

# CS Semester Project: Global Terrorism Database

Group 1: Brago Aboagye-Nyame,  
Grant Redfield, Mary Youssef

# Agenda

- Data Overview
- Events From a Time Perspective
- Events From a Location Perspective
- Events from a Motive/Attack Type Perspective
- Conclusion

# What Is the Global Terrorist Database?

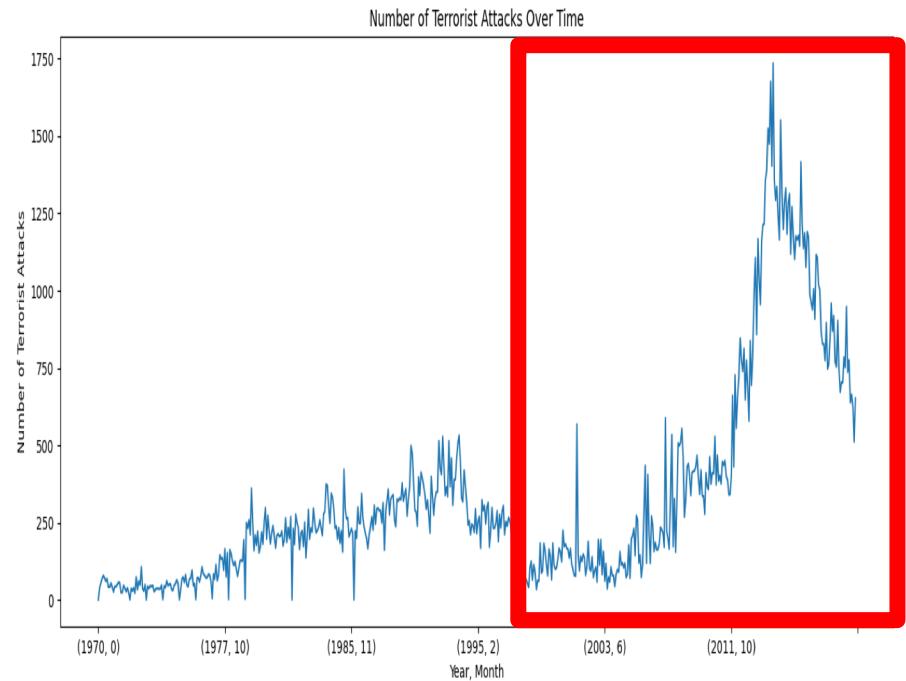
The Global Terrorism Database (GTD) is an open-source database including information on domestic and international terrorist attacks around the world from 1970 through 2019

What is considered a “terrorist event”?

- “The threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation.”

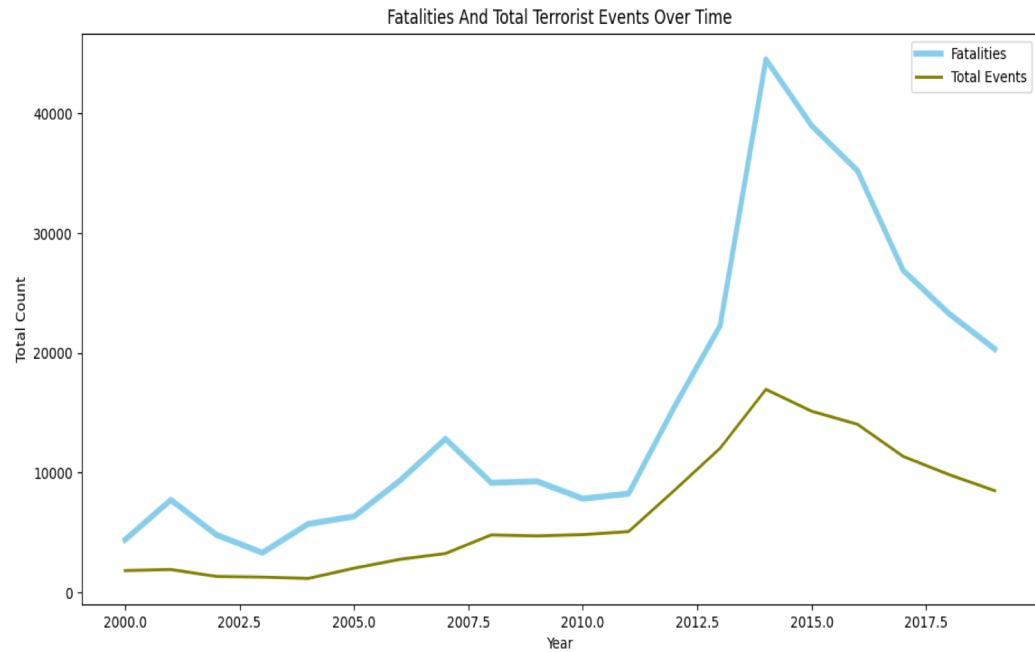
What time range did we focus on?

