

MILITARIZED INTERSTATE DISPUTES ANALYSIS

GA DATA SCIENCE- CHELSIE ZHANG



LIBYA CIVIL WAR



NORTH KOREA CRISIS



SOUTH CHINA
SEA DISPUTE



ISIS



PROBLEM:
WHICH DISPUTES ARE LIKELY ESCALATE TO
WAR?

DATA SOURCE:

THE CORRELATES OF WAR PROJECT
THE PENNSYLVANIA STATE UNIVERSITY
[HTTP://WWW.CORRELATESOFWAR.ORG/](http://WWW.CORRELATESOFWAR.ORG/)

DATA

ALL INSTANCES OF WHEN ONE STATE THREATENED, DISPLAYED, OR USED FORCE AGAINST ANOTHER BETWEEN 1816 AND 2010

OBSERVATION:

2,586 RECORDS OF DISPUTE

6,132 RECORDS OF PARTICIPANTS

VARIABLES:

- YEAR OF DISPUTE
- DURATION
- COUNTRIES INVOLVED
- RECIPROCATED DISPUTE
- NUMBER OF COUNTRIES ON EACH SIDE
- REVISIONIST STATE
- REV. TYPE: TERRITORY, POLICY, REGIME, OTHER

OUTPUT:

- HOSTILITY LEVEL OF DISPUTE
 - 1. NO MILITARIZED ACTION (L)
 - 2. THREAT TO USE FORCE (L)
 - 3. DISPLAY OF FORCE (L)
 - 4. USE OF FORCE (H)
 - 5. WAR (H)

