# Exploring the Connection between Terrorism and Freedom Indices

Jiwon Kim, Tina Krulec

# Abstract

Ever since 9/11 terrorism has become synonymous with transnational terrorsim in popular media. A terrorist has often times become equal to a Muslim. Furthermore, the popular public perception of terrorism is the one of opression. This research project aims to provide broad overview of that. It does not explore this complicated relationship in detail, but rather provides broad overview of how the available data on terrorist attacks correlates with different indices for freedom. We specifically look into freedom of expression, religion, movement and conflicts in general.

Platform where the application runs: NYU Dumbo cluster

# Introduction

## Motivation

- We have lived in a Muslim country for the past 3 years and get asked a lot about terrorism in relation to Islam and the lack of freedom. Hence, we wanted to **explore for ourselves how terrorism relates to the various indices of freedom**.
- There has been a lot of research done with regards to transnational terrorism, but **we were not able to find substantial data on how terrorism relates to freedom in general.**

## Users

- **General public** to inform themselves of the relations, **popular media** to inform themselves when representing terrorism, **researchers** to investigate this more in depth than we were able to.

# Introduction

## Beneficiaries

- **General Public** to explore implicit biases we hold when it comes to talking about terrorism.

## Importance

- We show that there is weak correlation between religion and the acts of terrorism, which is important given the public perceptions of terrorism.
- We show that terrorism is mostly correlated with the conflict itself, which is supported by the existing research.
- Our results support the claim that terrorism is hard to define, hence, we cannot understand it simply as the loss of lives.
- We show that there is a very weak correlation between freedom of press (sign of democratic and non-opressive governments) and acts of terrorism, which go against the popular perceptions of terrorism existing in oppresive regimes.

# Goodness

- We took time in understanding the components of the datasets and the domain insights to frame the various hypotheses.
- Given the multi-dimensional nature of the topic, we made sure to explore multitudes of different factors to gain a fuller insight of the topic.
- We referred to journals and articles to extend the analytics and to provide a context to each results.
- Took extra care in bridging the two datasets in terms of matching up the data and provide integrity in the combined data.

# Data Sources

Global Terrorism Database (GTD) 1970-2018 | 173.5 MB

- A comprehensive dataset on terrorist attacks in the world by Global National Consortium for the Study of Terrorism and Responses to Terrorism for 1970-2018.
- https://gtd.terrorismdata.com/files/gtd-1970-2018/

Human Freedom Index (HFI) 2019 | 672 B

- A dataset on human freedom in the world, using 76 distinct indicators of personal and economic freedom, of 2008-2017.
- https://www.cato.org/human-freedom-index-new