- The 21st century has allowed researchers to track events more closely due to the age of the internet. We focused on the last two decades (2000-2019) rather than the complete 1970 - 2019 to gather more recent information on terrorist activity.



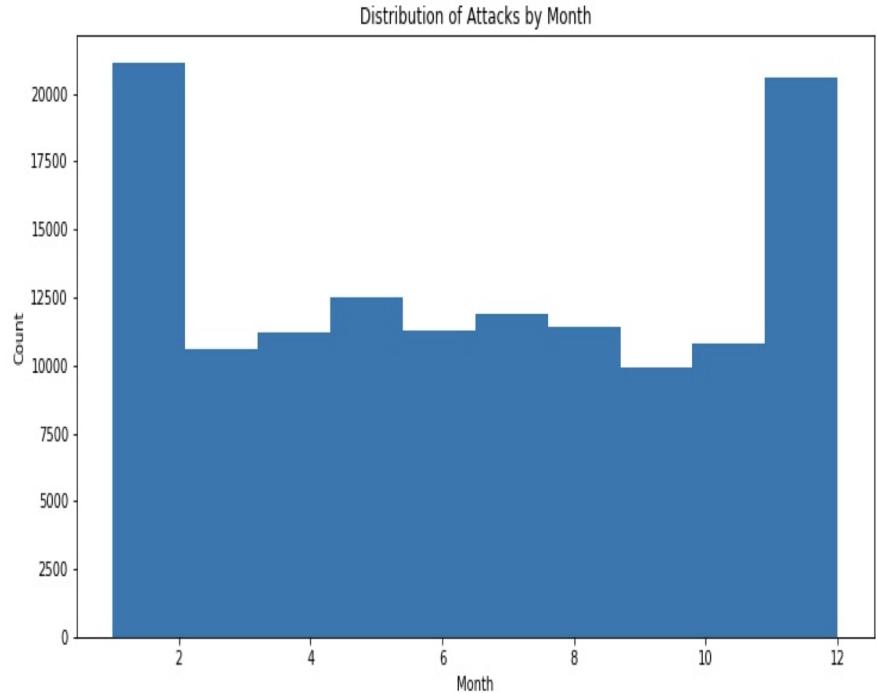
# When : Terrorist Events Over Time

Starting in 2000, there has been an increasing number of events in the world, peaking in 2014 where a record 44,524 fatalities were recorded. There has been a steady decrease since, but the numbers of events and fatalities remain high.



# When: Terrorist Events Over The Course Of the Year

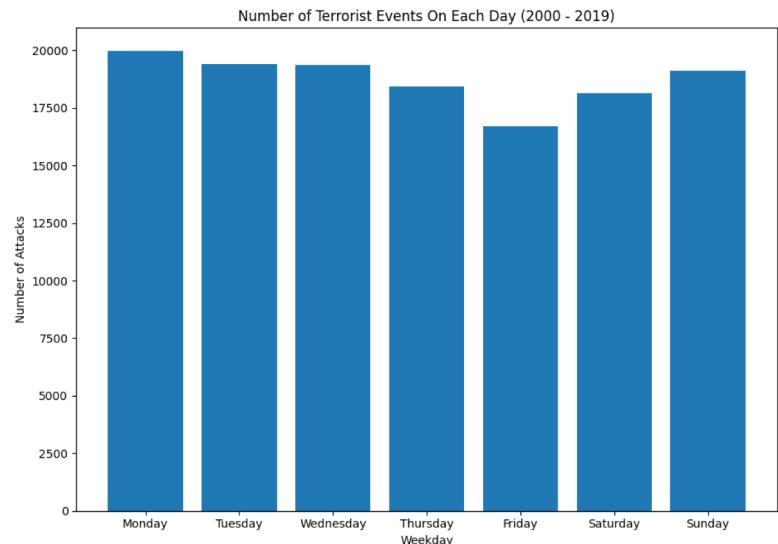
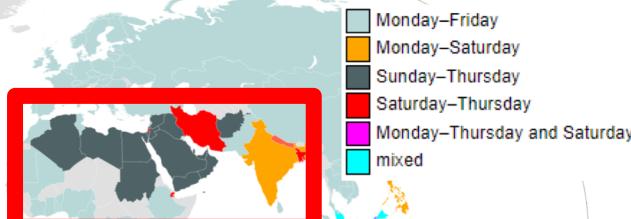
- Spikes appear at the beginning and end of the year
- This could be due to
  - Religious Holidays
  - Extended Time off for citizens
  - Election Cycles
- Further research needs to be conducted to quantify the reasoning behind these spikes



