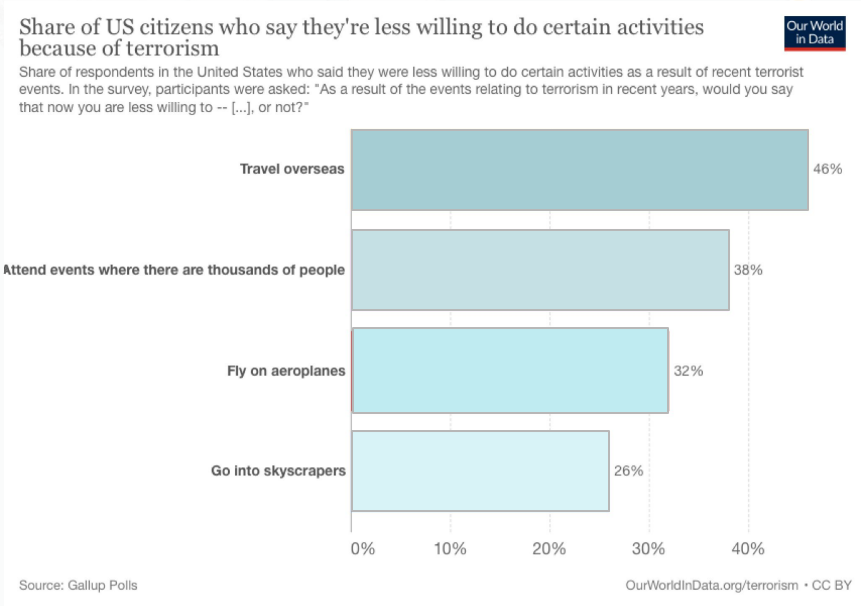
Global Terrorism

Classification of Terrorist Groups

Problem Statement



Terrorism affects people across the world on a daily basis.

According to statistics from the Global Terrorism Database, more than *200,000* terrorist attacks have been recorded from 1970 to the present day.

Attacks typically involve multiple injuries, deaths and directly cause massive property losses.

In addition, they bring tremendous psychological fear and anxiety.

To report an imminent threat call **999** or ring the Anti-Terrorist Hotline on **0800 789 321** Current national threat level: **SEVERE** Find out more SEARCH

SECURITYSERVICE MI5

Home Who we are What we do How we work What you can do News Visit Careers

Fact or Fiction

Terrorism is the biggest current threat to national security

Fact

Fiction

Awards & Support

TOP 75 EMPLOYER

TOP 100 GRADUATE EMPLOYERS

THE TIMES GRADUATE RECRUITMENT AWARDS 2021 BEST WEBSITE

Stonewall TOP 100 EMPLOYER 2019

NatWest BRITISH LGBT AWARDS 2018

disability confident LEADER

Stonewall STAR PERFORMER

To report an imminent threat call **999** or ring the Anti-Terrorist Hotline on **0800 789 321** Current national threat level: **SEVERE** Find out more SEARCH

SECURITYSERVICE MI5

Home Who we are What we do How we work What you can do News Visit Careers

Fact or Fiction

Terrorism is the biggest current threat to national security

Fact Fiction

Awards & Support

TOP 75 EMPLOYER

TOP 100 GRADUATE EMPLOYERS

AWARDS 2021 BEST WEBSITE

Stonewall TOP 100 EMPLOYER 2019

NatWest BRITISH LGBT AWARDS 2018

disability confident LEADER

Stonewall STAR PERFORMER

Correct!

The biggest threat we currently face comes from international terrorist groups and individuals inspired by them. Terrorist organisations in Northern Ireland also continue to pose a serious threat. Espionage (including cyber-espionage) also remains a significant problem, with at least 20 foreign intelligence services active against UK interests.

Show Me Another Question

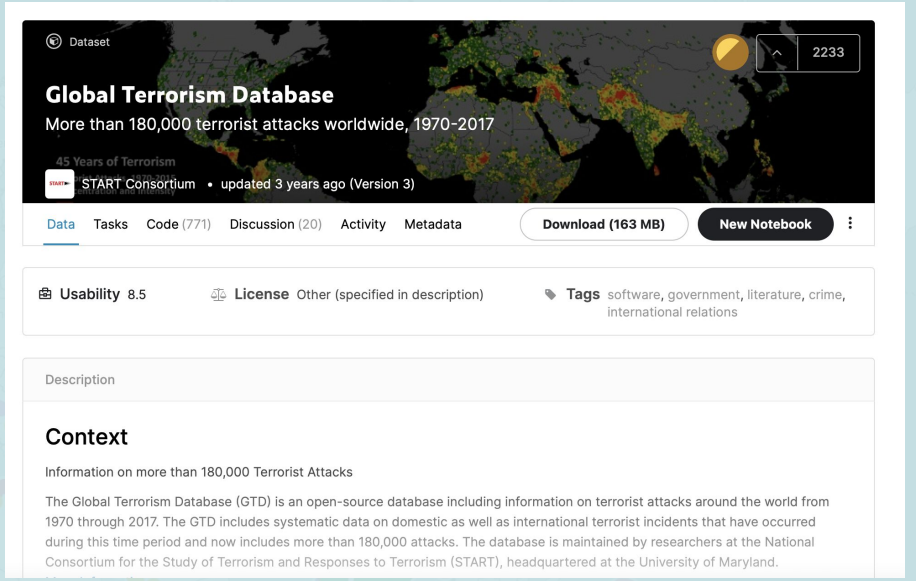
Potential Audience

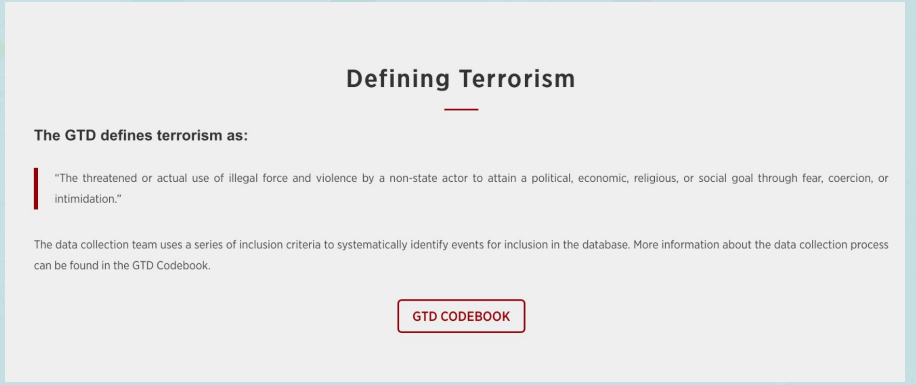
The analysis and prediction of terrorist groups would provide valuable information for antiterrorism and terrorism prevention operations, enabling authorities to assign attacks whose organisation are currently marked as ‘unknown’ and make them accountable.

Datasource

- Global Terrorism Database
- derived from Kaggle as csv
- over 180,00 observations
- collected from 1970 - 2017
- 135 features

(including attack type, weapon type, target type, country, date and ideologies.)





Hypothesis + Success Measures

Hypothesis

We can predict the 'terrorist group' who performed the attack.

For this type of prediction we will be using classification models. These types of models will look at the characteristics of each attack and predict who was responsible using the data provided.

Success Metric

Scoring our models.

In order to say we have a good enough model to predict each organisation correctly we will be measuring using an accuracy score. This is simply the amount of correctly predicted observations over the total observations. A perfect score would be 1 or 100% accuracy.

Cleaning The Data

Dates

Some of our dates were either missing or hand-written in other columns, using an inbuilt method call pd.todatetime we were able to collect the correct dates and apply them to the observations.

Locations

We also found that some of the location details were only partly filled, so we could use another built in function called geopandas to gather latitude and longitudes from city names and visa versa.

Unknown Values

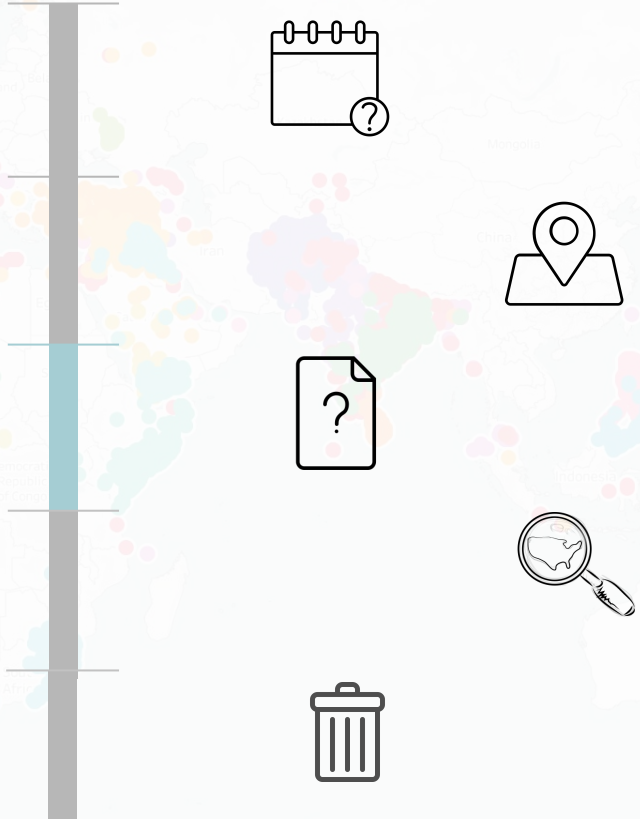
45% of the dataset's terrorist organisation column was marked as 'unknown', we'll remove these observations for now as they won't help with designing the model. But we'll keep them for later.

Null or Missing Values

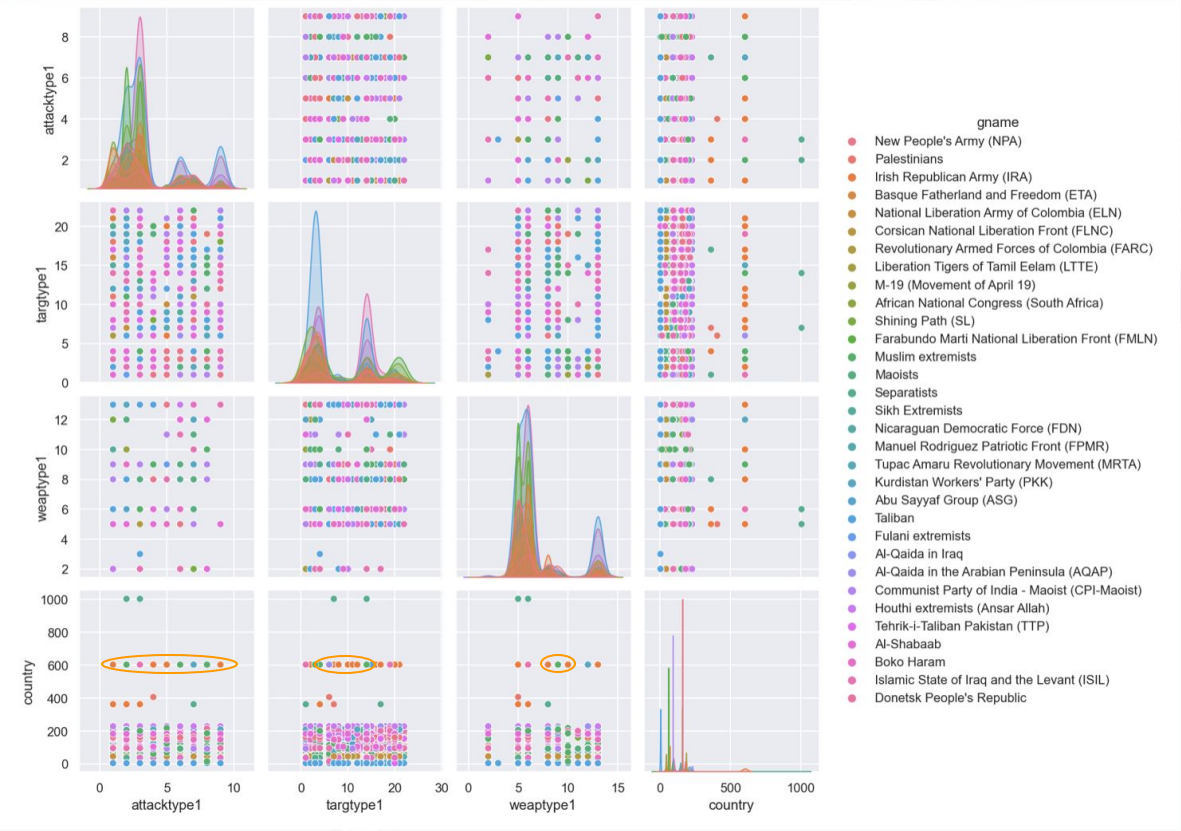
Any values missing that we could not find elsewhere in the data and were unusable were then dropped.

Removing Unnecessary Columns

Any columns that were deemed 'subjective', held too many missing values, or had 'descriptive' text which duplicated data in other columns were also dropped.