DATA OVERVIEW

DATA SOURCE: WIKIPEDIA

DATA TYPE: COUNTRY INFORMATION

DATA FORMAT: JSON

DATA SIZE: 1000+ COUNTRIES

DATA PAGES: 1000+ COUNTRIES

DATA COLUMNS: 1000+ COUNTRIES

DATA ROWS: 1000+ COUNTRIES

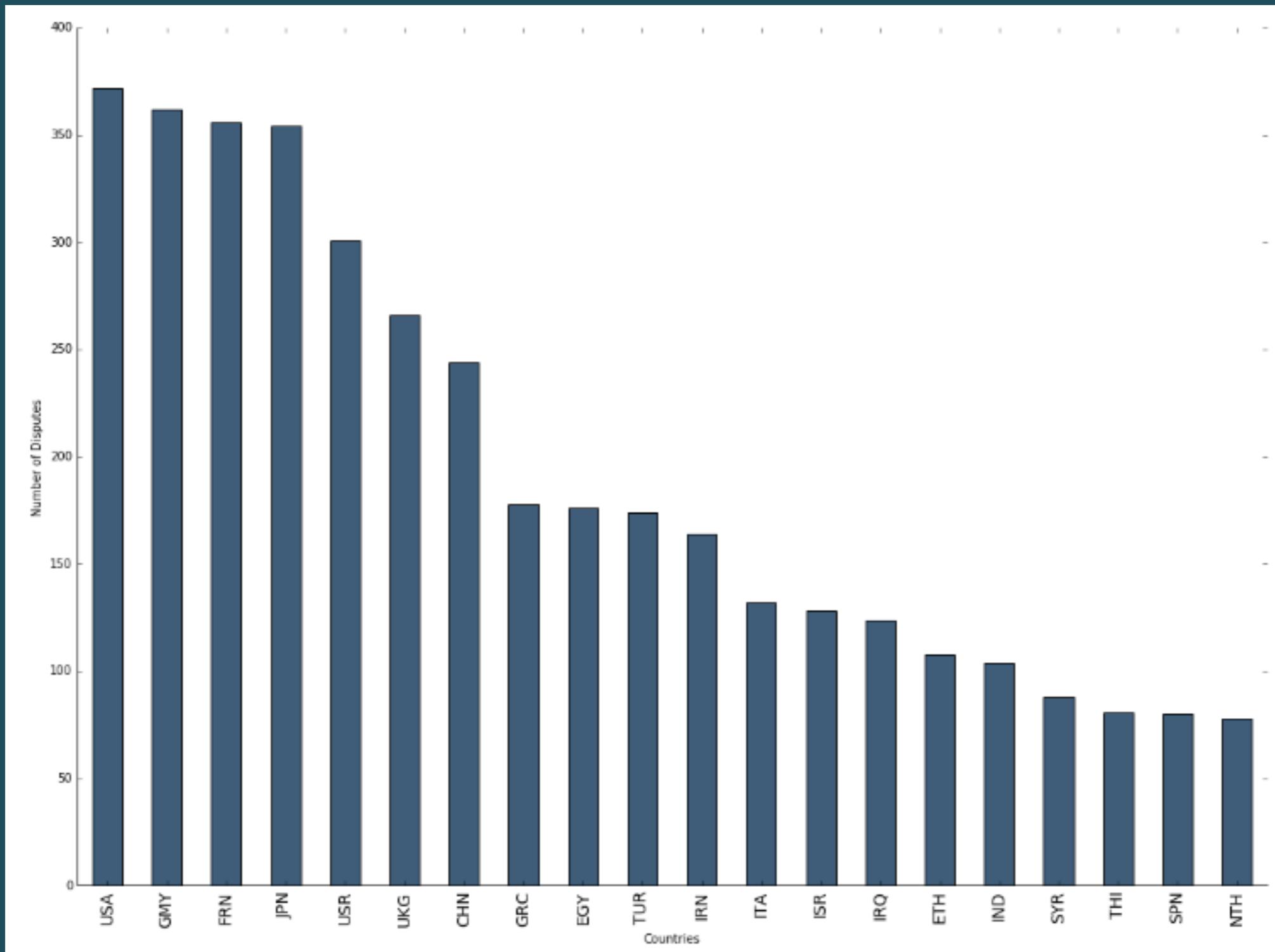
DATA INDEX: 1000+ COUNTRIES

DATA SUMMARY: 1000+ COUNTRIES

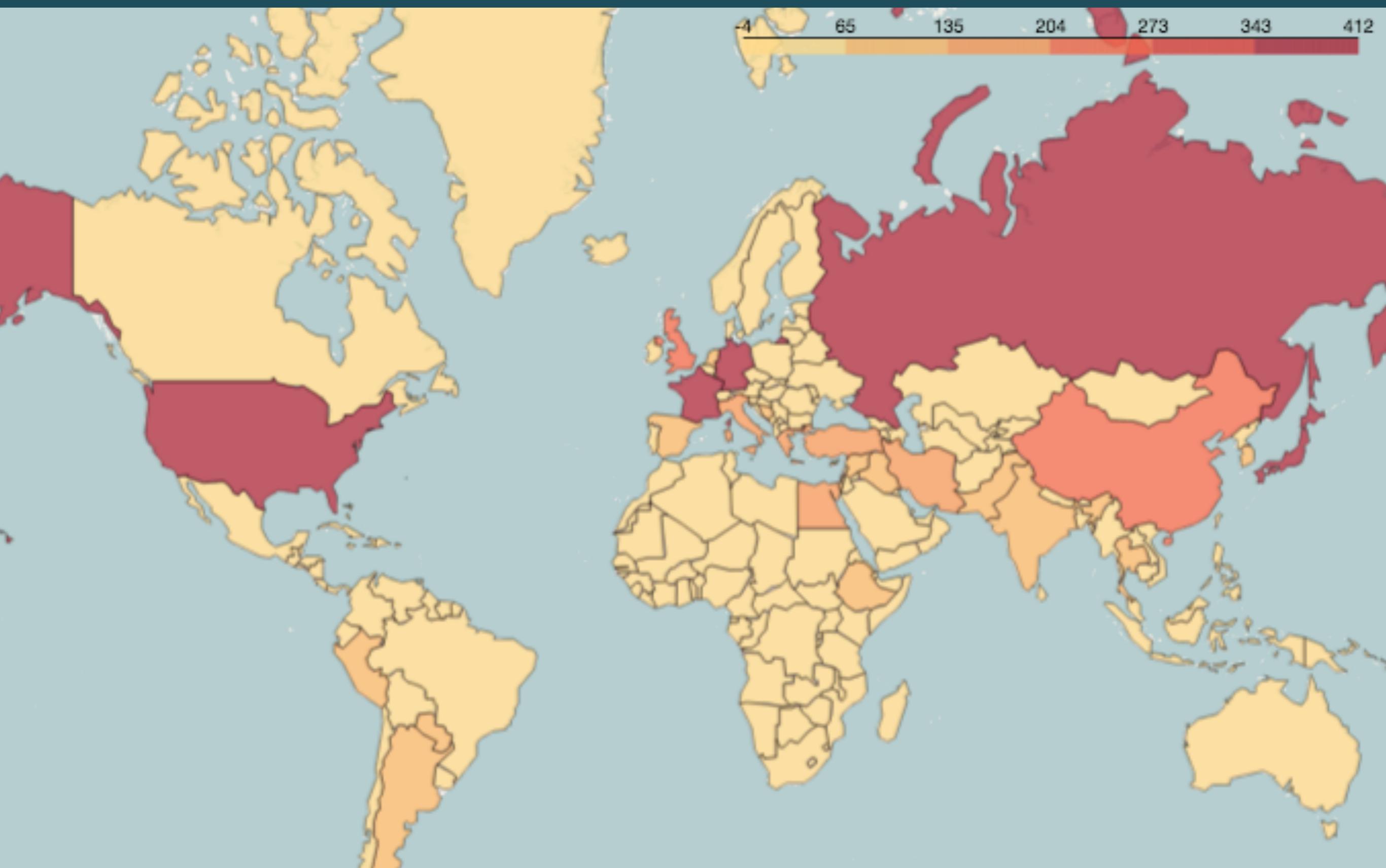