# When: Terrorist Events Over The Course Of the Week

- It is observed that terrorist like events are more likely to occur on Mondays as opposed to Fridays.

Typical Work Week By Country



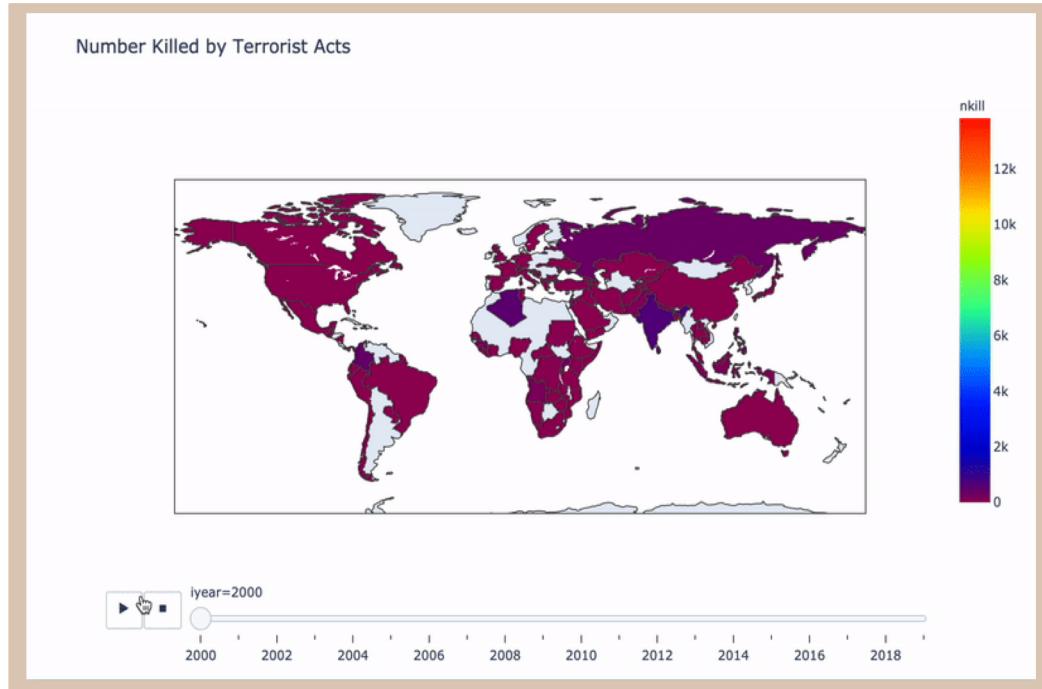
[Source](#)

# Where: Where Are These Events Happening?

Countries in the East tend to have the most number of people killed in their total attacks per year.

To the right is a table of the top 15 instances of people killed in a certain year.