Relationships in the Data



After cleaning the data we were left with:

57990 observations

32 terrorist organisations

41 columns

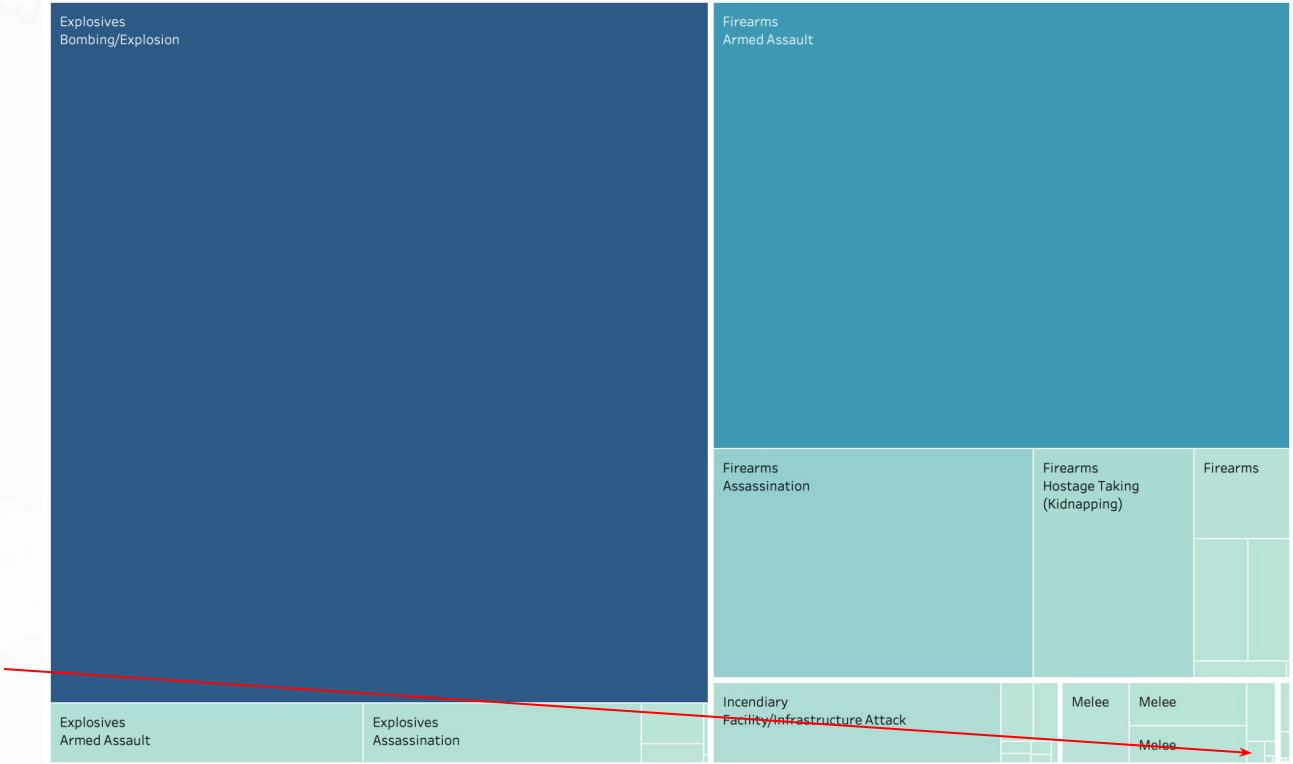
(39 as categorical)

Relationships in the Data

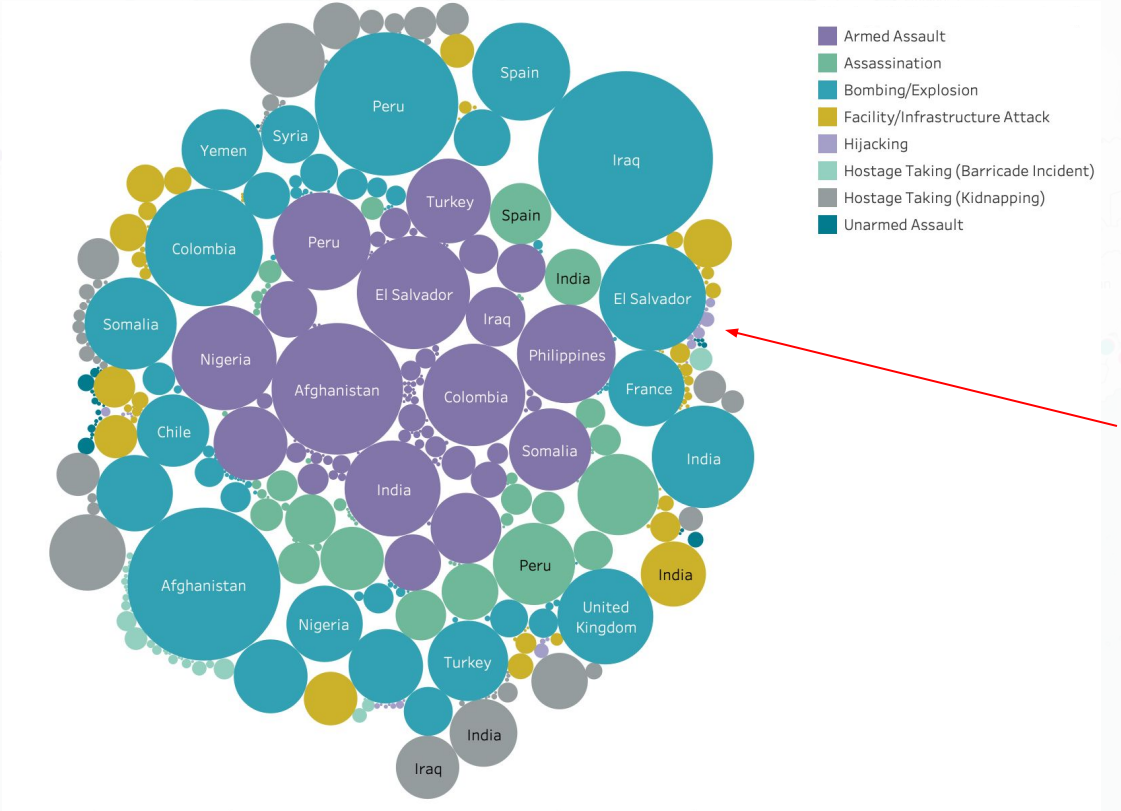
We can see that the highest number of attacks observed were from attack type ‘bombing/explosion’ using the weapon type ‘explosives’.

With ‘armed assault’ using ‘firearms’ as the second most likely means of attack.

Attack types such as ‘nuclear’, ‘biological’ and ‘chemical’ were much rarer occurrences.



Relationships in the Data

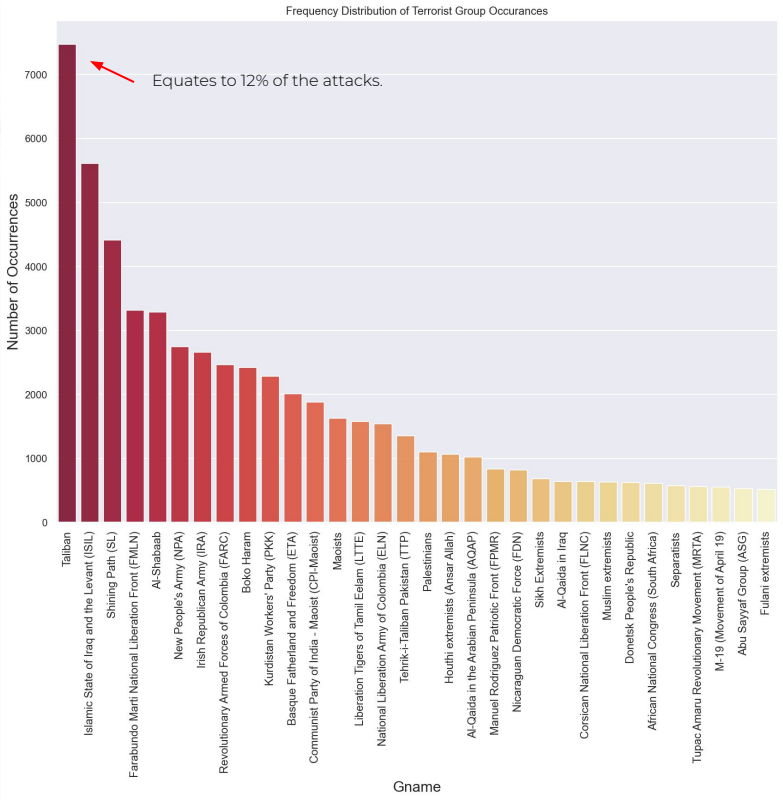


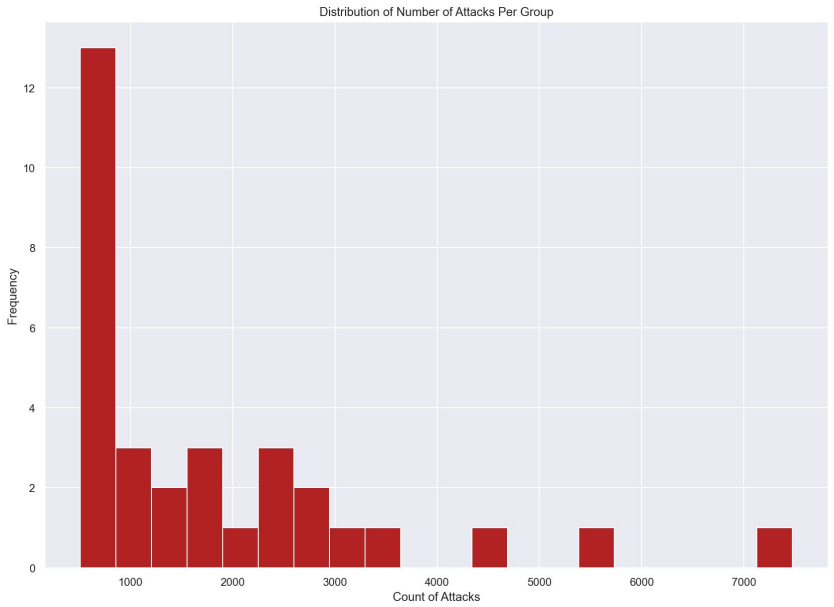
We can see that ‘bombing / explosion’ again holds the highest number of occurrances, appearing more often in countries such as Iraq, Afghanistan and Peru.

With countries such as Philippines and Columbia showing the largest number of ‘hijackings’.

And Peru and United Kingdom ranking highest for number of ‘assassination’.

Reviewing The Target

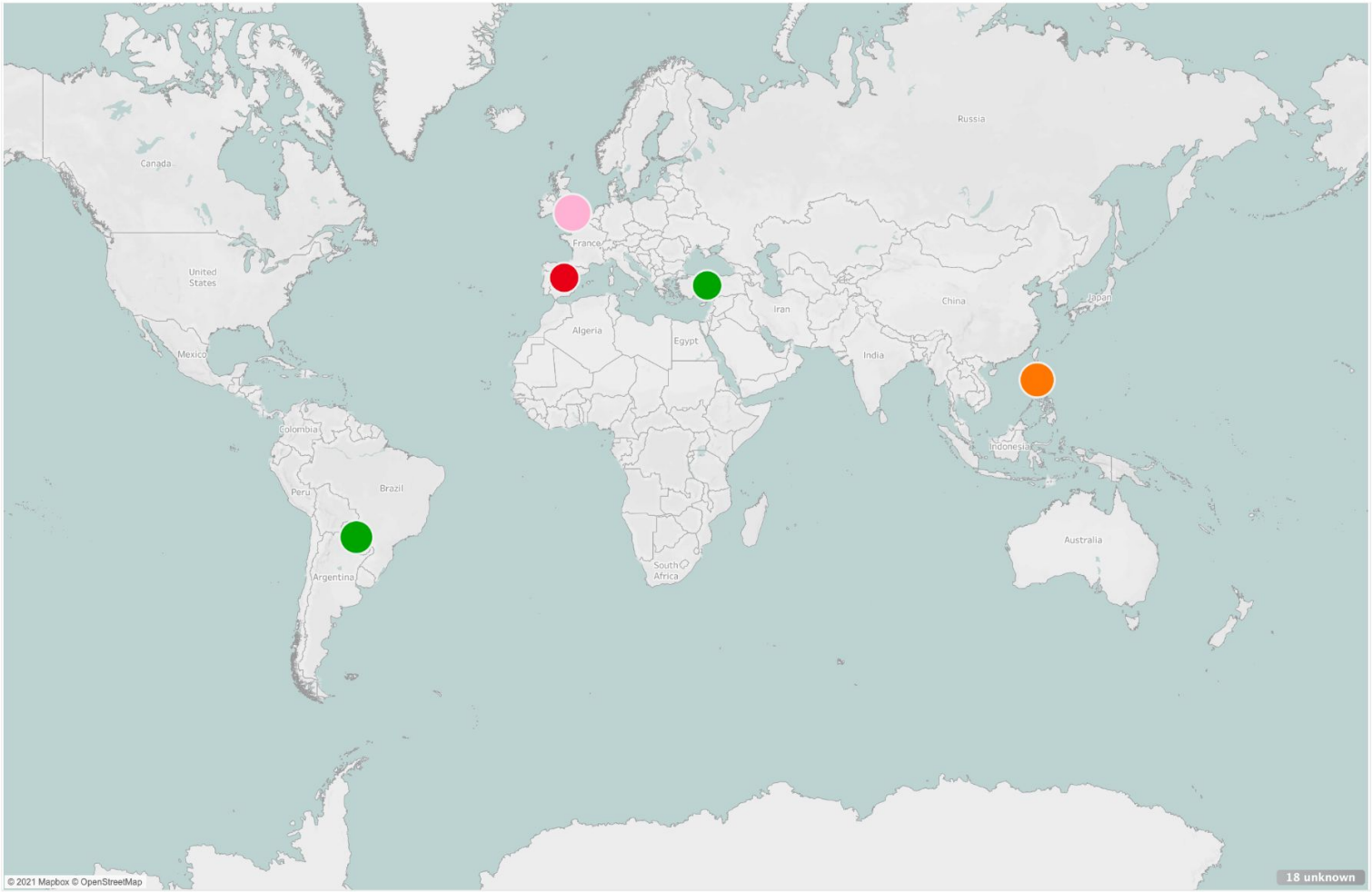




The terrorist organisation with the largest number of attacks were the Taliban with over 7000 events.