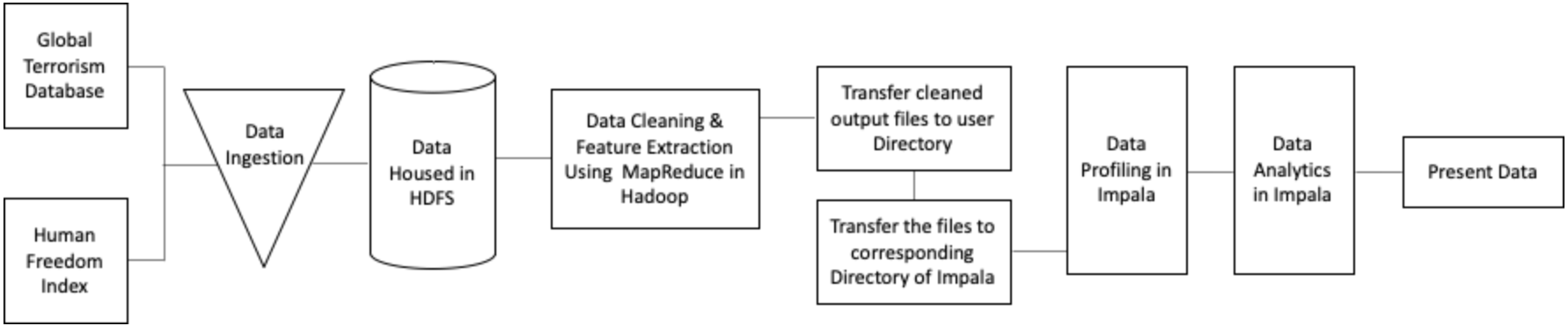
# Data Sample #1: GTD 1970-2018

| event_id | year | month | day | country | region | doubt_terrorist | success | suicide | attack_type | target_type | nationality | num_kill |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 200801010001 | 2008 | 1 | 1 | Philippines | Southeast Asia | 0 | 1 | 0 | Armed Assault | Business | Australia | 0 |
| 200801010002 | 2008 | 1 | 1 | Kenya | Sub-Saharan Africa | 0 | 1 | 0 | Armed Assault | Religious Figures/Institutions | Kenya | 50 |
| 200801010003 | 2008 | 1 | 1 | India | South Asia | 0 | 1 | 0 | Armed Assault | Police | India | 8 |
| 200801010004 | 2008 | 1 | 1 | Sudan | Sub-Saharan Africa | 1 | 1 | 0 | Armed Assault | Government (Diplomatic) | United States | 2 |
| 200801010005 | 2008 | 1 | 1 | Nigeria | Sub-Saharan Africa | 0 | 1 | 0 | Armed Assault | Police | Nigeria | 18 |
| 200801010006 | 2008 | 1 | 1 | Sri Lanka | South Asia | 0 | 1 | 0 | Armed Assault | Government (General) | Sri Lanka | 2 |
| 200801010008 | 2008 | 1 | 1 | Iraq | Middle East & North Africa | 0 | 1 | 1 | Bombing/Explosion | Private Citizens & Property | Iraq | 37 |
| 200801010009 | 2008 | 1 | 1 | Iraq | Middle East & North Africa | 0 | 1 | 0 | Armed Assault | Police | Iraq | 1 |
| 200801010010 | 2008 | 1 | 1 | Iraq | Middle East & North Africa | 0 | 1 | 0 | Assassination | Police | Iraq | 5 |
| 200801010011 | 2008 | 1 | 1 | Iraq | Middle East & North Africa | 0 | 1 | 0 | Armed Assault | Private Citizens & Property | Iraq | 1 |
| 200801010012 | 2008 | 1 | 1 | Iraq | Middle East & North Africa | 1 | 1 | 0 | Bombing/Explosion | Military | Iraq | 1 |
| 200801010013 | 2008 | 1 | 1 | Kosovo | Eastern Europe | 0 | 1 | 0 | Bombing/Explosion | Business | Serbia-Montenegro | 0 |
| 200801010014 | 2008 | 1 | 1 | Pakistan | South Asia | 1 | 1 | 0 | Hostage Taking (Kidnapping) | Military | Pakistan | 0 |

# Data Sample #2: HFI 2019

| country | region | hf_score | pf_score | ef_score |
|---|---|---|---|---|
| Albania | Eastern Europe | 7.679999828338623 | 8.069999694824219 | 7.28000020980835 |
| Algeria | Middle East & North Africa | 5.289999961853027 | 5.349999904632568 | 5.239999771118164 |
| Angola | Sub-Saharan Africa | 5.019999980926514 | 5.440000057220459 | 4.590000152587891 |
| Argentina | Latin America & the Caribbean | 7.099999904632568 | 8.189999580383301 | 6.010000228881836 |
| Armenia | Caucasus & Central Asia | 7.420000076293945 | 7.210000038146973 | 7.630000114440918 |
| Australia | Oceania | 8.729999542236328 | 9.289999961853027 | 8.180000305175781 |
| Austria | Western Europe | 8.420000076293945 | 9.020000457763672 | 7.820000171661377 |
| Azerbaijan | Caucasus & Central Asia | 6.440000057220459 | 6.730000019073486 | 6.150000095367432 |
| Bahamas | Latin America & the Caribbean | 7.96999979019165 | 8.319999694824219 | 7.610000133514404 |
| Bahrain | Middle East & North Africa | 7.150000095367432 | 6.880000114440918 | 7.409999847412109 |
| Bangladesh | South Asia | 5.800000190734863 | 5.599999904632568 | 6.010000228881836 |
| Barbados | Latin America & the Caribbean | 7.210000038146973 | 8.220000267028809 | 6.199999809265137 |
| Belarus | Eastern Europe | -1 | -1 | -1 |
| Belgium | Western Europe | 8.279999732971191 | 9.060000419616699 | 7.5 |
| Belize | Latin America & the Caribbean | 7.369999885559082 | 7.400000095367432 | 7.329999923706055 |
| Benin | Sub-Saharan Africa | 6.849999904632568 | 7.460000038146973 | 6.239999771118164 |
| Bhutan | South Asia | -1 | -1 | -1 |
| Bolivia | Latin America & the Caribbean | 6.75 | 7.460000038146973 | 6.039999961853027 |
| Bosnia and Herzegovina | Eastern Europe | 7.269999980926514 | 8 | 6.539999961853027 |
| Botswana | Sub-Saharan Africa | 6.809999942779541 | 6.75 | 6.869999885559082 |

# Design Diagram



# Code Challenge

# Code Challenge: Merging the data sets

When we were merging the data sets, we realized that the names of the countries did not match. To overcome this, we simply hard coded the changed names to match. We chose to use the names from the GTD dataset.