DATA ANALYSIS: 1000+ COUNTRIES

DATA PREDICTION: 1000+ COUNTRIES

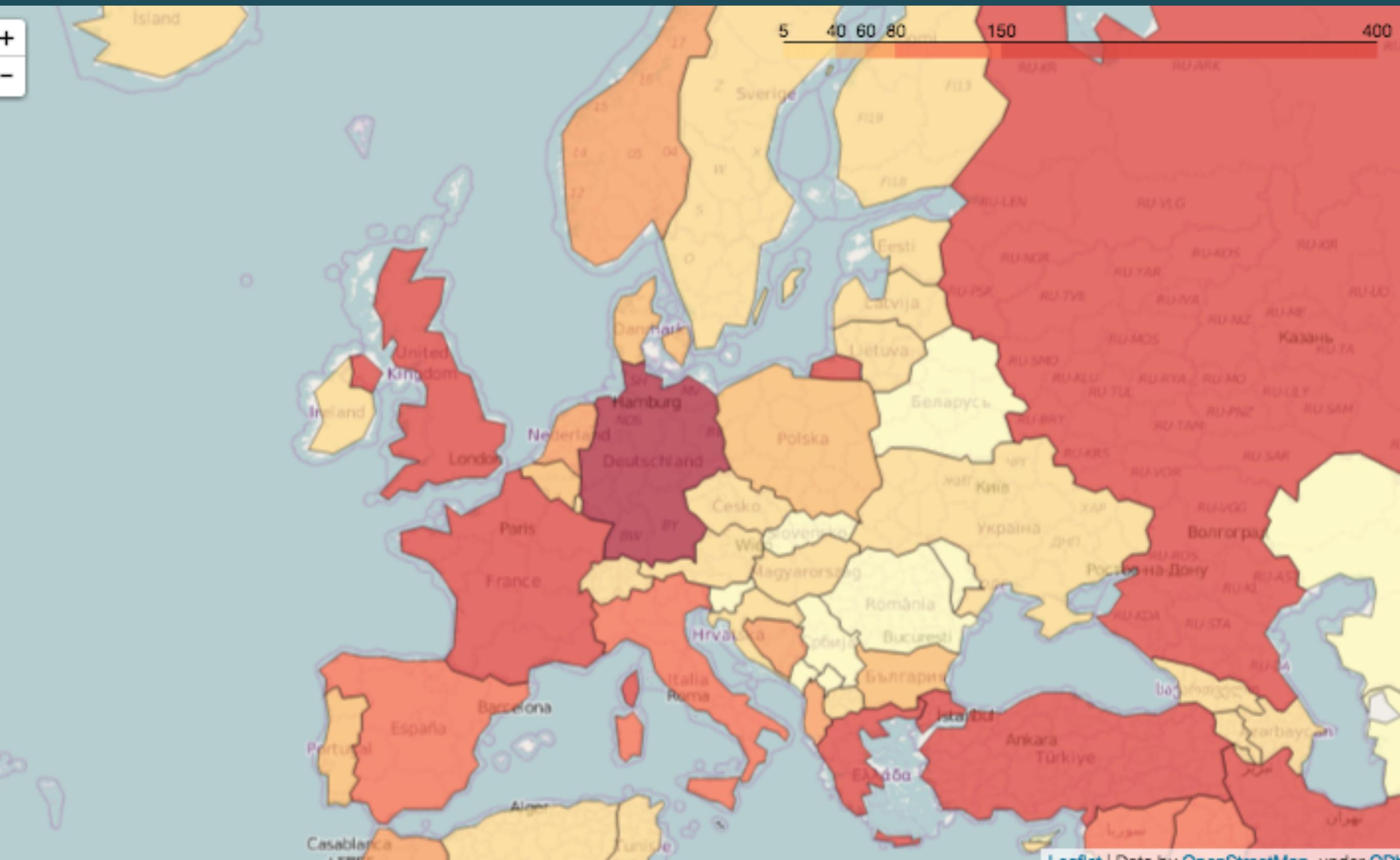
COUNTRIES WITH MOST NUMBER OF DISPUTES



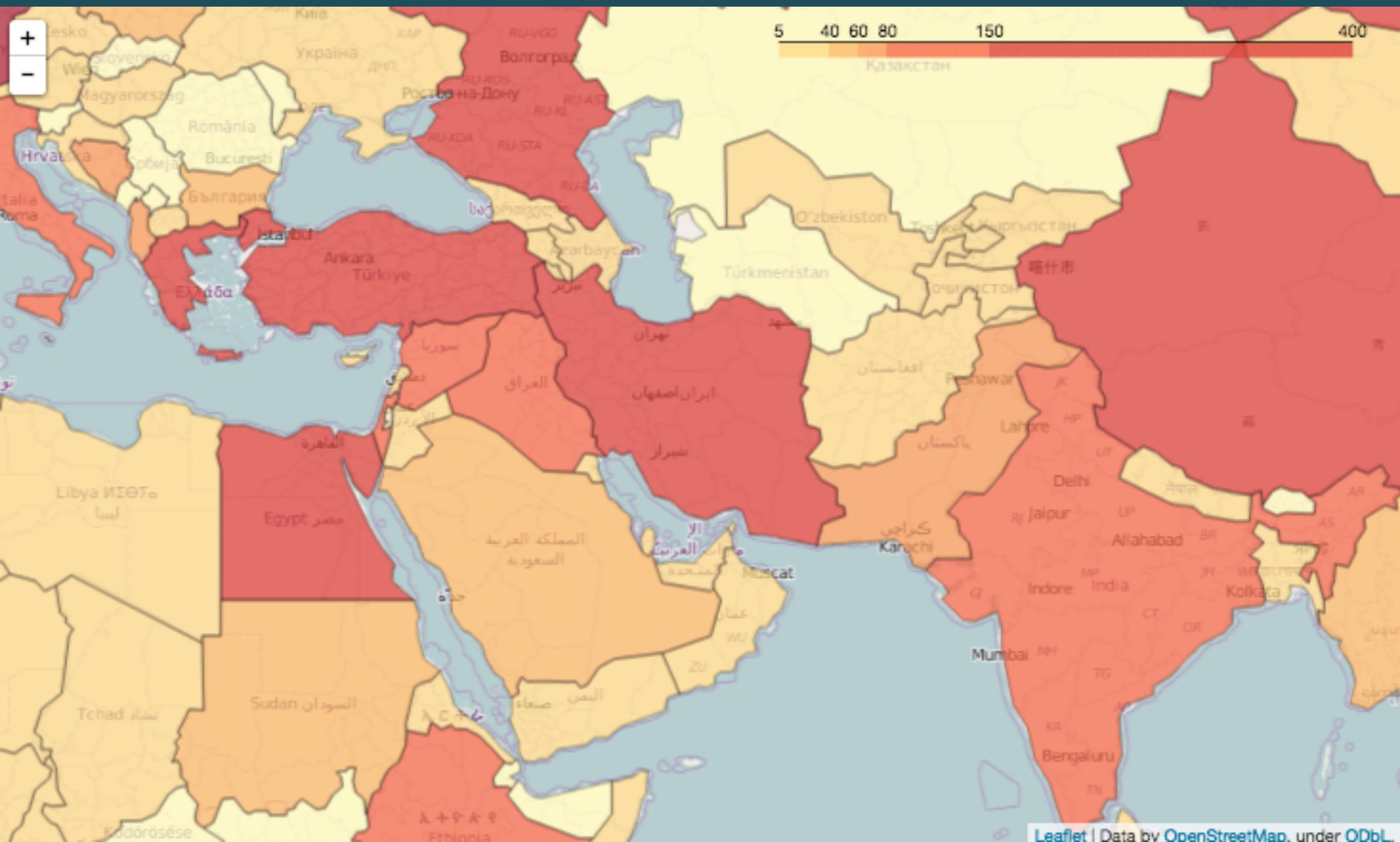