However, as we can see from the distribution plotted above, the majority of the groups performed less than 1000 attacks each, leading to an imbalance in the data.

Map Showing No. of Kills per Group (every 5yrs) - 1970



Gname

- Abu Sayyaf Group ..
- African National C..
- Al-Qaida in Iraq
- Al-Qaida in the Ar..
- Al-Shabaab
- Basque Fatherlan..
- Boko Haram
- Communist Party ..
- Corsican National ..
- Donetsk People's ..
- Farabundo Marti ..
- Fulani extremists
- Irish Republican A..
- Islamic State of Ir..
- Kurdistan Worker..
- Liberation Tigers ..
- M-19 (Movement ..
- Manuel Rodriguez..
- Maoists
- Muslim extremists
- National Liberatio..
- New People's Arm..
- Nicaraguan Demo..
- Palestinians
- Revolutionary Ar..
- Separatists
- Shining Path (SL)
- Sikh Extremists
- Taliban
- Tehrik-i-Taliban P..

YEAR(Date)

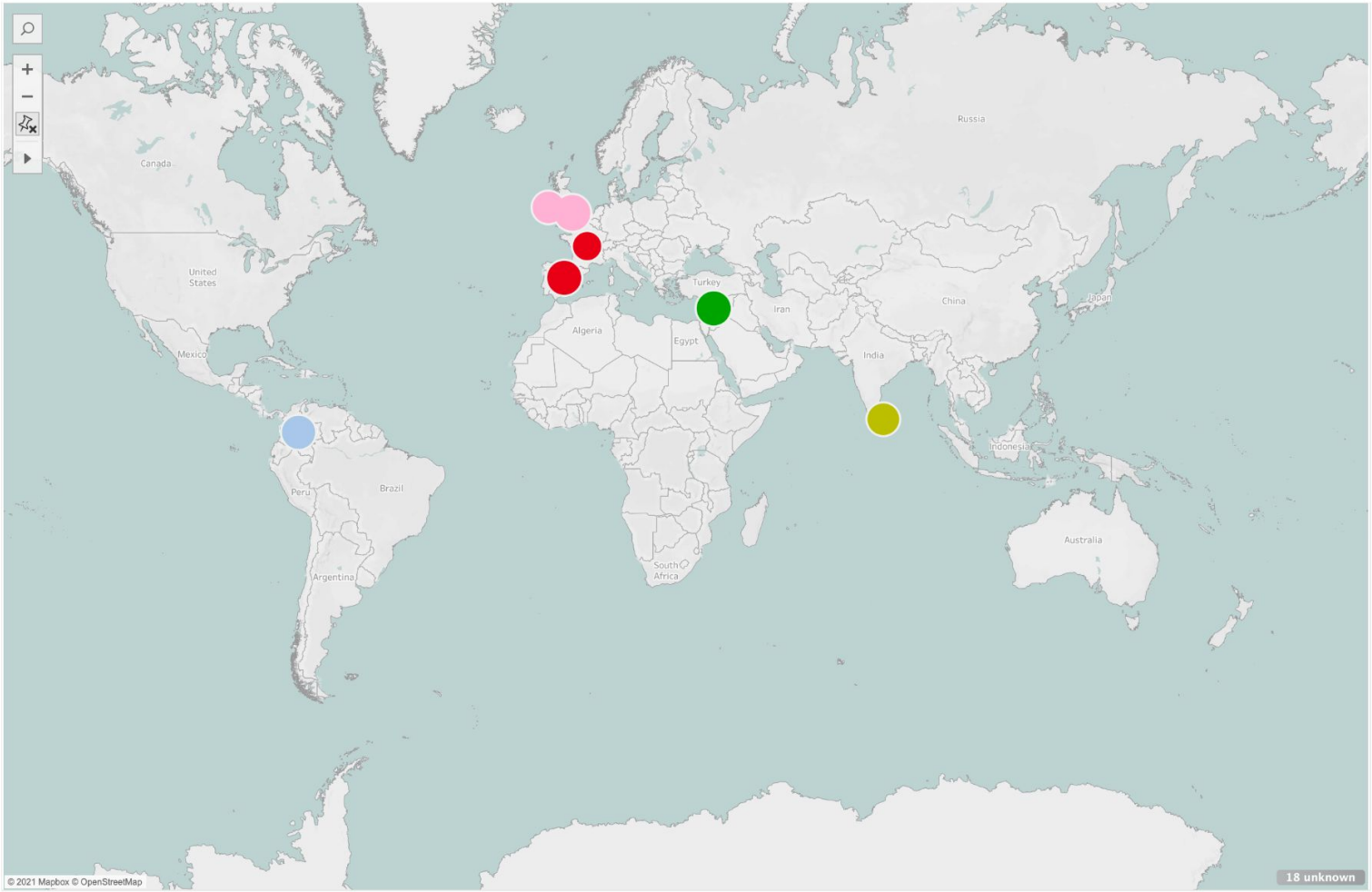
1970

Show history

Nkill

0.0
100.0
200.0
329.0

Map Showing No. of Kills per Group (every 5yrs) - 1975



Gname

- Abu Sayyaf Group ..
- African National C..
- Al-Qaida in Iraq
- Al-Qaida in the Ar..
- Al-Shabaab
- Basque Fatherlan..
- Boko Haram
- Communist Party ..
- Corsican National ..
- Donetsk People's ..
- Farabundo Marti ..
- Fulani extremists
- Irish Republican A..
- Islamic State of Ir..
- Kurdistan Worker..
- Liberation Tigers ..
- M-19 (Movement ..
- Manuel Rodriguez..
- Maoists
- Muslim extremists
- National Liberatio..
- New People's Arm..
- Nicaraguan Demo..
- Palestinians
- Revolutionary Ar..
- Separatists
- Shining Path (SL)
- Sikh Extremists
- Taliban
- Tehrik-i-Taliban P..

YEAR(Date)

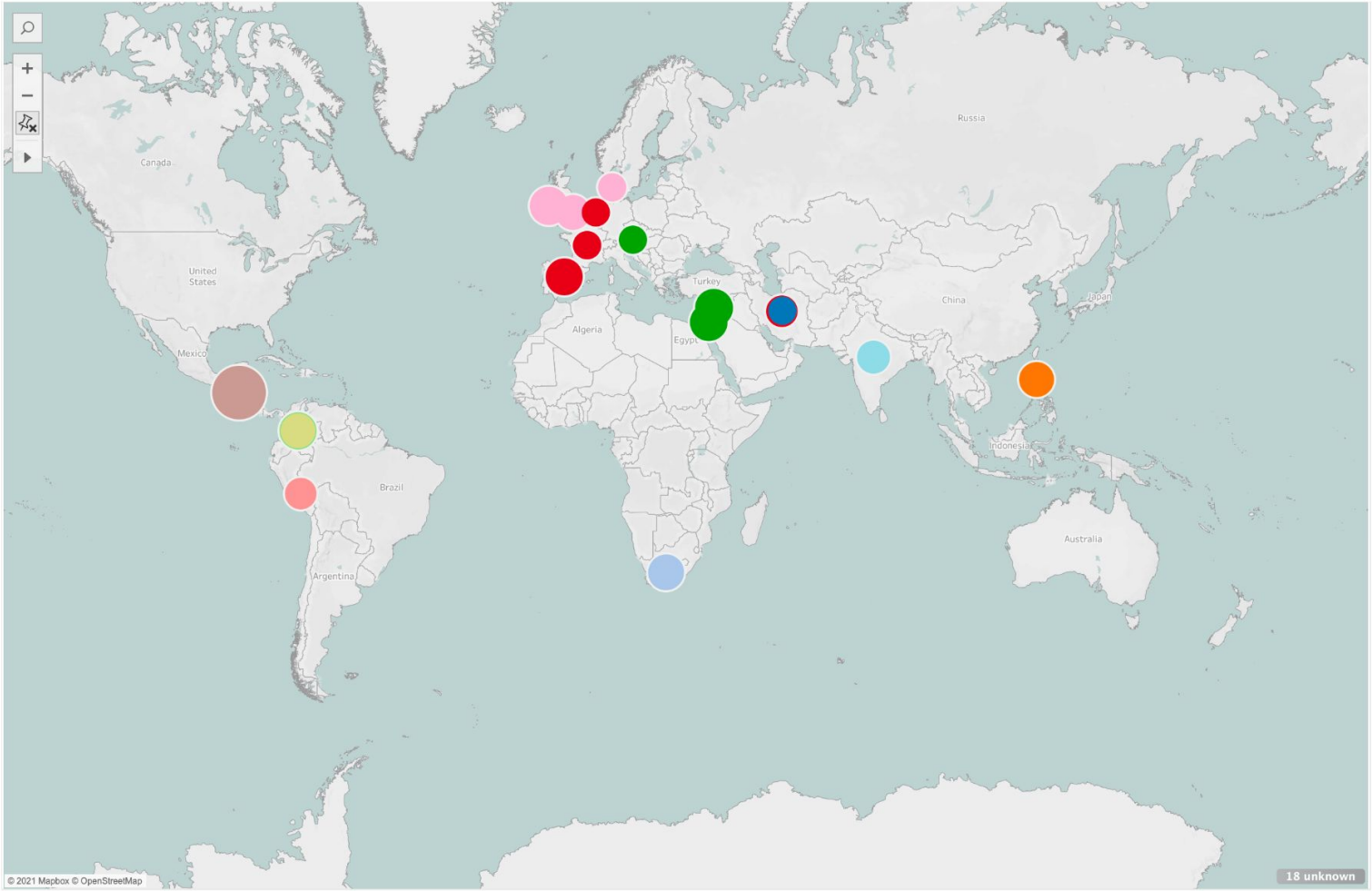
1975

Show history

Nkill

0.0
100.0
200.0
329.0

Map Showing No. of Kills per Group (every 5yrs) - 1980



Gname

- Abu Sayyaf Group ..
- African National C..
- Al-Qaida in Iraq
- Al-Qaida in the Ar..
- Al-Shabaab
- Basque Fatherlan..
- Boko Haram
- Communist Party ..
- Corsican National ..
- Donetsk People's ..
- Farabundo Marti ..
- Fulani extremists
- Irish Republican A..
- Islamic State of Ir..
- Kurdistan Worker..
- Liberation Tigers ..
- M-19 (Movement ..
- Manuel Rodriguez..
- Maoists
- Muslim extremists
- National Liberatio..
- New People's Arm..
- Nicaraguan Demo..
- Palestinians
- Revolutionary Ar..
- Separatists
- Shining Path (SL)
- Sikh Extremists
- Taliban
- Tehrik-i-Taliban P..

YEAR(Date)