- First we found the country names that did not match, looked through them and manually paired up the names we believed belonged to the same country.

```
drop table if exists distinct_country;
create external table distinct_country (country string);
truncate table distinct_country;
insert into distinct_country select gtd_country from(select distinct gtd_country from gtd_distinct_country
    union ALL select distinct hfi_country from hfi_distinct_country) as tbl group by gtd_country having count(*) =1;
```

# Code Challenge: Merging the data sets

- Then we manually updated the HFI table to match the expressions from the GTD table.

```
drop table if exists hfi_master1;
create external table hfi_master1 (year int,country string, region string, hf_score float, pf_rol float,pf_ss_disappearances_disap float,
    pf_ss_disappearances_violent float,pf_ss_disappearances_organized float,pf_ss_disappearances_fatalities float,pf_ss_disappearances_injuries float,
    pf_ss_disappearances float,pf_ss_women_fgm float,pf_ss_women_inheritance_widows float,pf_ss_women_inheritance_daughters float,pf_ss_women_inheritance float,
    pf_ss_women float,pf_ss float,pf_movement_domestic float,pf_movement_foreign float,pf_movement float,pf_religion_estop_establish float,
    pf_religion_estop_operate float,pf_religion_estop float,pf_religion_harassment float,pf_religion_restrictions float,pf_religion float,
    pf_association float,pf_expression_killed float,pf_expression_jailed float,pf_expression_influence float,pf_expression_control float,
    pf_expression_cable float,pf_expression_newspapers float, pf_expression_internet float,pf_expression float,pf_identity float,pf_score float,
    ef_trade_tariffs_revenue float, ef_trade_tariffs float,ef_trade_black float,ef_trade_movement_foreign float,ef_trade_movement_capital float,
    ef_trade_movement_visit float,ef_trade_movement float,ef_trade float, ef_regulation_labor float, ef_regulation_business float,ef_regulation float,
    ef_score float);
truncate table hfi_master1;
insert into hfi_master1 (year,country, region, hf_score, pf_rol,pf_ss_disappearances_disap,pf_ss_disappearances_violent,pf_ss_disappearances_organized,
    pf_ss_disappearances_fatalities,pf_ss_disappearances_injuries,pf_ss_disappearances,pf_ss_women_fgm,pf_ss_women_inheritance_widows,pf_ss_women_inheritance_daughters,
    pf_ss_women_inheritance,pf_ss_women,pf_ss,pf_movement_domestic,pf_movement_foreign,pf_movement,pf_religion_estop_establish,pf_religion_estop_operate,
    pf_religion_estop,pf_religion_harassment,pf_religion_restrictions,pf_religion,pf_association,pf_expression_killed,pf_expression_jailed,pf_expression_influence,
    pf_expression_control,pf_expression_cable,pf_expression_newspapers, pf_expression_internet,pf_expression,pf_identity,pf_score, ef_trade_tariffs_revenue,
    ef_trade_tariffs,ef_trade_black,ef_trade_movement_foreign,ef_trade_movement_capital,ef_trade_movement_visit,ef_trade_movement,ef_trade, ef_regulation_labor,
    ef_regulation_business,ef_regulation,ef_score)
select year, (case
    when country='Yemen, Rep.' Then 'Yemen'
    when country='Slovak Rep.' Then 'Slovak Republic'
    when country='Pap. New Guinea' Then 'Papua New Guinea'
    when country='Kyrgyz Republic' Then 'Kyrgyzstan'
    when country='Timor-Leste' Then 'East Timor'
    when country='Korea, South' Then 'South Korea'
    when country='Bosnia and Herzegovina' Then 'Bosnia-Herzegovina'
    when country='Dominican Rep.' Then 'Dominican Republic'
    when country='Central Afr. Rep.' Then 'Central African Republic'
    when country='Czech Rep.' Then 'Czech Republic'
    when country='Congo, Dem. R.' Then 'Democratic Republic of the Congo'
    when country='Congo, Rep. Of' Then 'Republic of the Congo' else country end) as country, (case
    when country= 'Greece' Then 'Western Europe'
    when country='Cyprus' Then 'Western Europe' else region end) as region, hf_score, pf_rol,pf_ss_disappearances_disap,pf_ss_disappearances_violent,
    pf_ss_disappearances_organized,pf_ss_disappearances_fatalities,pf_ss_disappearances_injuries,pf_ss_disappearances,pf_ss_women_fgm,
    pf_ss_women_inheritance_widows,pf_ss_women_inheritance_daughters,pf_ss_women_inheritance,pf_ss_women,pf_ss,pf_movement_domestic,
    pf_movement_foreign,pf_movement,pf_religion_estop_establish,pf_religion_estop_operate,pf_religion_estop,pf_religion_harassment,pf_religion_restrictions,
    pf_religion,pf_association,pf_expression_killed,pf_expression_jailed,pf_expression_influence,pf_expression_control,pf_expression_cable,pf_expression_newspapers,
    pf_expression_internet,pf_expression,pf_identity,pf_score, ef_trade_tariffs_revenue, ef_trade_tariffs,ef_trade_black,ef_trade_movement_foreign,ef_trade_movement_capital,
    ef_trade_movement_visit,ef_trade_movement,ef_trade, ef_regulation_labor, ef_regulation_business,ef_regulation,ef_score from hfi_raw;
```