FOLIUM WORLD MAP



EUROPE



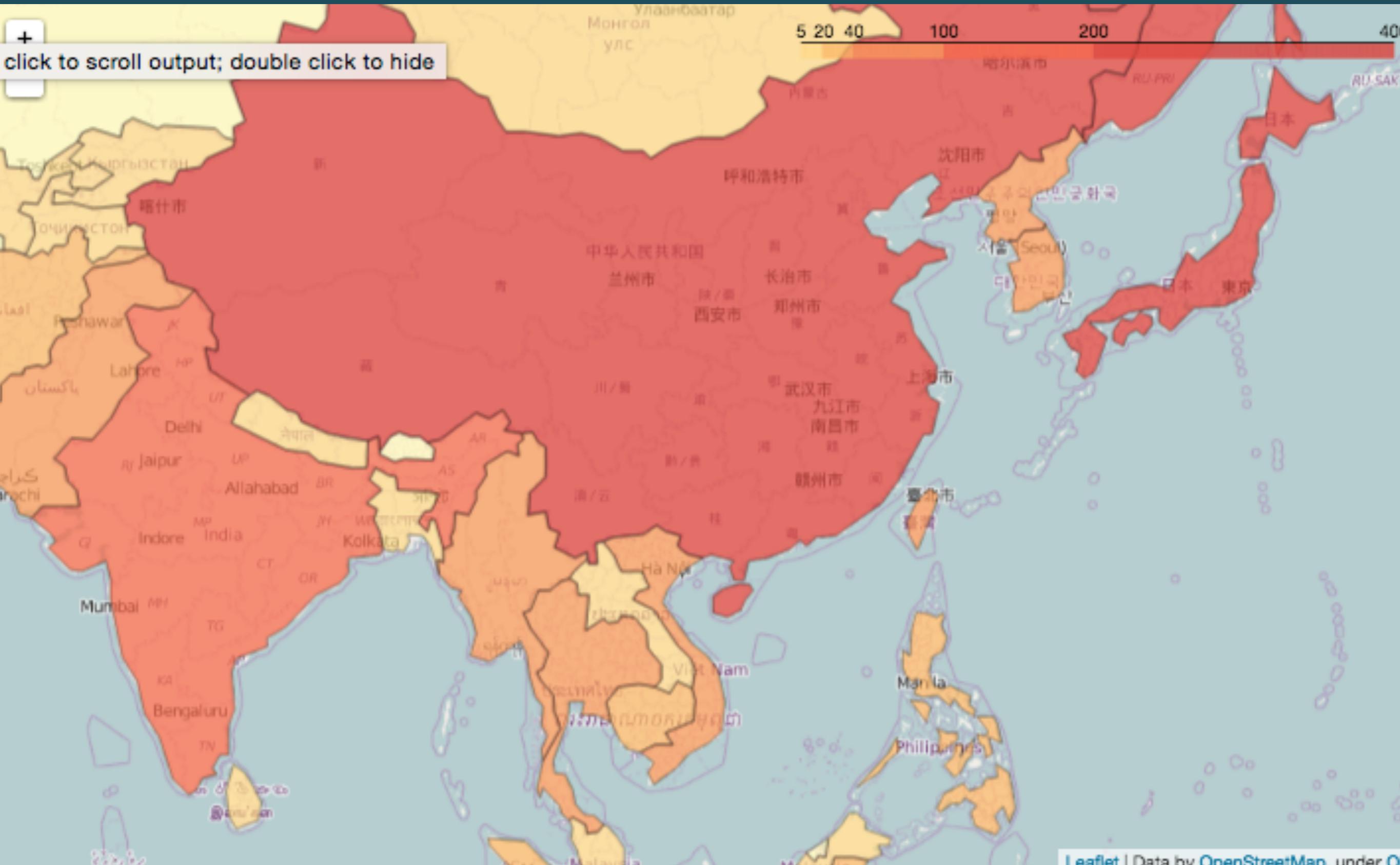
MIDDLE EAST AND NORTH AFRICA



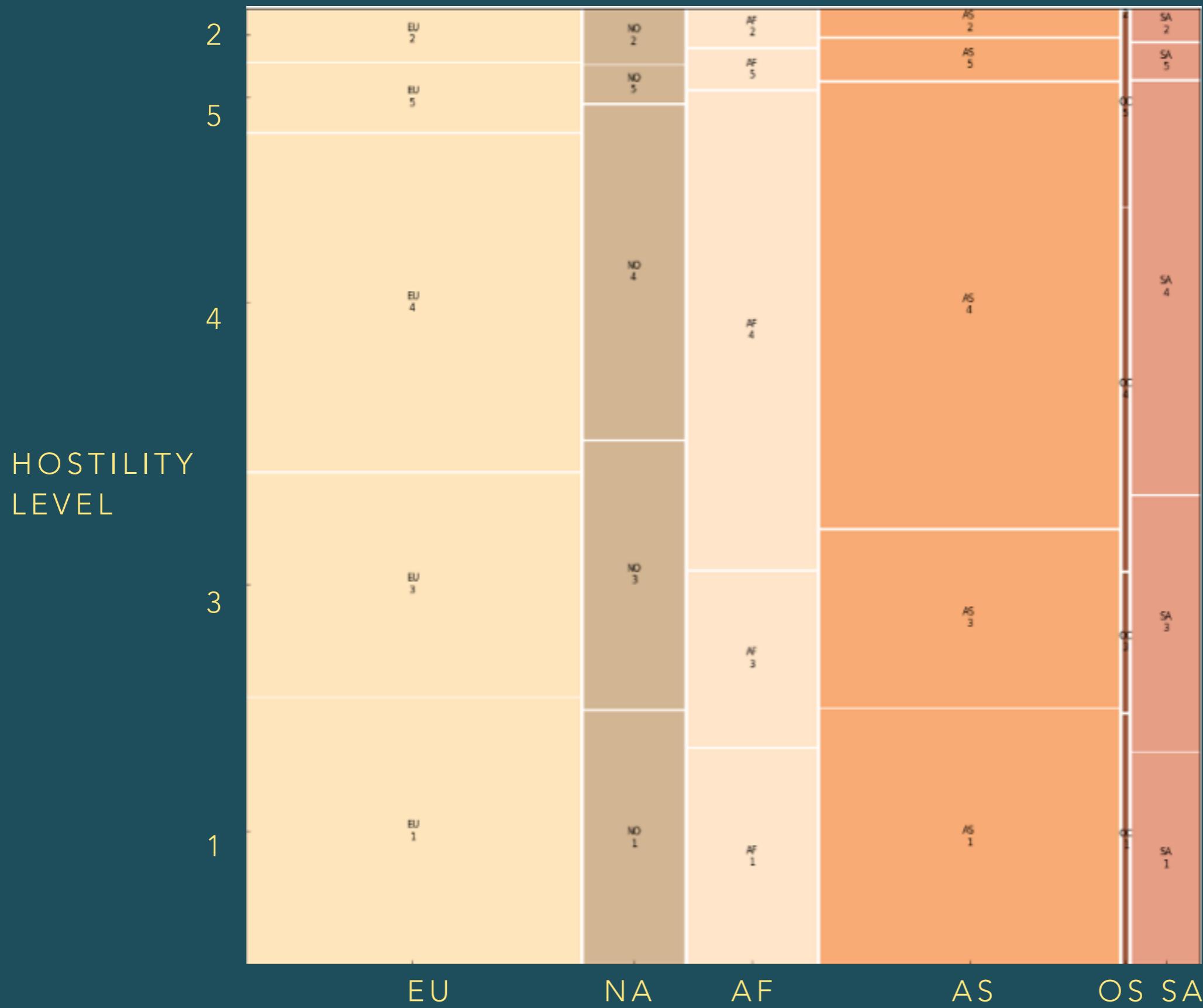
ASIA