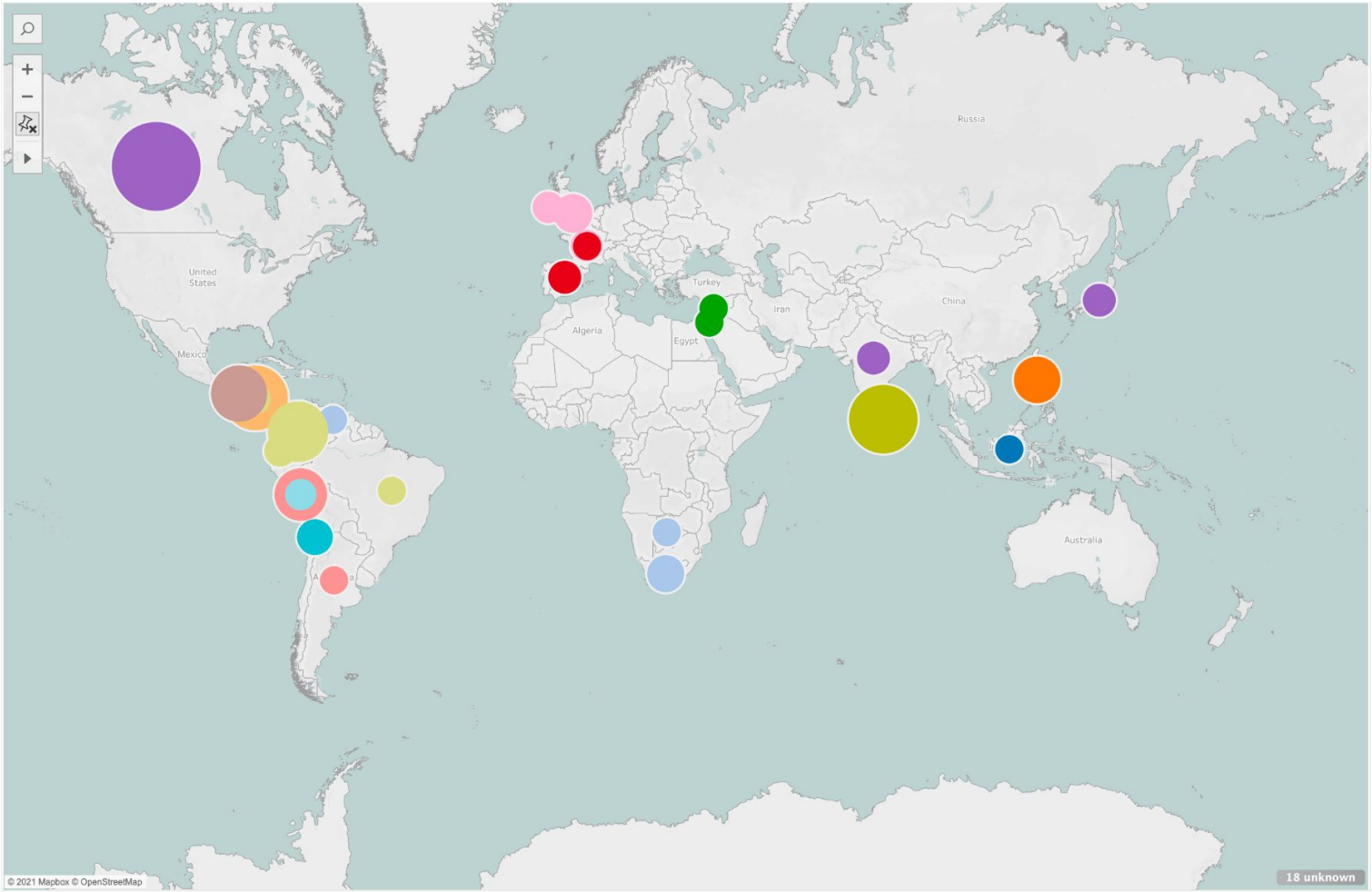
1980

Show history

Nkill

0.0
100.0
200.0
329.0

Map Showing No. of Kills per Group (every 5yrs) - 1985



Gname

- Abu Sayyaf Group ..
- African National C..
- Al-Qaida in Iraq
- Al-Qaida in the Ar..
- Al-Shabaab
- Basque Fatherlan..
- Boko Haram
- Communist Party ..
- Corsican National ..
- Donetsk People's ..
- Farabundo Marti ..
- Fulani extremists
- Irish Republican A..
- Islamic State of Ir..
- Kurdistan Worker..
- Liberation Tigers ..
- M-19 (Movement ..
- Manuel Rodriguez..
- Maoists
- Muslim extremists
- National Liberatio..
- New People's Arm..
- Nicaraguan Demo..
- Palestinians
- Revolutionary Ar..
- Separatists
- Shining Path (SL)
- Sikh Extremists
- Taliban
- Tehrik-i-Taliban P..

YEAR(Date)

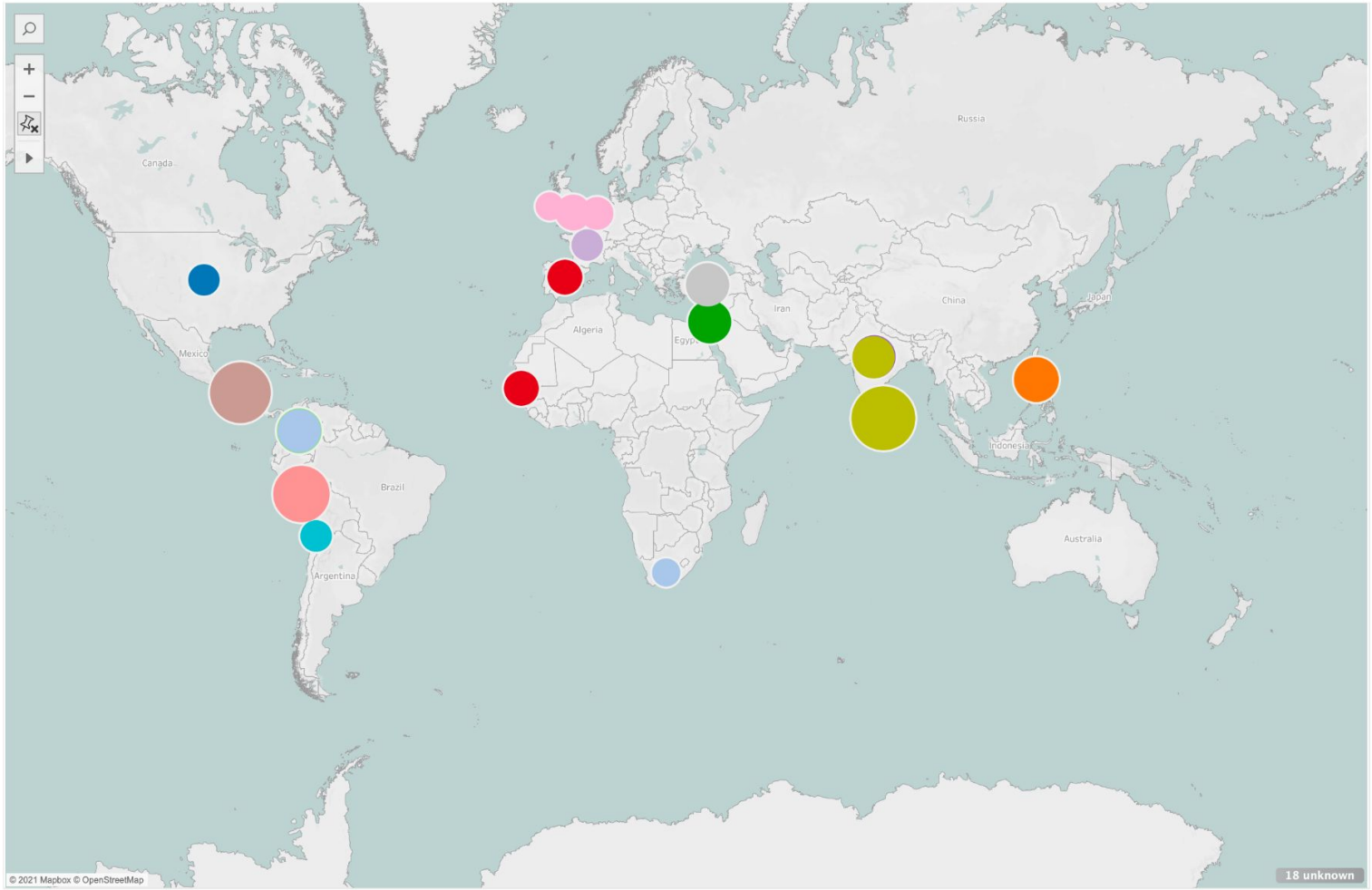
1985

Show history

Nkill

0.0
100.0
200.0
329.0

Map Showing No. of Kills per Group (every 5yrs) - 1990



Gname

- Abu Sayyaf Group ..
- African National C..
- Al-Qaida in Iraq
- Al-Qaida in the Ar..
- Al-Shabaab
- Basque Fatherlan..
- Boko Haram
- Communist Party ..
- Corsican National ..
- Donetsk People's ..
- Farabundo Marti ..
- Fulani extremists
- Irish Republican A..
- Islamic State of Ir..
- Kurdistan Worker..
- Liberation Tigers ..
- M-19 (Movement ..
- Manuel Rodriguez..
- Maoists
- Muslim extremists
- National Liberatio..
- New People's Arm..
- Nicaraguan Demo..
- Palestinians
- Revolutionary Ar..
- Separatists
- Shining Path (SL)
- Sikh Extremists
- Taliban
- Tehrik-i-Taliban P..

YEAR(Date)

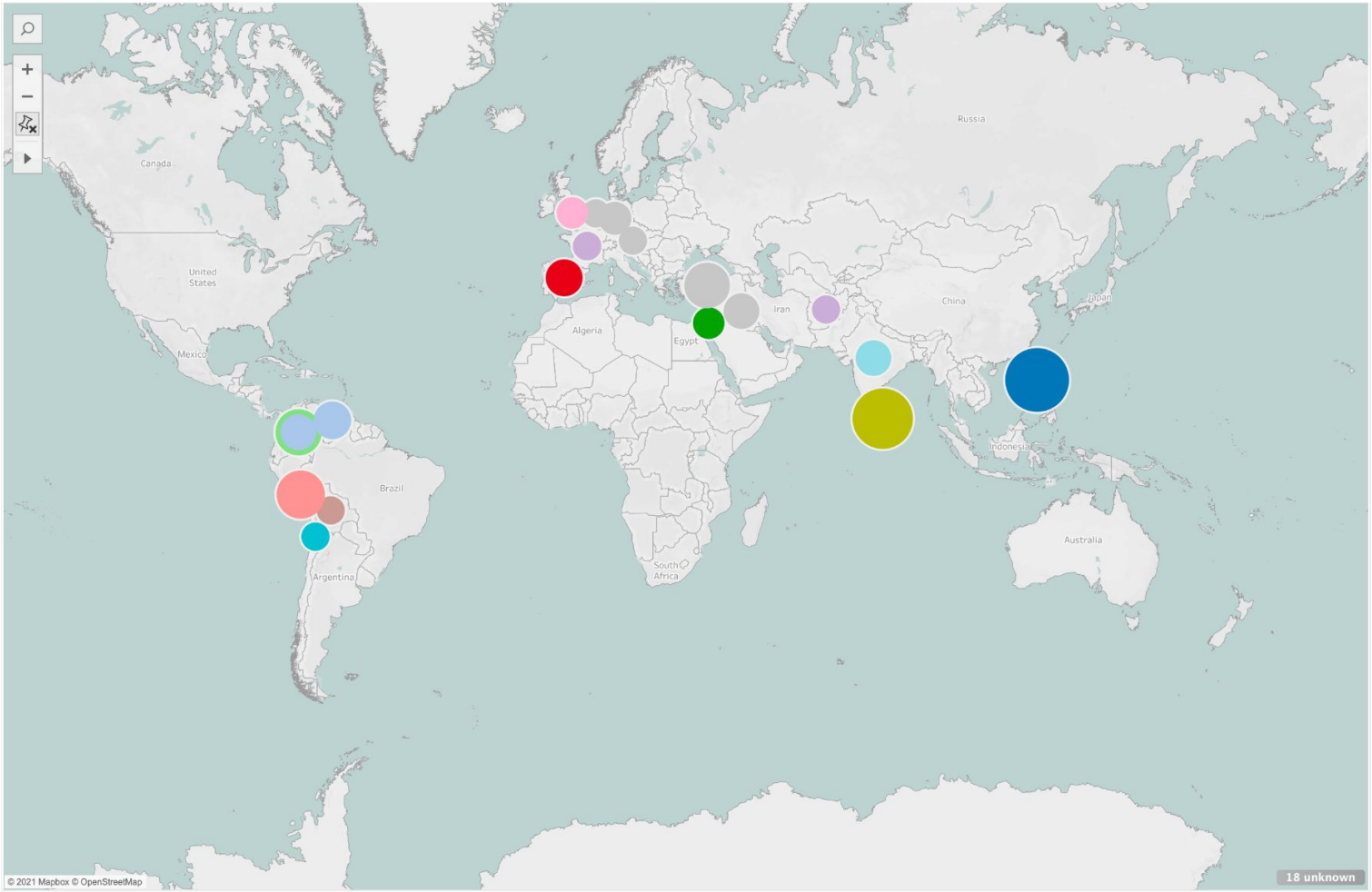
1990

Show history

Nkill

0.0
100.0
200.0
329.0

Map Showing No. of Kills per Group (every 5yrs) - 1995



Gname

- Abu Sayyaf Group ..
- African National C..
- Al-Qaida in Iraq
- Al-Qaida in the Ar..
- Al-Shabaab
- Basque Fatherlan..
- Boko Haram
- Communist Party ..
- Corsican National ..
- Donetsk People's ..
- Farabundo Marti ..
- Fulani extremists
- Irish Republican A..
- Islamic State of Ir..
- Kurdistan Worker..
- Liberation Tigers ..
- M-19 (Movement ..
- Manuel Rodriguez..
- Maoists
- Muslim extremists
- National Liberatio..
- New People's Arm..
- Nicaraguan Demo..
- Palestinians
- Revolutionary Ar..
- Separatists
- Shining Path (SL)
- Sikh Extremists
- Taliban
- Tehrik-i-Taliban P..