# Code Challenge: Merging the data sets

- Lastly, we made an intersection table, which we then used that make ‘master’ tables hfi_master and gtd_master by checking if the year, country combination exists in this intersection table.

```
drop table if exists cy_intersection;
create table cy_intersection (year int, country string, region string, year2 int, country2 string, region2 string);
truncate table cy_intersection;
insert into cy_intersection select * from cy_intersection_gtd inner join cy_intersection_hfi on cy_intersection_gtd.year=cy_intersection_hfi.year
    and cy_intersection_gtd.country=cy_intersection_hfi.country and cy_intersection_gtd.region=cy_intersection_hfi.region;
```

```
drop table if exists hfi_master;
create external table hfi_master(year int,country string, region string, hf_score float, pf_rol float,pf_ss_disappearances_disap float,
    pf_ss_disappearances_violent float,pf_ss_disappearances_organized float,pf_ss_disappearances_fatalities float,pf_ss_disappearances_injuries float,
    pf_ss_disappearances float,pf_ss_women_fgm float,pf_ss_women_inheritance_widows float,pf_ss_women_inheritance_daughters float,pf_ss_women_inheritance float,
    pf_ss_women float,pf_ss float,pf_movement_domestic float,pf_movement_foreign float,pf_movement float,pf_religion_estop_establish float,
    pf_religion_estop_operate float,pf_religion_estop float,pf_religion_harassment float,pf_religion_restrictions float,pf_religion float,
    pf_association float,pf_expression_killed float,pf_expression_jailed float,pf_expression_influence float,pf_expression_control float,
    pf_expression_cable float,pf_expression_newspapers float, pf_expression_internet float,pf_expression float,pf_identity float,pf_score float,
    ef_trade_tariffs_revenue float, ef_trade_tariffs float,ef_trade_black float,ef_trade_movement_foreign float,ef_trade_movement_capital float,
    ef_trade_movement_visit float,ef_trade_movement float,ef_trade float, ef_regulation_labor float, ef_regulation_business float,ef_regulation float,ef_score float);
truncate table hfi_master;
insert into hfi_master select * from hfi_master1 where EXISTS (select year, country FROM cy_intersection WHERE hfi_master1.year=cy_intersection.year
    and hfi_master1.country=cy_intersection.country);
```

# Code Challenge: Merging the data sets

Furthermore, the region names also did not match. GTD had more regions listed than HFI, which is why we decided to merge some of GTD ones to match the ones in HFI.

```
drop table if exists gtd_master1;
create external table gtd_master1 (event_id string, year int, month int, day int, country string, region string, doubt_terrorist int,
    success int, suicide int, attack_type string, target_type string, nationality string, num_kill int);
truncate table gtd_master1;
insert into gtd_master1 (event_id, year, month, day, country, region, doubt_terrorist, success, suicide, attack_type, target_type, nationality, num_kill)
select event_id, year, month, day, country, (case when region='Southeast Asia' Then 'South Asia'
        when region ='Australasia & Oceania' Then 'Oceania'
        when region = 'Central America & Caribbean' Then 'Latin America & the Caribbean'
        when region='South America' Then 'Latin America & the Caribbean'
        when region='Central Asia' Then 'Caucasus & Central Asia'
        when country ='Mexico' Then 'Latin America & the Caribbean' else region end) as region, doubt_terrorist, success,
    suicide, attack_type, target_type, nationality, num_kill from gtd_raw;
```

# Code Challenge: Getting aggregate results in GTD

We did not have any demographic data to control for number of events in relation to the population or the size of the country, so we decided to take **percentages of some other factors such as successful events, events including suicide, based on the total number of events.**

```
drop table if exists agg_gtd_master;
create external table agg_gtd_master (year int, country string, region string, num_events int, sum_mismatch int,sum_success int,
    sum_doubt int, sum_suicide int, sum_num_kill int, perc_suicide float, perc_doubt float, perc_success float, perc_mismatch float, killed_per_event float);
truncate table agg_gtd_master;
insert into agg_gtd_master select year, country, region, cast(count(event_id) as int), cast(sum(mismatch) as int), cast(sum(success) as int),
    cast(sum(doubt_terrorist) as int), cast(sum(suicide) as int), cast(sum(num_kill) as int), cast(((cast(sum(suicide)as int)/(cast(count(event_id) as int)))) as float),
    cast(((cast(sum(doubt_terrorist)as int)/(cast(count(event_id) as int)))) as float), cast(((cast(sum(success)as int)/(cast(count(event_id) as int)))) as float),
    cast(((cast(sum(mismatch)as int)/(cast(count(event_id) as int)))) as float), cast(((cast(sum(num_kill)as int)/(cast(count(event_id) as int)))) as float)
    from gtd_agg group by year, country, region;
```

# Insights

# Insights: General Insights

We ran correlations on different levels

- Year, country pairs available in both data sets - **834 unique inputs**.
- Countries across all the available years - **139 unique inputs**.
- Regions across all the countries and years - **10 unique inputs**.