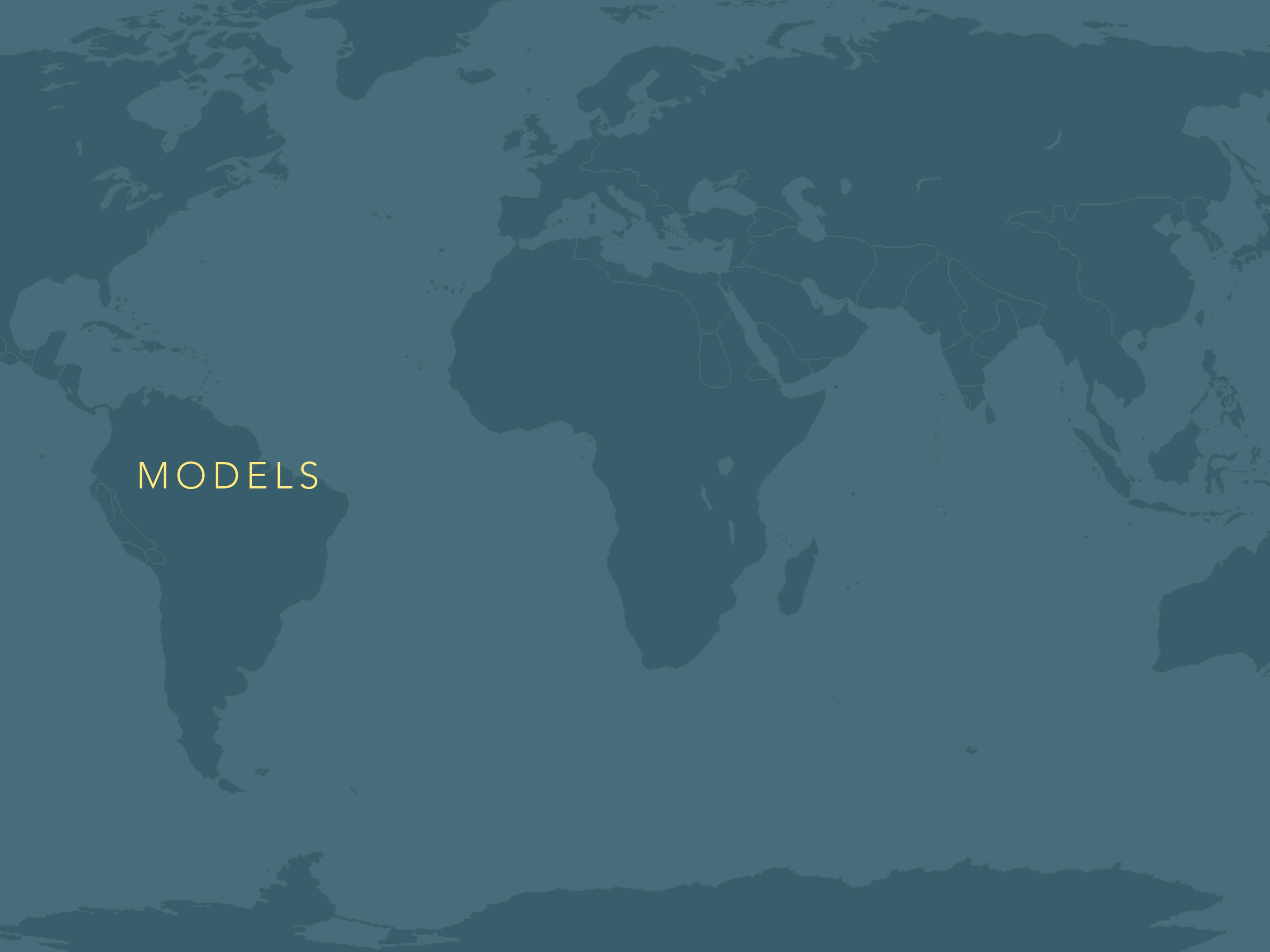
1

click to scroll output; double click to hide



CONTINENTS X HOSTILITY LEVEL





MODELS

MODEL - DECISION TREE

BENCHMARK: .71 (1,649 HIGH HOSTILITY LEVEL OUT OF 2303)