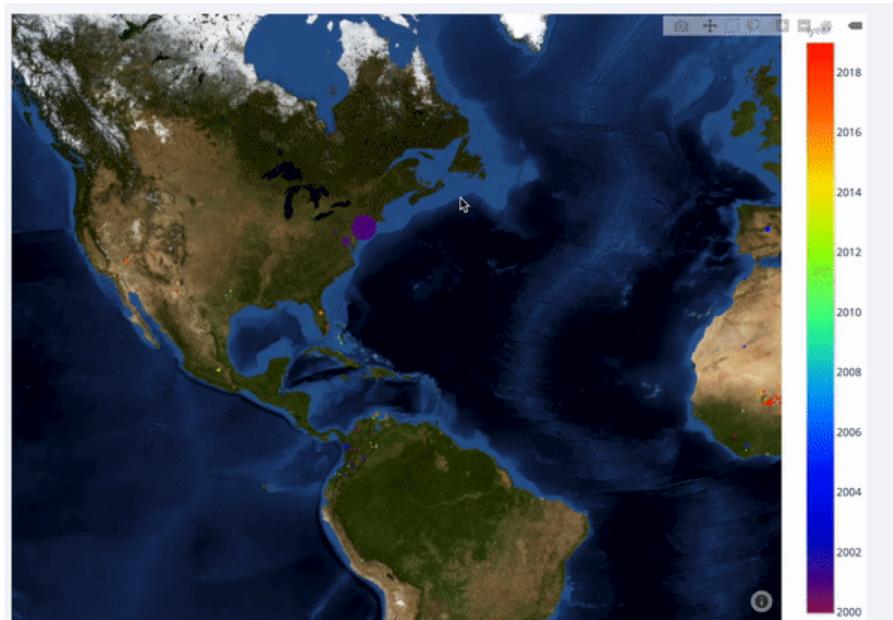
	iyear	country_txt	nkill
1160	2014	Iraq	13799.0
1363	2016	Iraq	12114.0
1530	2018	Afghanistan	9685.0
1259	2015	Iraq	8769.0
1639	2019	Afghanistan	8180.0
1188	2014	Nigeria	7703.0
1067	2013	Iraq	7002.0
569	2007	Iraq	6660.0
1465	2017	Iraq	6481.0
1221	2015	Afghanistan	6129.0
1319	2016	Afghanistan	6069.0
1427	2017	Afghanistan	6013.0
1287	2015	Nigeria	5512.0
1121	2014	Afghanistan	5360.0
498	2006	Iraq	4579.0



# Where: Certain Hotspots Have Appeared

This map shows the largest individual attacks across the globe from 2000-2019. Below are the top 10 attacks.

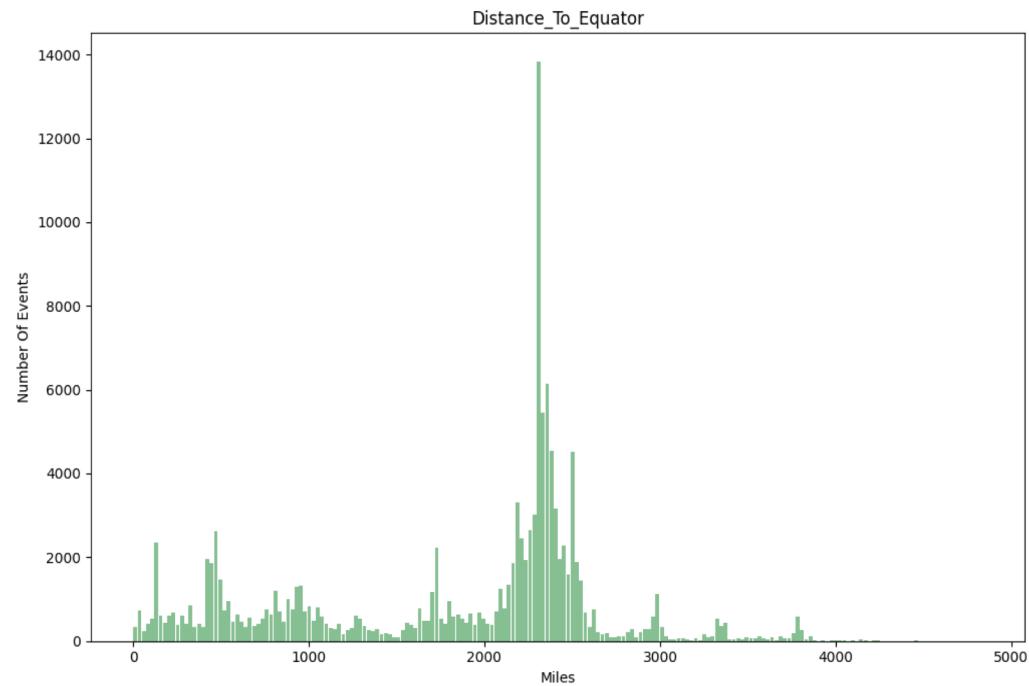
	iyear	imonth	iday	country_txt	nkill	attacktype1_txt
4696	2001	9	11	United States	1385.0	Hijacking
28463	2014	6	10	Iraq	670.0	Armed Assault
38453	2017	10	14	Somalia	588.0	Bombing/Explosion
7919	2004	3	21	Nepal	518.0	Armed Assault
22461	2016	7	3	Iraq	383.0	Bombing/Explosion
8524	2004	9	1	Russia	344.0	Hostage Taking (Barricade Incident)
9502	2018	12	16	Democratic Republic of the Congo	339.0	Hostage Taking (Kidnapping)
3663	2018	5	15	Afghanistan	330.0	Bombing/Explosion
39522	2017	11	24	Egypt	311.0	Bombing/Explosion
16600	2016	2	7	Iraq	300.0	Hostage Taking (Kidnapping)