Across all the correlations we ran, we found that the highest scores were obtained when exploring on the regional level. However, there are some outliers, which we will be present.

# Insights: General Insights

We found very weak correlation of the general indices number of events, number of suicides, number of successful events with the general indices of Human Freedom score, Personal Freedom score and Economic Freedom score. This correlations was done on year, country pairs.

| | num_events | sum_num_killed |
|---|---|---|
| hf_score | -0.259 | -0.282 |
| pf_score | -0.298 | -0.318 |
| ef_score | -0.149 | -0.175 |

It seems that although week, the correlations are the strongest with the Personal Freedom score.

# Insights #1: There is minimal relationship between freedom of expression and # of terrorism

On year, country pairs there is **very minimal relationship between number of events and the scores for freedom of expression** - Expression & Information (pf_expression), Laws and Regulations that Influence Media Content (pf_expression_influence), Press Killed (pf_expression_killed), Press Jailed (pf_expression_jailed), Political Pressures and Controls on Media Content (pf_expression_control), State Control over Internet Access (pf_expression_internet). The **strongest (negative) correlation is with the score for Press Killed**.

| | pf_expression | pf_expression_influence | pf_expression_killed | pf_expression_control | pf_expression_internet |
|---|---|---|---|---|---|
| num_events | -0.151 | -0.088 | -0.297 | -0.183 | -0.063 |

# Insights #1: There is minimal relationship between freedom of expression and # of terrorism

| country | year | num_events | pf_expression | pf_expression_influence |
|---|---|---|---|---|
| Iraq | 2016 | 3383 | 5.300000190734863 | 2.299999952316284 |
| Iraq | 2017 | 2503 | 5.400000095367432 | 2.299999952316284 |
| Pakistan | 2013 | 2214 | 7.099999904632568 | 3.700000047683716 |
| Pakistan | 2014 | 2151 | 7.099999904632568 | 3.700000047683716 |
| Pakistan | 2012 | 1654 | 7.099999904632568 | 3.700000047683716 |
| Pakistan | 2015 | 1243 | 7.300000190734863 | 3.700000047683716 |
| India | 2016 | 1026 | 8 | 6.300000190734863 |
| Pakistan | 2011 | 1012 | 7.099999904632568 | 4 |
| India | 2017 | 967 | 8 | 6.300000190734863 |
| Ukraine | 2014 | 895 | 6.599999904632568 | 5.300000190734863 |

# Insights #1: There is minimal relationship between freedom of expression and # of terrorism

On the country level the correlation increases, however, it is still quite weak.

| | pf_expression | pf_expression_influence | pf_expression_killed | pf_expression_jailed | pf_expression_control | pf_expression_internet |
|---|---|---|---|---|---|---|
| num_events | -0.198 | -0.130 | -0.542 | -0.004 | -0.204 | -0.099 |

| country | num_events | pf_expression_c | pf_expression_influence_c |
|---|---|---|---|
| Iraq | 21590 | 5.350000381469727 | 2.299999952316284 |
| Pakistan | 11805 | 7.550000190734863 | 4.360000133514404 |
| India | 7559 | 8.079999923706055 | 6.630000114440918 |
| Philippines | 4403 | 8.140000343322754 | 5.789999961853027 |
| Nigeria | 3691 | 8.239999771118164 | 4.820000171661377 |
| Yemen | 3191 | 5.800000190734863 | 1 |
| Thailand | 2894 | 6.840000152587891 | 2.599999904632568 |
| Libya | 2236 | 5.860000133514404 | 3.099999904632568 |
| Syria | 2058 | 3.950000047683716 | 0.300000011920929 |
| Egypt | 1995 | 6.630000114440918 | 2.329999923706055 |

# Insights #1: There is minimal relationship between freedom of expression and # of terrorism

On the region level the **correlations increase**. The biggest increase is seen in the State control of the Internet. Given the steep jump, **we are assuming that here the correlation is not a result of causation but there are other factors influencing the scores that we are just not aware of.**

| | pf_expression | pf_expression_influence | pf_expression_killed | pf_expression_jailed | pf_expression_control | pf_expression_internet |
|---|---|---|---|---|---|---|
| num_events | -0.578 | -0.561 | -0.713 | -0.565 | -0.468 | -0.674 |

| region | num_events | pf_expression_r | pf_expression_influence_r |
|---|---|---|---|
| Middle East & North Africa | 35128 | 6.373831748962402 | 2.652336359024048 |
| South Asia | 29234 | 7.371716976165771 | 4.011111259460449 |
| Sub-Saharan Africa | 7944 | 7.386792659759521 | 3.832075595855713 |
| Eastern Europe | 2999 | 8.459375381469727 | 5.667187690734863 |
| Western Europe | 2144 | 9.376859664916992 | 8.031405448913574 |
| Latin America & the Caribbean | 1805 | 8.089691162109375 | 5.061855792999268 |
| North America | 341 | 9.447368621826172 | 8.315789222717285 |
| Caucasus & Central Asia | 144 | 6.491666793823242 | 2.629166603088379 |
| East Asia | 143 | 7.516666412353516 | 4.545833587646484 |
| Oceania | 56 | 9.445454597473145 | 8.300000190734863 |