TEST ACCURACY: .73

CONFUSION MATRIX:

```
ARRAY([[ 68,  86],  
      [ 58, 306]])
```

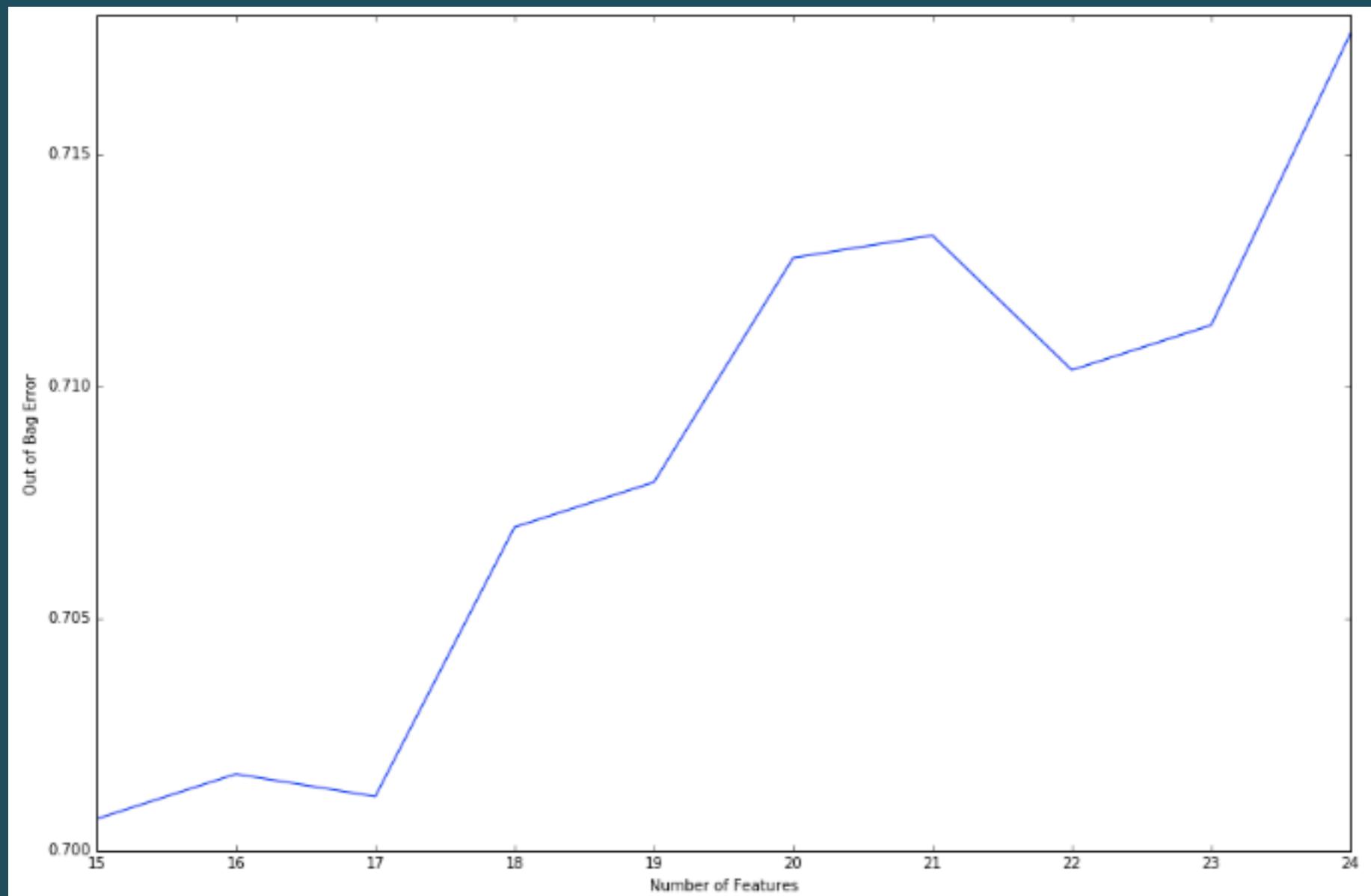
IMPORTANT FEATURES:

	Feature	Importance
0	Recip	0.406289
4	Revstate_Policy	0.093370
184	Tur	0.073506
204	Revstate	0.061391
67	Frn	0.058505
229	Fra	0.053087
225	Deu	0.045708

MODEL - RANDOM FOREST

BENCHMARK: .71 (1,649 HIGH HOSTILITY LEVEL OUT OF 2303)

OOB ACCURACY: .72



MODEL - BOOSTING & NAÏVE BAYES

BENCHMARK: .71 (1,649 HIGH HOSTILITY LEVEL OUT OF 2303)

BOOSTING

CV SCORE: .71

BERNOULLINB

TEST ACCURACY: .70



PREDICTION

NORTH KOREA CRISIS

KEY VARIABLES:

- RECIPROCATED DISPUTE = 1
- NUMBER OF COUNTRIES ON SIDE A = 1
- NUMBER OF COUNTRIES ON SIDE B = 1
- REVISION TYPE _ REGIME = 1
- REVISION TYPE _ TERRITORY = 1
- REVISION STATE = 1
- COUNTRIES/REGIONS INVOLVED: NORTH KOREA, SOUTH KOREA

OUTPUT:

[0]

HOWEVER...

IF WE INCLUDE CHINA AND US IN THE MODEL, THE OUTPUT INDICATES WAR





SOUTH CHINA SEA DISPUTE

KEY VARIABLES:

- RECIPROCATED DISPUTE = 0
- NUMBER OF COUNTRIES ON SIDE A = 1
- NUMBER OF COUNTRIES ON SIDE B = 8
- REVISION TYPE _ TERRITORY = 8
- REVISION STATE = 8
- COUNTRIES/REGIONS INVOLVED: CHINA, PHILIPPINES, VIETNAM, SINGAPORE, MALAYSIA, CAMBODIA, TAIWAN, INDONESIA

OUTPUT:

[0]



ISIS

KEY VARIABLES:

- RECIPROCATED DISPUTE = 1
- NUMBER OF COUNTRIES ON SIDE A = 1
- NUMBER OF COUNTRIES ON SIDE B = 2
- REVISION TYPE _ REGIME = 8
- REVISION STATE = 2
- COUNTRIES/REGIONS INVOLVED: US, IRAQ, SYRIA

OUTPUT:

[1]

CAVEATS & FURTHER STUDY

- HISTORICAL DATA ONLY
- OTHER DATA INPUTS
 - GDP GROWTH
 - POPULATION GROWTH
 - TRADE RECORDS
 - POLITICAL STABILITY
- NPL
- SENTIMENT ANALYSIS FROM SOCIAL MEDIA

THANKS