YEAR(Date)

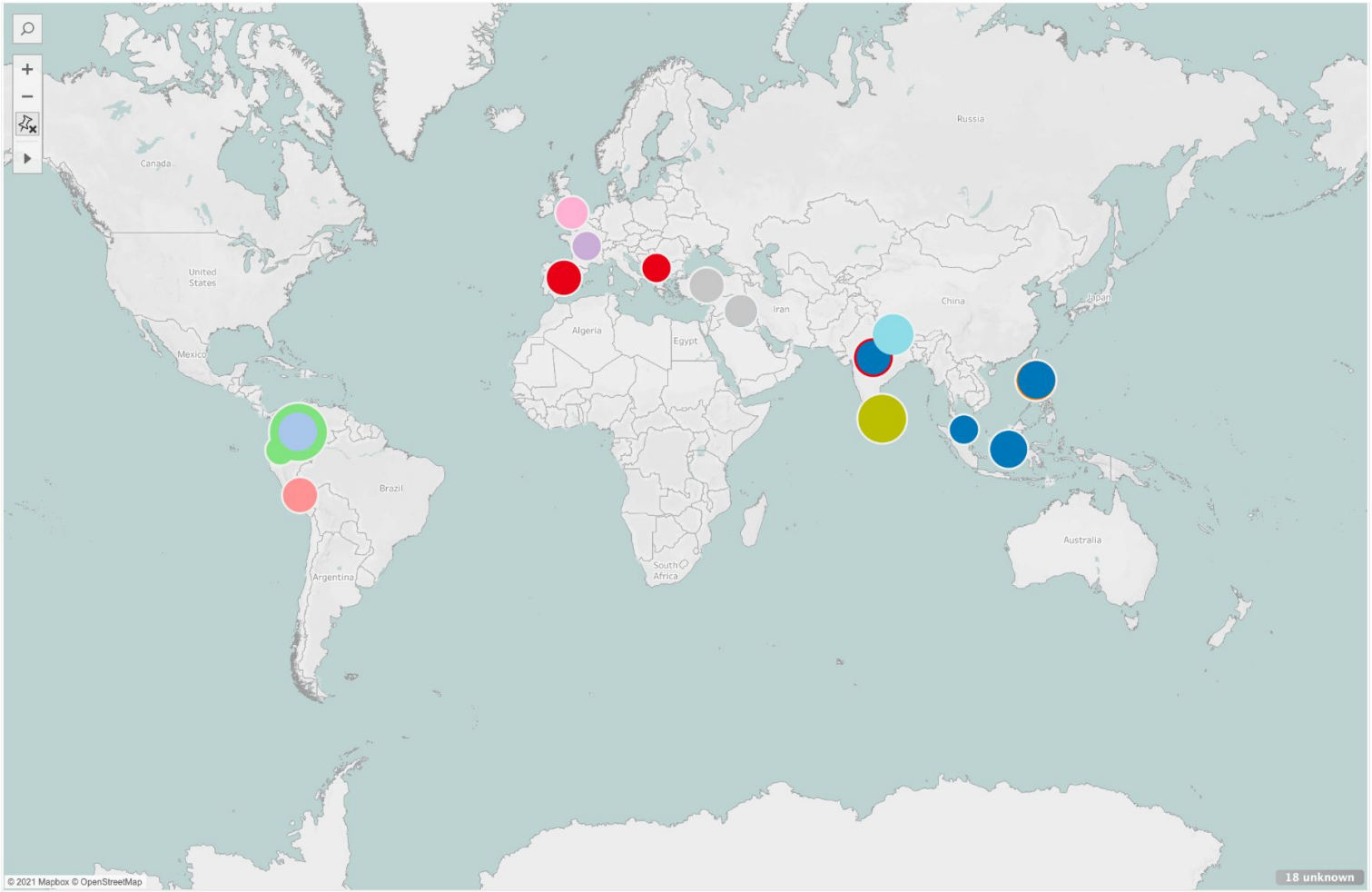
1995

Show history

Nkill

0.0
100.0
200.0
329.0

Map Showing No. of Kills per Group (every 5yrs) - 2000



Gname

- Abu Sayyaf Group ..
- African National C..
- Al-Qaida in Iraq
- Al-Qaida in the Ar..
- Al-Shabaab
- Basque Fatherlan..
- Boko Haram
- Communist Party ..
- Corsican National ..
- Donetsk People's ..
- Farabundo Marti ..
- Fulani extremists
- Irish Republican A..
- Islamic State of Ir..
- Kurdistan Worker..
- Liberation Tigers ..
- M-19 (Movement ..
- Manuel Rodriguez..
- Maoists
- Muslim extremists
- National Liberatio..
- New People's Arm..
- Nicaraguan Demo..
- Palestinians
- Revolutionary Ar..
- Separatists
- Shining Path (SL)
- Sikh Extremists
- Taliban
- Tehrik-i-Taliban P..

YEAR(Date)

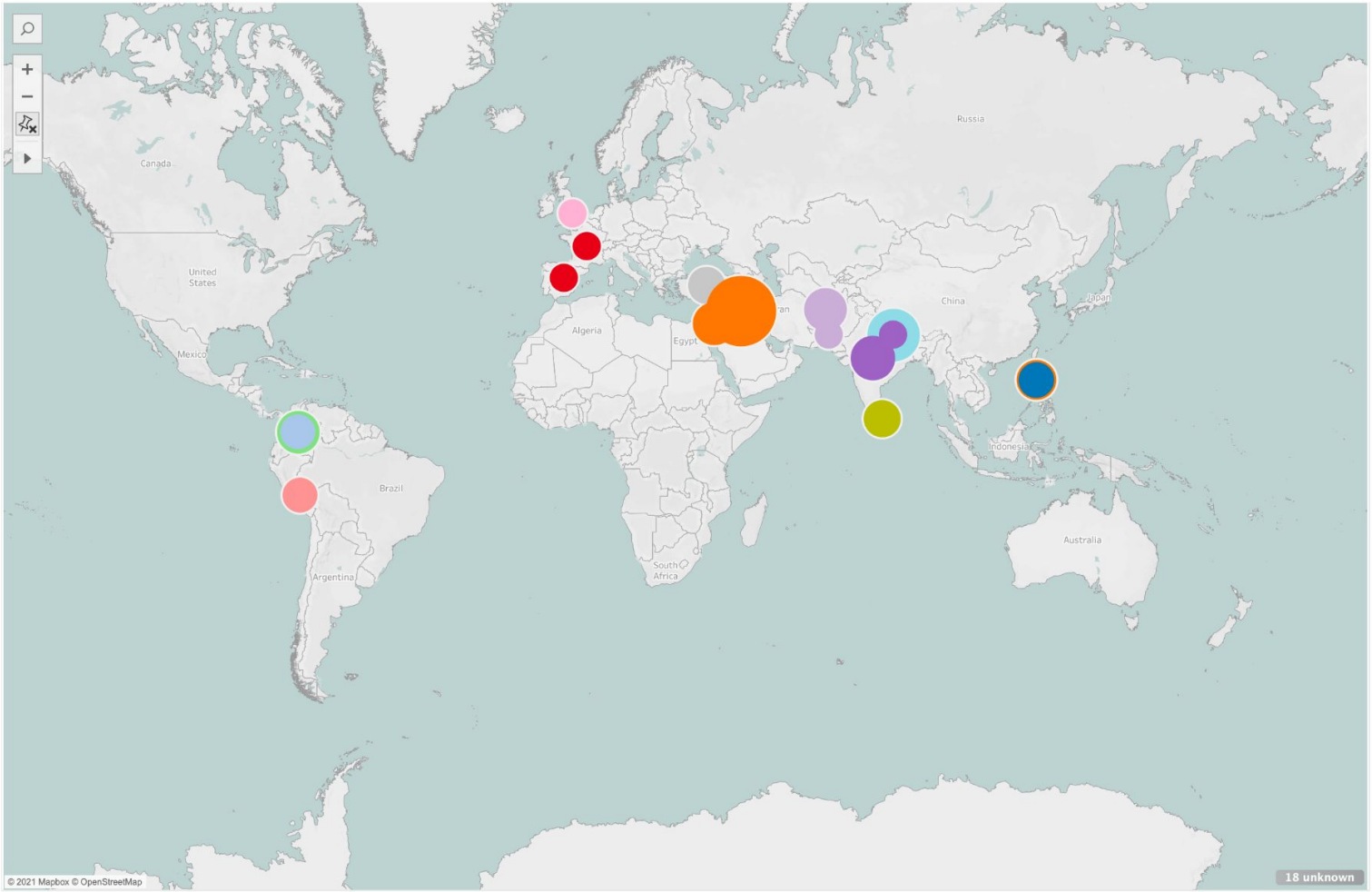
2000

Show history

Nkill

0.0
100.0
200.0
329.0

Map Showing No. of Kills per Group (every 5yrs) - 2005



Gname

- Abu Sayyaf Group ..
- African National C..
- Al-Qaida in Iraq
- Al-Qaida in the Ar..
- Al-Shabaab
- Basque Fatherlan..
- Boko Haram
- Communist Party ..
- Corsican National ..
- Donetsk People's ..
- Farabundo Marti ..
- Fulani extremists
- Irish Republican A..
- Islamic State of Ir..
- Kurdistan Worker..
- Liberation Tigers ..
- M-19 (Movement ..
- Manuel Rodriguez..
- Maoists
- Muslim extremists
- National Liberatio..
- New People's Arm..
- Nicaraguan Demo..
- Palestinians
- Revolutionary Ar..
- Separatists
- Shining Path (SL)
- Sikh Extremists
- Taliban
- Tehrik-i-Taliban P..

YEAR(Date)

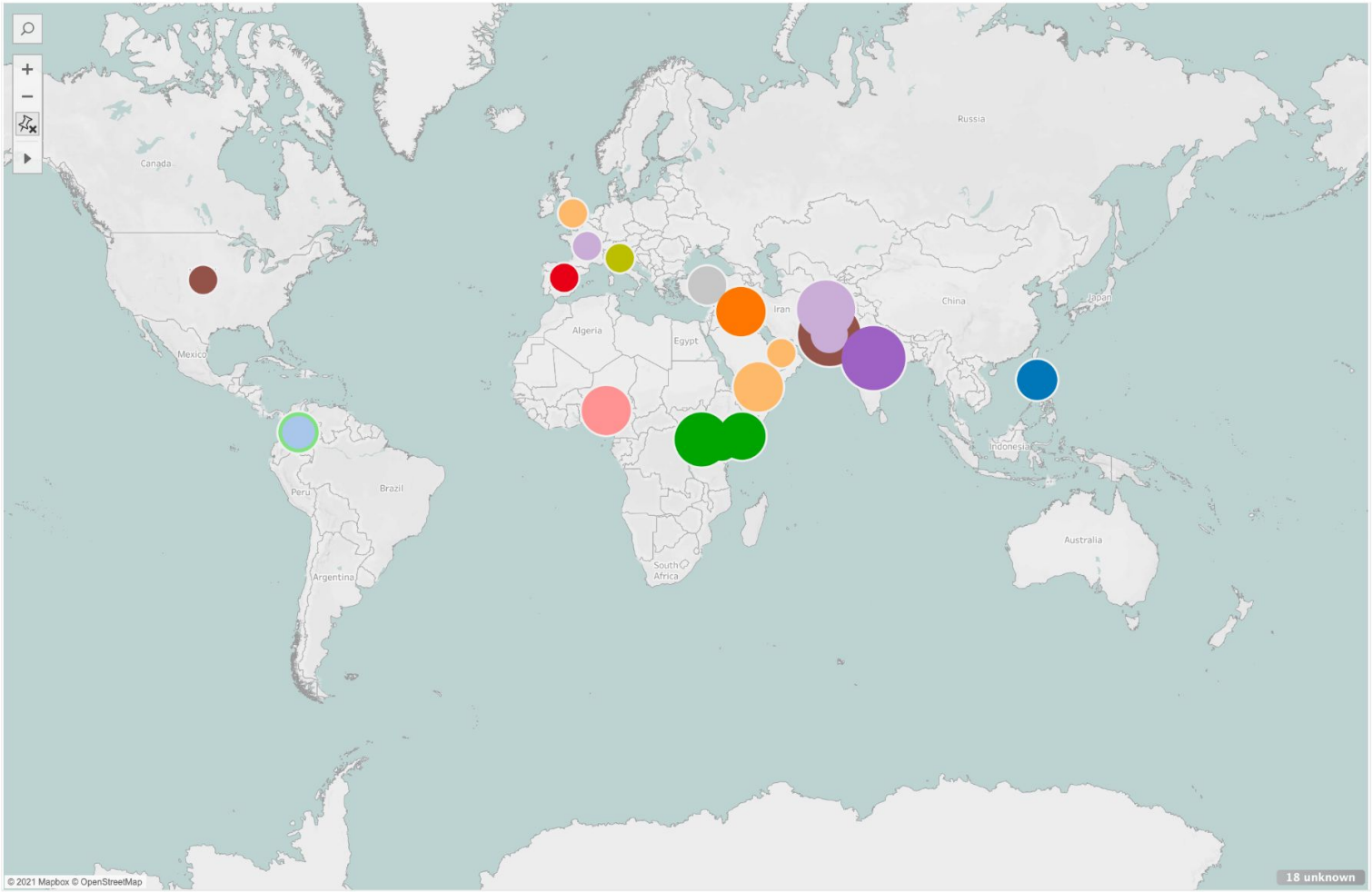
2005

Show history

Nkill

0.0
100.0
200.0
329.0

Map Showing No. of Kills per Group (every 5yrs) - 2010



Gname

- Abu Sayyaf Group ..
- African National C..
- Al-Qaida in Iraq
- Al-Qaida in the Ar..
- Al-Shabaab
- Basque Fatherlan..
- Boko Haram
- Communist Party ..
- Corsican National ..
- Donetsk People's ..
- Farabundo Marti ..
- Fulani extremists
- Irish Republican A..
- Islamic State of Ir..
- Kurdistan Worker..
- Liberation Tigers ..
- M-19 (Movement ..
- Manuel Rodriguez..
- Maoists
- Muslim extremists
- National Liberatio..
- New People's Arm..
- Nicaraguan Demo..
- Palestinians
- Revolutionary Ar..
- Separatists
- Shining Path (SL)
- Sikh Extremists
- Taliban
- Tehrik-i-Taliban P..