# Insights #2: Terrorism shows high relationship with conflict itself

One of the **highest correlations we observed for year, country pairs** was for the correlations with the score for **Disapperances, Conflict & Terrorism (pf_ss_disappearances)**. Correlation is even higher on the country and region level. These are all **negatively correlated** - higher the number of the events, lower the score.

| | pf_ss | pf_ss_women | pf_ss_disappearances |
|---|---|---|---|
| num_events | -0.338 | -0.186 | -0.530 |

pf_ss is the Security and Safety score and pf_ss_women is the SS score for women. Both pf_ss_womena and pf_ss_dissapearences are aggregated with other scores in the pf_ss.

# Insights #2: Terrorism shows high relationship with conflict itself

Table showing top 5 year, country pairs ordered by the highest number of events.

| year | country | region | num_events | pf_ss_disappearances |
|---|---|---|---|---|
| 2016 | Iraq | Middle East & North Africa | 3383 | 0 |
| 2017 | Iraq | Middle East & North Africa | 2503 | 0 |
| 2013 | Pakistan | South Asia | 2214 | 3.5 |
| 2014 | Pakistan | South Asia | 2151 | 3.799999952316284 |
| 2012 | Pakistan | South Asia | 1654 | 3.5 |

Table showing bottom 5 year, country pairs ordered by the highest number of events.

| year | country | region | num_events | pf_ss_disappearances |
|---|---|---|---|---|
| 2016 | Haiti | Latin America & the Caribbean | 1 | 9 |
| 2014 | Iceland | Western Europe | 1 | 10 |
| 2012 | Mozambique | Sub-Saharan Africa | 1 | 9 |
| 2015 | Trinidad and Tobago | Latin America & the Caribbean | 1 | 9.5 |
| 2015 | Georgia | Caucasus & Central Asia | 1 | 9 |

# Insights #2: Terrorism shows high relationship with conflict itself

Correlation on level of countries (across all the available years). This is an increase from that on year, country pairs.

| | pf_ss | pf_ss_women | pf_ss_disapperances |
|---|---|---|---|
| num_events | -0.359 | -0.195 | -0.623 |

| country | region | num_events | pf_ss_disappearances_avg_c |
|---|---|---|---|
| Iraq | Middle East & North Africa | 21590 | 0 |
| Pakistan | South Asia | 11805 | 4.190000057220459 |
| India | South Asia | 7559 | 7.090000152587891 |
| Philippines | South Asia | 4403 | 6.679999828338623 |
| Nigeria | Sub-Saharan Africa | 3691 | 5.840000152587891 |

# Insights #2: Terrorism shows high relationship with conflict itself

Correlation on level of regions (across all the available years and countries). This is an increase from that on the country level or year, country pairs.

| | pf_ss | pf_ss_women | pf_ss_disapperances |
|---|---|---|---|
| num_events | -0.557 | -0.738 | -0.825 |

| region | num_events | pf_ss_disappearances_avg |
|---|---|---|
| Middle East & North Africa | 35128 | 6.558677673339844 |
| South Asia | 29234 | 7.584465980529785 |
| Sub-Saharan Africa | 7944 | 7.617391109466553 |
| Eastern Europe | 2999 | 8.971621513366699 |
| Western Europe | 2144 | 9.735537528991699 |

# Insights #3: Terrorism is rather harder to define, and not necessarily determined in relationship to the losses of lives

Currently, there is no universal, world-wide agreed upon definition for terrorism.

Global Terrorism Database (GTD), which defines terrorism as “acts of violence by non-state actors, perpetrated against civilian populations, intended to cause fear, in order to achieve a political objective”

Based on our analytics of doubts_terrorist and success, this concept of “fear” is not shown to be directly influenced by losses in human lives, ie. fatalities occurred or if the outcome has resulted into death of the actors or the victims.

