forms of social media and technology
- And ...

References

- https://www.livescience.com/16153-10-significant-political-protests.html

- https://www.globalcitizen.org/es/content/peace-protests-dallas-response/

- https://www.theatlantic.com/technology/archive/2020/06/why-protests-work/613420/