YEAR(Date)

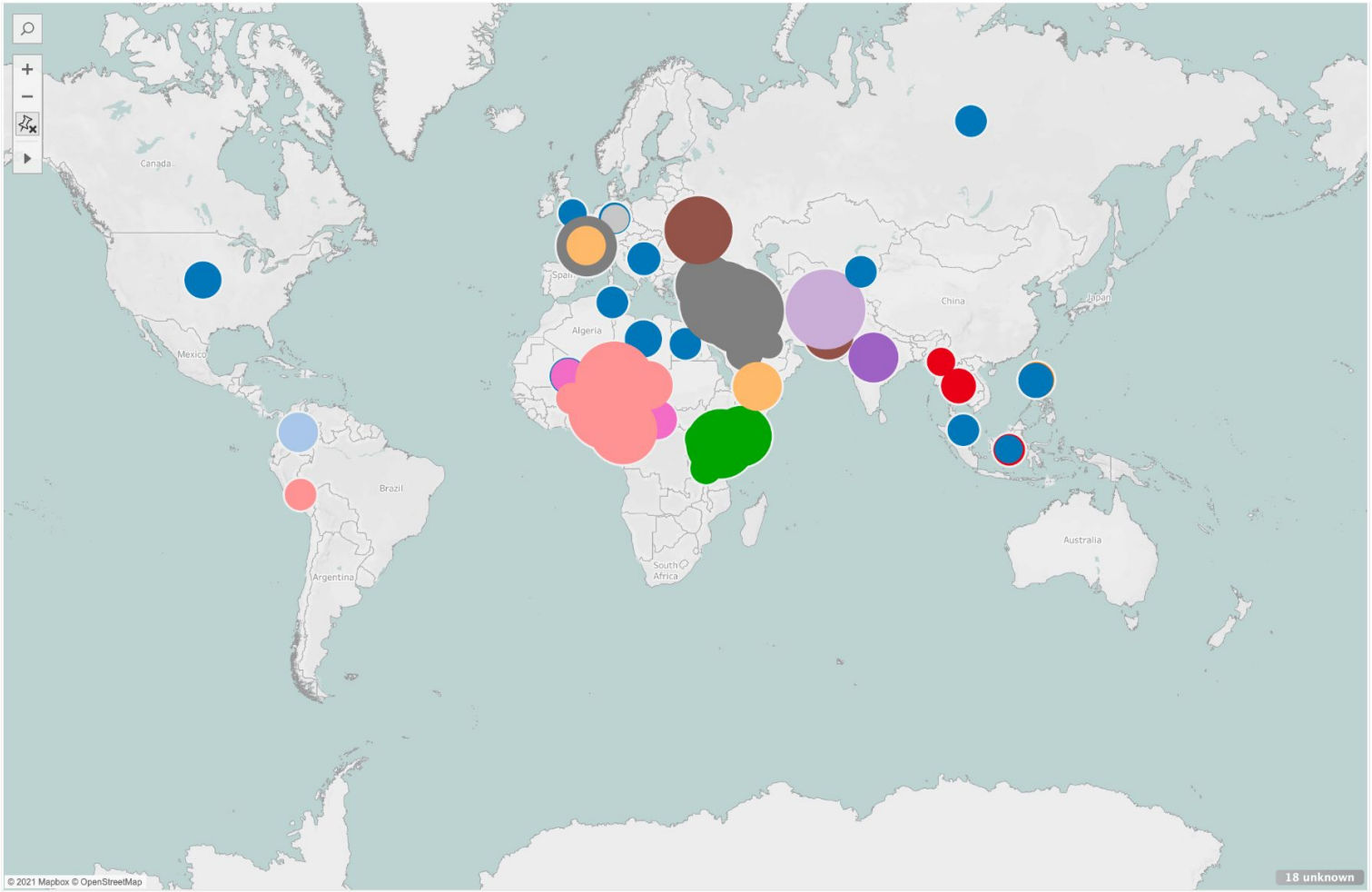
2010

Show history

Nkill

0.0
100.0
200.0
329.0

Map Showing No. of Kills per Group (every 5yrs) - 2015



Gname

- Abu Sayyaf Group ..
- African National C..
- Al-Qaida in Iraq
- Al-Qaida in the Ar..
- Al-Shabaab
- Basque Fatherlan..
- Boko Haram
- Communist Party ..
- Corsican National ..
- Donetsk People's ..
- Farabundo Marti ..
- Fulani extremists
- Irish Republican A..
- Islamic State of Ir..
- Kurdistan Worker..
- Liberation Tigers ..
- M-19 (Movement ..
- Manuel Rodriguez..
- Maoists
- Muslim extremists
- National Liberatio..
- New People's Arm..
- Nicaraguan Demo..
- Palestinians
- Revolutionary Ar..
- Separatists
- Shining Path (SL)
- Sikh Extremists
- Taliban
- Tehrik-i-Taliban P..

YEAR(Date)

2015

Show history

Nkill

0.0
100.0
200.0
329.0

Preprocessing Stages

‘Related’ Column

In order to clearly see the amount of attacks that were related to each observation I edited the column ‘related’ to show the number of related events.

Categorical Columns

So that our model knew the values we were dealing with in the 39 columns were ‘categorical’ we made sure to label these columns with .astype(‘category’).

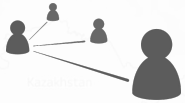
Performed PCA

In order to reduce the amount of columns (which perhaps held data that was irrelevant or held no importance) I performed a test on the dataset to only return the columns to me that held the most relevant information, once performed it gave me a subset of just 10 columns.

Performed Chi Square Test

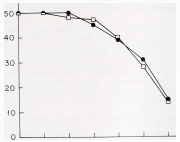
As plotting correlations can be difficult with categorical data I also ran a ‘Chi Square Test’ which returned to me the level of ‘dependency’ each column had on the target variable. Those who had little dependency (and therefore no real relationship to tell) could be dropped.

Final Data



CAT:

300 +



Final Data

Cleaned data = 41 columns.
Cleaned data = 57990 observations.

Subset A

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 57990 entries, 0 to 57989
Data columns (total 10 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   year      57990 non-null  category
 1   month     57990 non-null  category
 2   day       57990 non-null  category
 3   extended  57990 non-null  category
 4   country   57990 non-null  category
 5   region    57990 non-null  category
 6   vicinity  57990 non-null  category
 7   crit1     57990 non-null  category
 8   crit2     57990 non-null  category
 9   crit3     57990 non-null  category
dtypes: category(10)
```

After running the PCA test.

Subset B

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 57990 entries, 39 to 181687
Data columns (total 25 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   year           57990 non-null  category
 1   extended       57990 non-null  category
 2   country        57990 non-null  category
 3   region         57990 non-null  category
 4   vicinity       57990 non-null  category
 5   doubtterr      57990 non-null  category
 6   multiple       57990 non-null  category
 7   suicide        57990 non-null  category
 8   attacktype1    57990 non-null  category
 9   targtype1      57990 non-null  category
 10  targsubtype1   57990 non-null  category
 11  natlty1        57990 non-null  category
 12  targsubtype2   57990 non-null  category
 13  weaptype1      57990 non-null  category
 14  weapsubtype1   57990 non-null  category
 15  weapsubtype2   57990 non-null  category
 16  nkill          57990 non-null  float64
 17  property       57990 non-null  category
 18  ishostkid      57990 non-null  category
 19  ransom         57990 non-null  category
 20  INT_LOG        57990 non-null  category
 21  INT_IDEO       57990 non-null  category
 22  INT_MISC       57990 non-null  category
 23  INT_ANY        57990 non-null  category
 24  related        57990 non-null  category
dtypes: category(24), float64(1)
```

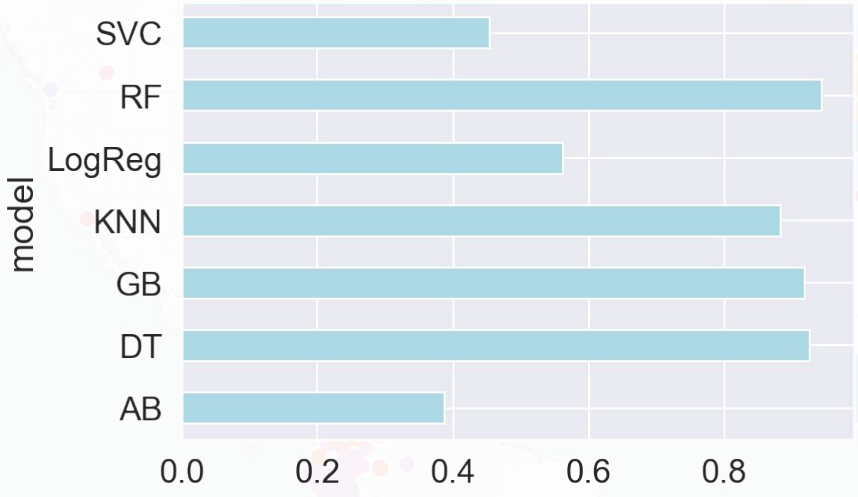
After running the Chi Squared test.

Subset C

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 57990 entries, 0 to 57989
Data columns (total 22 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   year           57990 non-null  category
 1   extended       57990 non-null  category
 2   country        57990 non-null  category
 3   region         57990 non-null  category
 4   vicinity       57990 non-null  category
 5   doubtterr      57990 non-null  category
 6   multiple       57990 non-null  category
 7   suicide        57990 non-null  category
 8   attacktype1    57990 non-null  category
 9   targtype1      57990 non-null  category
 10  targsubtype1   57990 non-null  category
 11  natlty1        57990 non-null  category
 12  targsubtype2   57990 non-null  category
 13  weaptype1      57990 non-null  category
 14  weapsubtype1   57990 non-null  category
 15  weapsubtype2   57990 non-null  category
 16  nkill          57990 non-null  float64
 17  property       57990 non-null  category
 18  ishostkid      57990 non-null  category
 19  ransom         57990 non-null  category
 20  INT_LOG        57990 non-null  category
 21  INT_IDEO       57990 non-null  category
dtypes: category(21), float64(1)
```

After running the Chi Squared test then a PCA test.

Models Tested



** RF = 0.94 (94% accuracy)

Baseline = 12%

Logistic Regression

Decision Tree Classifier

Random Forest Classifier **

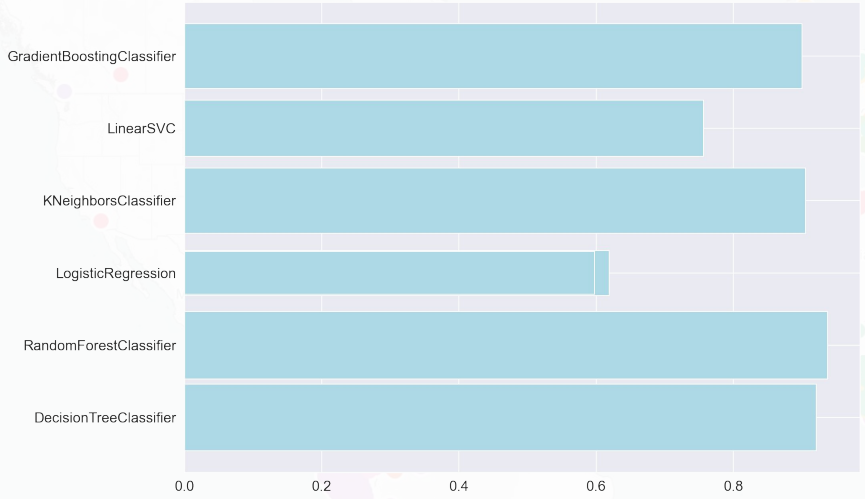
KNeighbors Classifier

LinearSVC

AdaBoost Classifier

Gradient Boosting Classifier

Model Optimisation



** RF = 0.94 (94% accuracy)

Baseline = 12%

Using the following tools from the SKlearn library

-pipeline

-SMOTE

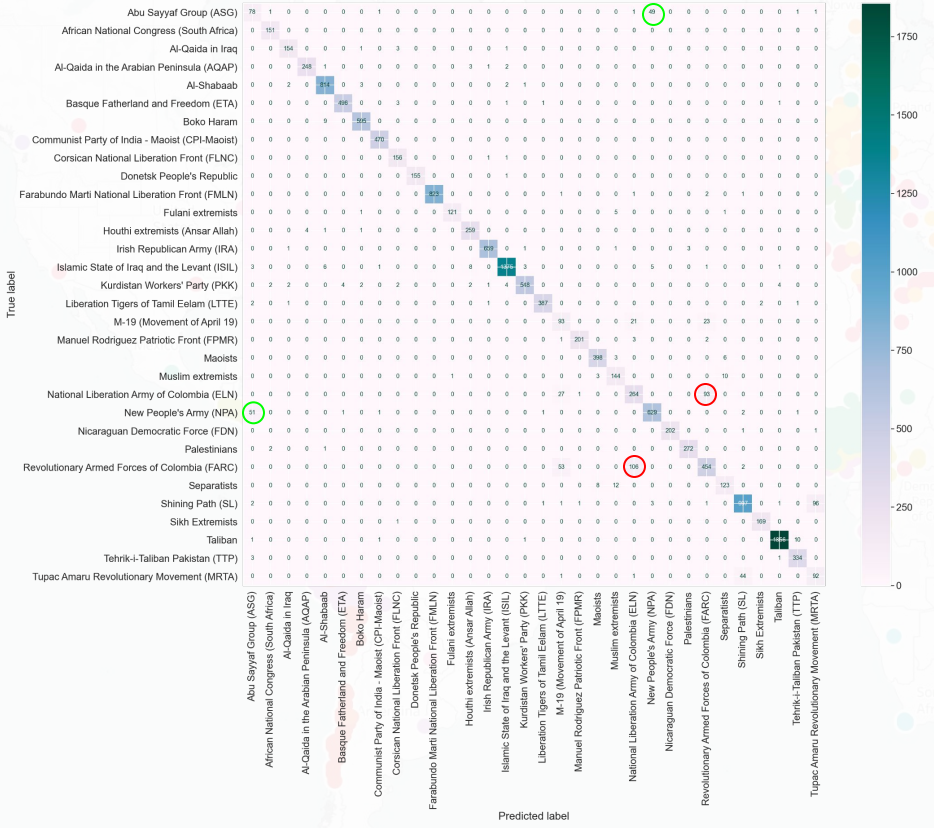
-gridsearch

-cross-validation

we were able to tune our models to see if we could gain any improvement in our scores, or if our Random Forest Classifier would still perform the best.

Of which it did.

Analysing Best Performing Model



The Random Forest model I have evaluated as my most trusted was derived from subset C - with 22 columns.