# Insights #3: Terrorism is rather harder to define, and not necessarily determined in relationship to the losses of lives

- High correlation with doubt in the event was a terrorist attack and whether the event was successful, number of fatalities and whether it had led to suicide
- High correlation with success in a terrorist attack and whether the event was doubted in it being a terrorist attack, number of fatalities and whether it had led to suicide
- Both doubts in the terrorist attack and whether an event was successful or not were not heavily but still moderately correlated to the attacker being a non-national

| | success | num_killed | suicide | Non_national terrorist |
|---|---|---|---|---|
| doubt_terrorist | 0.857 | 0.750 | 0.716 | 0.540 |
| success | | 0.886 | 0.847 | 0.443 |

# Insights #4: There is minimal correlation between events of terrorism and freedom of religion

| | pf_religion | pf_religion_estop | pf_religion_harassment | pf_religion_restrictions |
|---|---|---|---|---|
| num_events | -0.172 | -0.095 | -0.050 | -0.086 |
| sum_num_kill | -0.158 | -0.048 | -0.091 | -0.042 |

Pf_religion in the target type of Private Citizens & Property, Business Government (General), Government (Diplomatic), Military, Religious Figures /Institutions

| | Private Citizens & Property | Business | Government (General) | Government (Diplomatic) | Military | Religious Figures /Institutions |
|---|---|---|---|---|---|---|
| num_events | -0.125 | -0.082 | 0.001 | 0.115 | -0.168 | -0.085 |

# Obstacles

# Obstacles #1: Missing Data

HFI Dataset has missing data, that was simply denoted by '-'. To get around that we replaced all the '-' with -1 in the cleaning stage and made sure to control for that in our analytic.

```
for line in sys.stdin:
        line = line.strip()
        words=line.split(';')
        x = len(words)
        count=0
        list=''
        for a in words:
                count+=1
                if a!='-':
                        list+=a
                else:
                        list+='-1'

                if count!=x:
                        list+=';'

        print '%s'%(list)
```

```
drop table if exists hfi_country_diff;
create external table hfi_country_diff (country string, min_hf float,
    max_hf float, diff float);
truncate table hfi_country_diff;
insert into hfi_country_diff select country, cast(min(hf_score) as float),
    cast(max(hf_score) as float), cast ((max(hf_score)- min(hf_score)) as float)
    from hfi_main where hf_score!=-1 group by country;
select * from hfi_country_diff order by diff desc;
```

# Obstacles #2: Non-compatible Data Sets

As described in the Data Challenges our data sets **were not compatible.** The **names of the countries and regions were different**. Some countries were event **classified in different regions** between the two datasets (Greece, Cyprus and Mexico). Furthermore, GTD had a **list of events, where as HFI has aggregate results by year and country.**

Hence, we had to aggregate results in GTD by year, country combinations and only keep those *year, country* combinations that were present in both. This ended up being **834 combinations across 10 regions and years 2008 - 2017, with 139 distinct countries** (average 6 years of data per country).

Given how varied these data points were, some countries had more years of data than others, it definitely impacted our results. Hence, we found it hard to draw any conclusive outcomes.

# Obstacles #2: Non-compatible Data Sets

Example how varied the number of years per country is.

| | | | |
|---|---|---|---|
| 2008 | Greece | Western Europe | 53 |
| 2009 | Greece | Western Europe | 115 |
| 2010 | Greece | Western Europe | 49 |
| 2011 | Greece | Western Europe | 11 |
| 2012 | Greece | Western Europe | 23 |
| 2013 | Greece | Western Europe | 54 |
| 2014 | Greece | Western Europe | 26 |
| 2015 | Greece | Western Europe | 31 |
| 2016 | Greece | Western Europe | 30 |
| 2017 | Greece | Western Europe | 44 |
| 2012 | Guatemala | Latin America & the Caribbean | 1 |
| 2013 | Guatemala | Latin America & the Caribbean | 6 |
| 2013 | Guinea | Sub-Saharan Africa | 2 |
| 2015 | Guinea | Sub-Saharan Africa | 2 |
| 2016 | Guinea | Sub-Saharan Africa | 1 |
| 2009 | Guinea-Bissau | Sub-Saharan Africa | 1 |
| 2012 | Guinea-Bissau | Sub-Saharan Africa | 1 |
| 2008 | Guyana | Latin America & the Caribbean | 3 |
| 2016 | Guyana | Latin America & the Caribbean | 1 |
| 2016 | Haiti | Latin America & the Caribbean | 1 |

# Obstacles #2: Non-compatible Data Sets

## Name changes that had to be made to HFI to match GTD notation:

Yemen, Rep. (HFI) → Yemen (GTD)
Slovak Rep. (HFI) → Slovak Republic (GTD)
Pap. New Guinea (HFI) → Papua New Guinea (GTD)
Kyrgyz Republic (HFI)→ Kyrgyzstan (GTD)
Timor-Leste (HFI) → East Timor (GTD)
Korea, South (HFI) → South Korea (GTD)
Bosnia and Herzegovina (HFI)→ Bosnia-Herzegovina (GTD)
Dominican Rep. (HFI) → Dominican Republic (GTD)
Central Afr. Rep.(HFI) → Central African Republic (GTD)
Czech Rep.(HFI) → Czech Republic (GTD)
Congo, Rep. Of (HFI) → Republic of the Congo (GTD)
Congo, Dem. R. (HFI) → Democratic Republic of the Congo (GTD)