```
Cross-validation score: 0.9437137645901477
Test score: 0.9461305007587253
```

It returned an accuracy score of 0.95, or 95%.

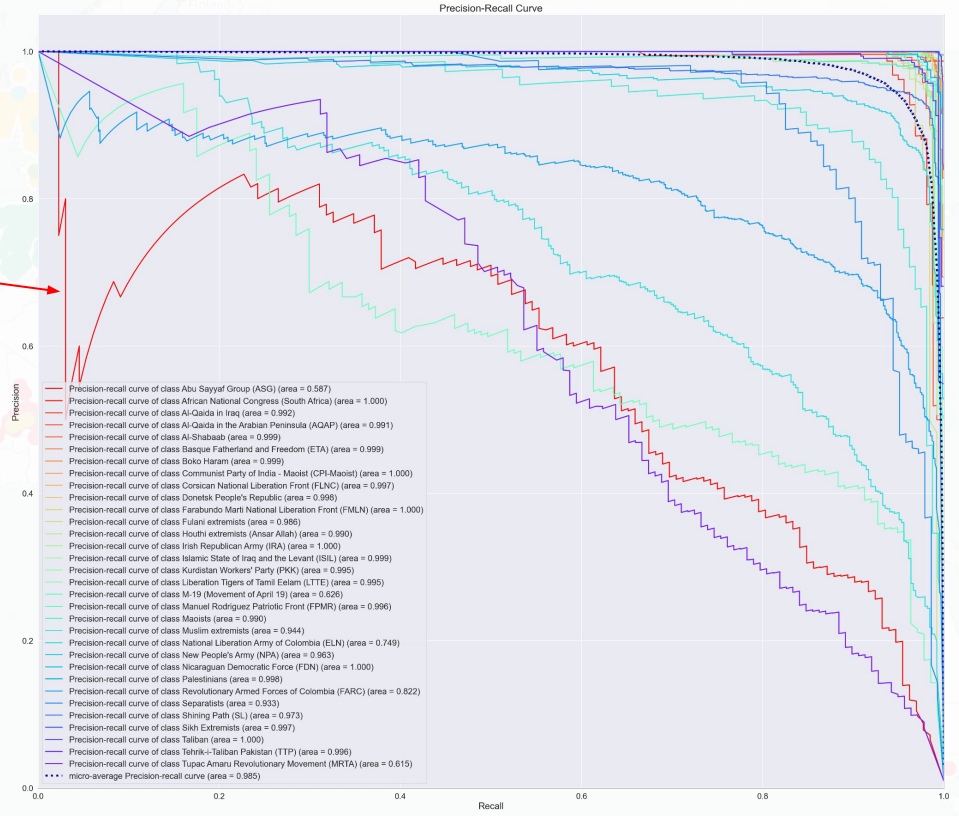
We can see incorrect predictions made for those groups in a underweighted class, such as the Abu Sayyaf Group (ASG) who had the smallest number of occurrences in our dataset.

Analysing Best Performing Model

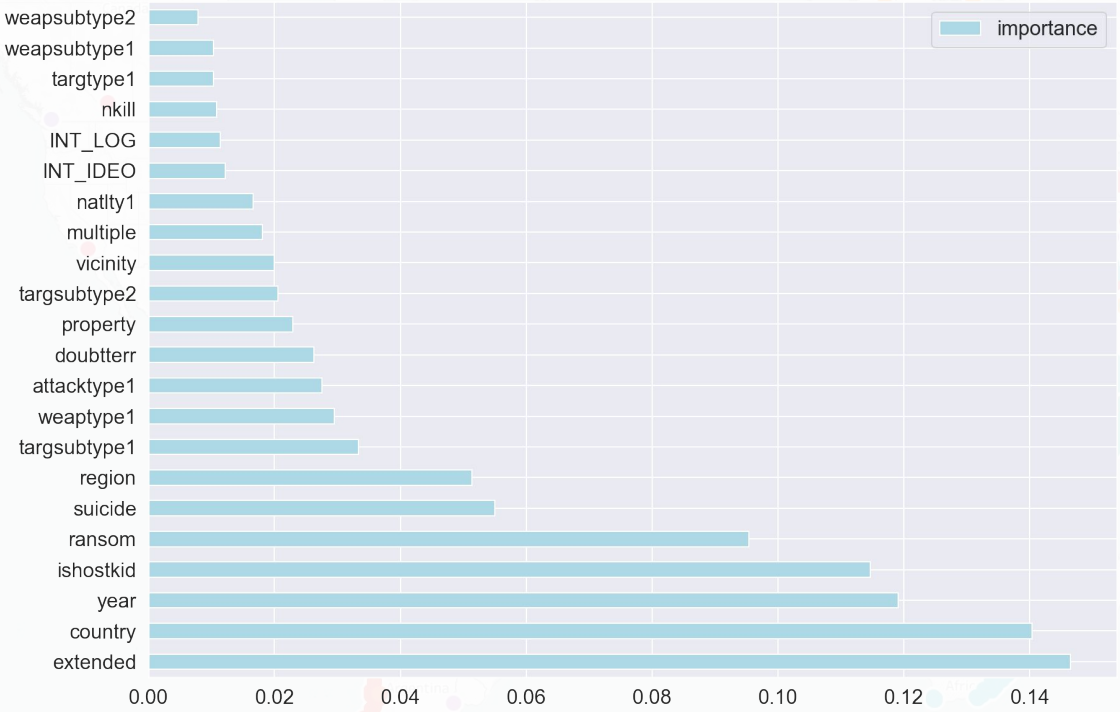
The model struggled to predict with the least represented groups.

For example, the precision and recall line for the Abu Sayyaf Group (ASG), is much lower than than the other better predicted groups.

This reinforces what we saw in the confusion matrix above that this group is being incorrectly predicted as others, and other attacks are being incorrectly predicted as being performed by them.



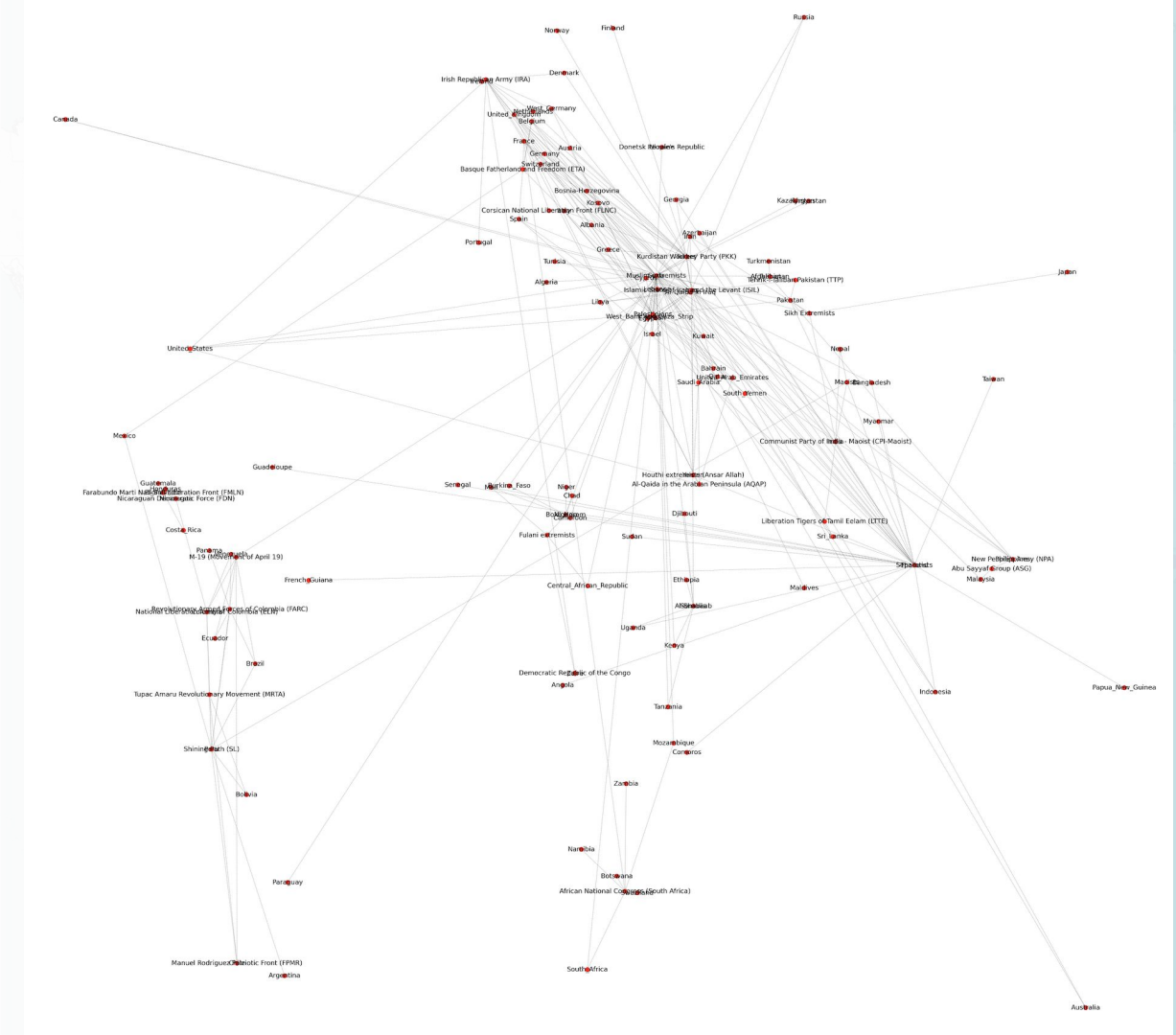
Feature Importances



Our top predictors were:

- Extended (did the event last more than 24hours).
- Country
- Year

For the other models we ran we did find other features came out on top but the main feature that ranked highly on most of them was *Country*.

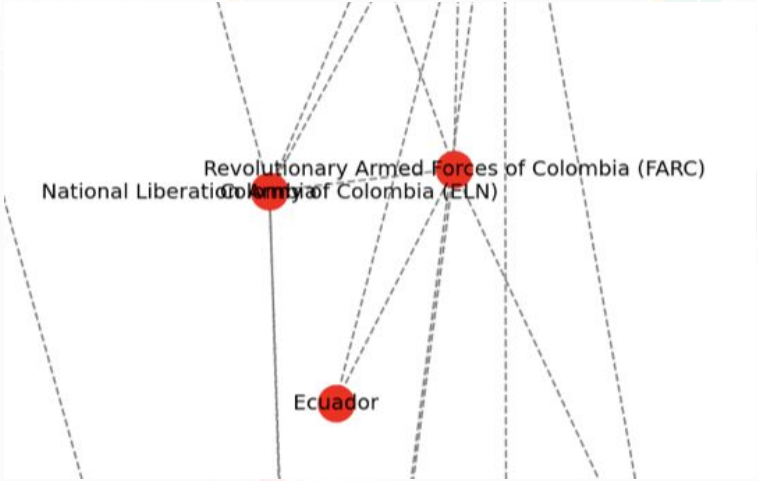


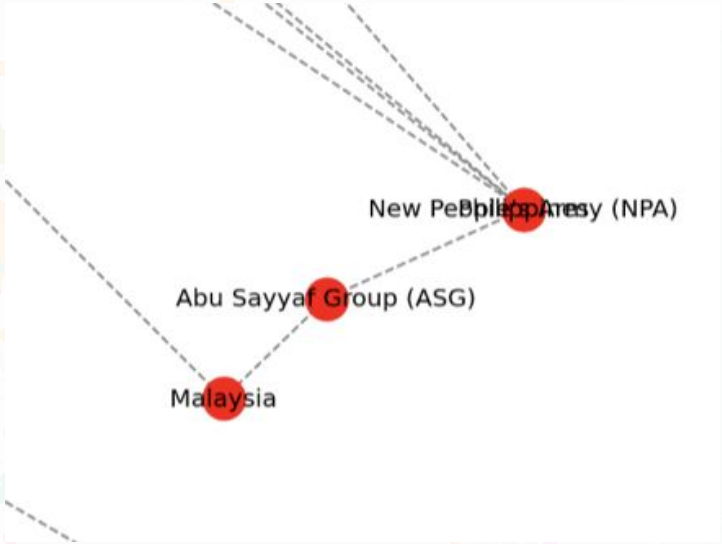
What about the incorrect predictions?

Exploring The Network