## Regions that had to be combined in the GTD to match HFI:

South Asia (HFI) → South Asia (GTD) + Southeast Asia (GTD)
Oceania (HFI) → Australasia & Oceania (GTD)
Latin America & the Caribbean (HFI) → Central America & Caribbean (GTD) + South America (GTD)
Caucasus & Central Asia (HFI) → Central Asia (GTD)

## Countries with mismatched regions:

Cyprus: Eastern Europe (HFI); Western Europe (GTD)
→ Western Europe
Greece: Eastern Europe (HFI); Western Europe (GTD)
→ Western Europe
Mexico: Latin America & the Caribbean (HFI); North America (GTD)
→ Latin America & the Caribbean

# Obstacles #3: Lacking Research Question and Poor Planning

We were not able to find a lot of research that would look into the relations of terrorism and freedom, besides the transnational terrorism, hence we decided to **focus on on getting wide array of different results instead of setting a specific research question**. In hindsight, this was not the best approach, as we now have a lot of general insights but very little of those that have more significant impact.

Furthermore, this was a project like this we have both undertaken. We **underestimated the importance of proper planning and data profiling**. Had we done that part better, we would likely have more insightful results.

# Obstacles #4: Conclusions

The **most conclusive results ended up being those that were aggregated by region** (10 in total) and not by year, country combinations. However, there are **other factors that could influence that**. Hence our results, while insightful, are hardly conclusive. Other factors, that we did not investigate but could have an effect on this are cultural, political, geographical, etc.

Furthermore, a lot of our conclusions are based on the correlations between number of terrorist events and freedom scores. Since we **do not control the number of events for the population** of the country/region or the number of the countries in the region, those are **likely skewed towards the regions with higher population.**

# Reflection

After the completing the analytics part of the project, we realized there were much room for improvement. Below are some of the things ***we would rethink***, were we starting this project all over again.

- We would **define the research question better**. That would allow us to focus this project more narrowly and go in depth instead of breadth.
- We would **set up the experiments better** and and find a way to compare the correlations as well by having them structured in a separate table.
- We would **explore the use of other tools**, besides impala, such a tableau or similar that would allow us to do some basic visualizations as well.
- We would **add demographic data** as an additional dataset. That would enable us to do better and more in depth correlations measures.
- We would do a **case study**. Exploring these correlations for a particular country or region.

# Summary

- Broad overview of the correlations between the instances of terrorism accross the world with different freedom indices.
- Due to the limited scope of the project, we cannot make any definite conclusions on the causations, but the explored correlations are very insightful.
- No significant correlation between freedom of press and number of terrorist attacks.
- No significant correlation between freedom of religion and number of terrorist attacks.

# References - papers/online resources

Alex P. Schmid (2005) Root Causes of Terrorism: Some Conceptual Notes, a Set of Indicators, and a Model, Democracy and Security, 1:2, 127-136, DOI: 10.1080/17419160500321139

Eskildsen LS, Bjørnskov C. Does Freedom of Expression Cause Less Terrorism? Political Studies. September 2020. doi:10.1177/0032321720950223

Greene A. DEFINING TERRORISM: ONE SIZE FITS ALL? International and Comparative Law Quarterly. 2017;66(2):411-440. doi:10.1017/S0020589317000070

Hannah Ritchie, Joe Hasell, Cameron Appel and Max Roser (2013) - "Terrorism". Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/terrorism' [Online Resource]

Institute for Economics & Peace. Global Terrorism Index 2020: Measuring the Impact of Terrorism, Sydney, November 2020. Available from: http://visionofhumanity.org/reports (accessed 02 12 2021). https://www.visionofhumanity.org/wp-content/uploads/2020/11/GTI-2020-web-1.pdf

# References - papers/online resources

Lobato RM, Moya M, Moyano M, Trujillo HM. From Oppression to Violence: The Role of Oppression, Radicalism, Identity, and Cultural Intelligence in Violent Disinhibition. Front Psychol. 2018;9:1505. Published 2018 Aug 20. doi:10.3389/fpsyg.2018.01505

SCHEININ, Martin, Limits to freedom of expression : lessons from counter-terrorism, in Tarlach MCGONAGLE and Yvonne DONDERS (eds), The United Nations and freedom of expression and information : critical perspectives, Cambridge : Cambridge University Press, 2015, pp. 428-442 Retrieved from Cadmus, European University Institute Research Repository, at: http://hdl.handle.net/1814/38808

Wesley S. McCann (2020) Who Said We Were Terrorists? Issues with Terrorism Data and Inclusion Criteria, Studies in Conflict & Terrorism, DOI: 10.1080/1057610X.2020.1816676

# References - data sources

Global Terrorism Database (GTD) 1970-2018
https://gtd.terrorismdata.com/files/gtd-1970-2018/

Global Terrorism Database (GTD) Code Book
start.umd.edu/gtd/downloads/Codebook.pdf

Human Freedom Index (HFI) 2019
https://www.cato.org/human-freedom-index-new