```
shortest_connection(B, 'National Liberation Army of Colombia (ELN)',
                       'Revolutionary Armed Forces of Colombia (FARC)')
```

```
The Network Path Length Between Them Is: 2
Shortest Connection Between Them Is: ['National Liberation Army of Colombia (ELN)', 'Colombia', 'Revolutionar
y Armed Forces of Colombia (FARC)']
```

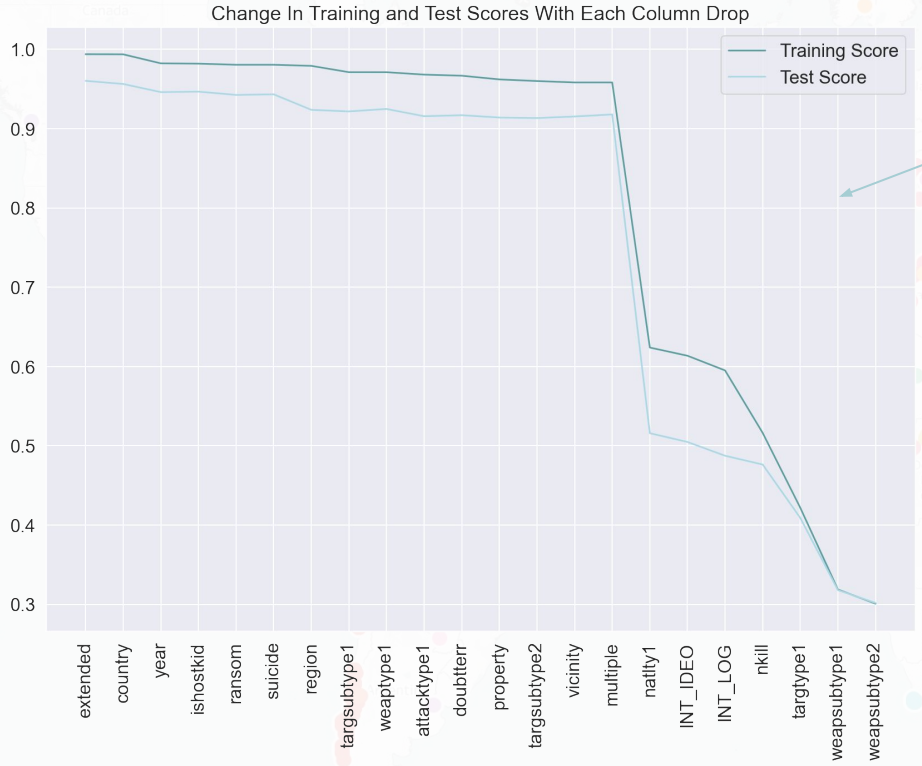





```
shortest_connection(B, 'Abu Sayyaf Group (ASG)', "New People's Army (NPA)")
```

```
The Network Path Length Between Them Is: 2
Shortest Connection Between Them Is: ['Abu Sayyaf Group (ASG)', 'Philippines', "New People's Army (NPA)"]
```

Model Overfitting



Scores after dropping each column by feature importance.

Scores after dropping just the single column.

column_name	train_score	test_score	cv_score
extended	0.993976	0.960822	0.959211
country	0.993884	0.955442	0.953003
year	0.983238	0.951855	0.948956
ishostkid	0.993884	0.961443	0.959602
ransom	0.993976	0.961098	0.959142
suicide	0.994068	0.961098	0.958797
region	0.994068	0.957925	0.954153
targsubtype1	0.991033	0.960891	0.957394
weaptype1	0.994068	0.960822	0.959119
attacktype1	0.993861	0.960822	0.959027
doubtterr	0.994068	0.961305	0.958590
property	0.993608	0.961443	0.958843
targsubtype2	0.993976	0.961236	0.958935
vicinity	0.993815	0.961443	0.959096
multiple	0.994068	0.961788	0.959211
natlty1	0.993953	0.960753	0.957555
INT_IDEO	0.994068	0.960339	0.958820
INT_LOG	0.994022	0.960063	0.958337
nkill	0.990481	0.961167	0.959050
targtype1	0.993999	0.961305	0.958475
weapsubtype1	0.993332	0.961029	0.960085
weapsubtype2	0.993930	0.961098	0.958429

Making Predictions

Take a look at this observation, which was originally marked as 'Unknown'.....

Which terrorist organisation would you predict from the data given?

- Took place in Afghanistan.
- Hostage Taking (Kidnapping).
- Non-suicidal.
- Year 2013.
- Not extended over 24 hours.
- The attack was part of a multiple incident.
- Main target was a telecommunications facility.
- Main weapons used were firearms.
- The incident resulted in property damage.

```
model.predict(unknown_std)
```

```
array(['Islamic State of Iraq and the Levant (ISIL)', 'Taliban',
       'Islamic State of Iraq and the Levant (ISIL)',
       'Islamic State of Iraq and the Levant (ISIL)', 'Maoists',
       'Islamic State of Iraq and the Levant (ISIL)', 'Maoists',
       'Separatists', 'Houthi extremists (Ansar Allah)', 'Taliban',
       "Donetsk People's Republic",
       'Islamic State of Iraq and the Levant (ISIL)', 'Taliban',
       'Taliban', 'Taliban', 'Taliban', 'Taliban', 'Taliban',
       'Communist Party of India - Maoist (CPI-Maoist)',
       'Revolutionary Armed Forces of Colombia (FARC)',
       'Communist Party of India - Maoist (CPI-Maoist)',
       'Irish Republican Army (IRA)',
       'Communist Party of India - Maoist (CPI-Maoist)',
       'Communist Party of India - Maoist (CPI-Maoist)',
       'Communist Party of India - Maoist (CPI-Maoist)'], dtype=object)
```

Did you predict the same?

Summary

Hypothesis

Our hypothesis was correct, we were able to predict the name of the terrorist organisation who performed an attack from the data provided.

Score Metric

Our baseline score or ‘best guess’ was a 12% accuracy, our model predicted at a 95% accuracy which is a huge improvement.

Features

In order to predict who performed the attack our model found that country, year and whether or not the duration of an incident extended more than 24 hours were the most important characteristics.

Project Extensions

❑ Time Series model

I would like to apply a time series model to the data for ‘number of kills’, to see if we are able to find trends in the dates as to when larger and more deadly attacks are more likely to happen.

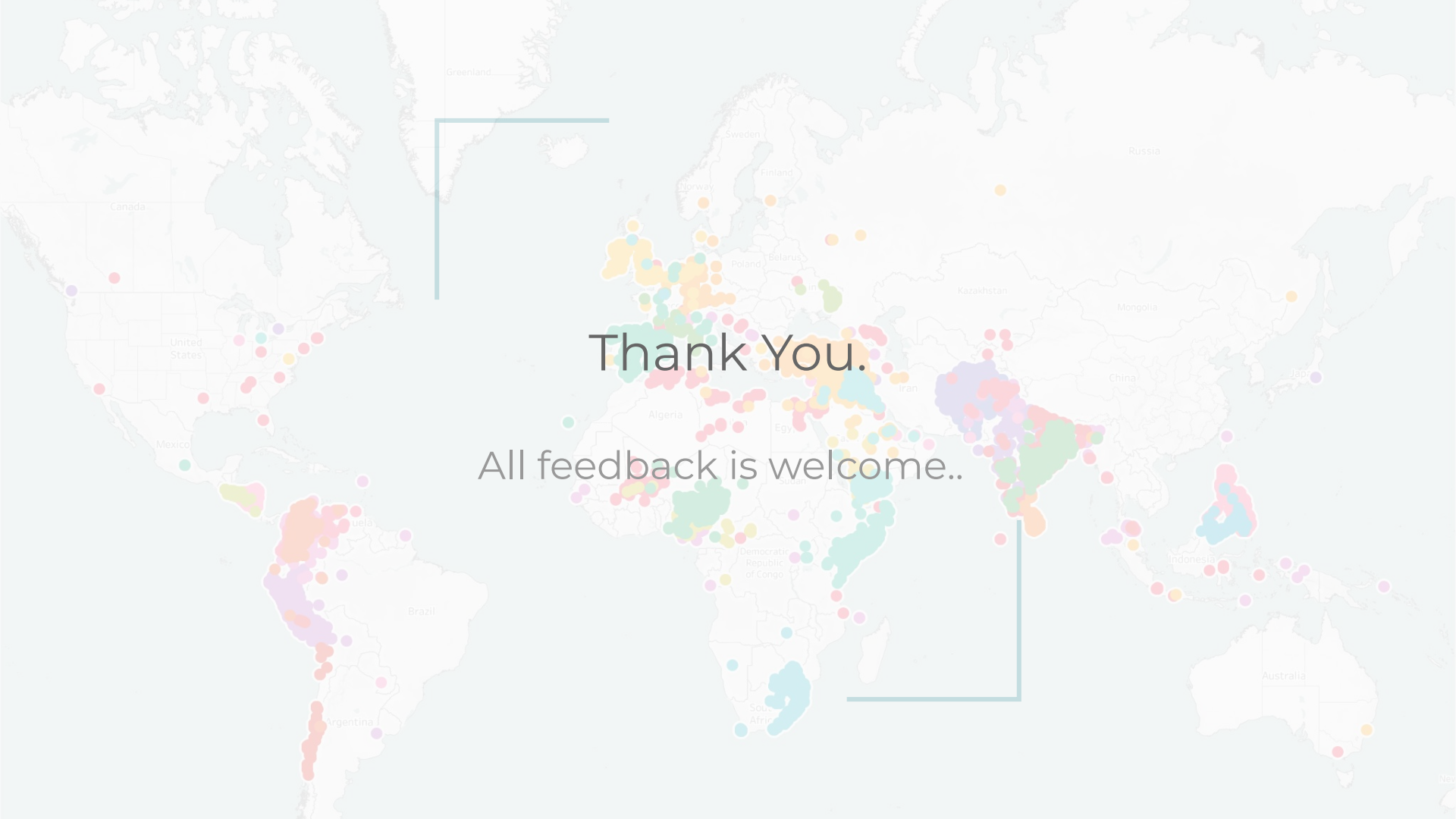
❑ Applying the initial hypothesis of movie data

In my original hypothesis I had wanted to include information on movies that were released around the date of the attacks. I have already scraped information from IMDB and I would apply this new data to look at the relationship between movies and terrorist attacks.

	domestic	domesticLifetimeGross	foreign	foreignLifetimeGross	rank	title	worldwideLifetimeGross	year	id
0	26.7%	$760,507,625	73.3%	$2,086,738,578	1	Avatar	$2,847,246,203	2009	tt0499549
1	30.7%	$858,373,000	69.3%	$1,939,128,328	2	Avengers: Endgame	$2,797,501,328	2019	tt4154796
2	30%	$659,363,944	70%	$1,542,283,320	3	Titanic	$2,201,647,264	1997	tt0120338
3	45.3%	$936,662,225	54.7%	$1,132,859,475	4	Star Wars: Episode VII - The Force Awakens	$2,069,521,700	2015	tt2488496
4	33.1%	$678,815,482	66.9%	$1,369,544,272	5	Avengers: Infinity War	$2,048,359,754	2018	tt4154756
...	...	...	...	...	...	...	...	...	...
195	35.7%	$193,595,521	64.3%	$348,468,325	196	Madagascar	$542,063,846	2005	tt0351283
196	37.5%	$202,807,711	62.5%	$337,648,165	197	World War Z	$540,455,876	2013	tt0816711
197	44%	$237,283,207	56%	$301,700,000	198	Brave	$538,983,207	2012	tt1217209
198	54.4%	$292,753,960	45.6%	$245,621,107	199	Star Wars: Episode V - The Empire Strikes Back	$538,375,067	1980	tt0080684
199	34.1%	$183,135,014	65.9%	$353,279,279	200	The Simpsons Movie	$536,414,293	2007	tt0462538

```
: # here we pull the needed information from IMDB
movies = {'domestic': [],
          'domesticLifetimeGross': [],
          'foreign': [],
          'foreignLifetimeGross': [],
          'rank': [],
          'title': [],
          'worldwideLifetimeGross': [],
          'year': [],
          'id' : []}
```

```
for item in json_['items']:
    try:
        movies['domestic'].append(item['domestic'])
    except:
        movies['domestic'].append('None')
    try:
        movies['domesticLifetimeGross'].append(item['domesticLifetimeGross'])
    except:
        movies['domesticLifetimeGross'].append('None')
    try:
        movies['foreign'].append(item['foreign'])
    except:
        movies['foreign'].append('None')
    try:
        movies['foreignLifetimeGross'].append(item['foreignLifetimeGross'])
    except:
        movies['foreignLifetimeGross'].append('None')
    try:
        movies['rank'].append(item['rank'])
    except:
        movies['rank'].append('None')
    try:
        movies['title'].append(item['title'])
    except:
        movies['title'].append('None')
    try:
        movies['worldwideLifetimeGross'].append(item['worldwideLifetimeGross'])
    except:
        movies['worldwideLifetimeGross'].append('None')
    try:
        movies['year'].append(item['year'])
    except:
        movies['year'].append('None')
    try:
        movies['id'].append(item['id'])
    except:
        movies['id'].append('None')
```


Thank You.
All feedback is welcome..