# Where: How Close Are Events To The Equator?

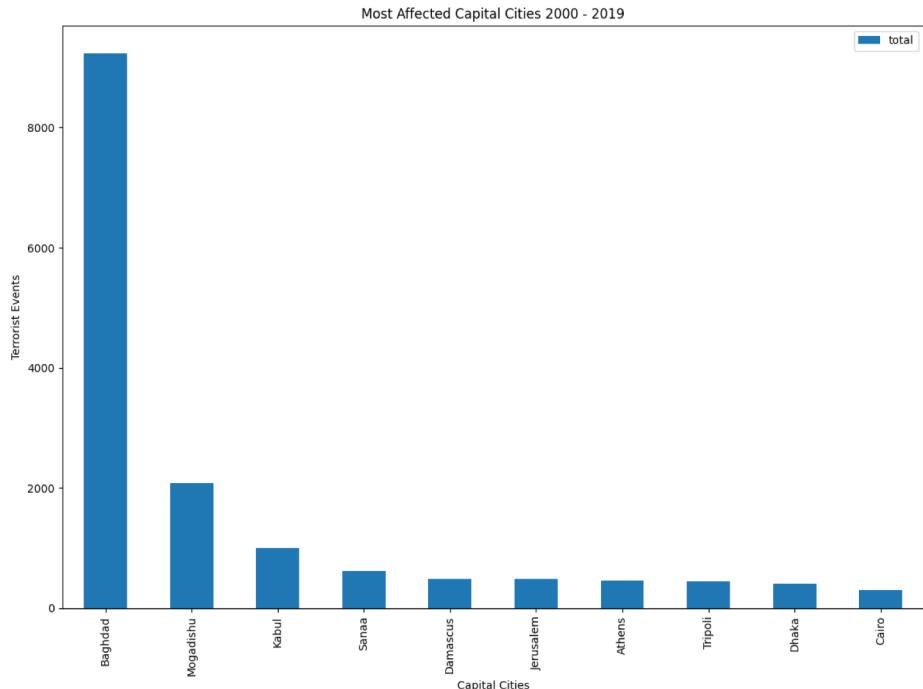
There is a left-skewed distribution of attacks when comparing distance of the event to the equator, with the majority between 2000 and 3000 which subsequently are countries in the middle east.

These spikes can also possibly be associated with population density as well as climate.



# Where: A Closer Look Into Capital Cities (cont.)

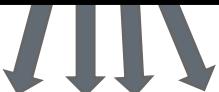
- It is observed that there are certain capital cities that are much more affected than others.
- Many capital cities around the middle east have been targeted over the past two decades



# Where: Using Vectorized Functions To Calculate Distances



Feed numpy arrays into function



```
def haversine_vectorize(lon1, lat1, lon2, lat2):  
    lon1, lat1, lon2, lat2 = map(np.radians, [lon1, lat1, lon2, lat2])  
    newlon = lon2 - lon1  
    newlat = lat2 - lat1  
    haver_formula = np.sin(newlat/2.0)**2 + np.cos(lat1) * np.cos(lat2) * np.sin(newlon/2.0)**2  
    dist = 2 * np.arcsin(np.sqrt(haver_formula))  
    miles = 3958 * dist # Multiplying Distance by Miles/Radius of the Earth  
    index_location = np.where(miles == miles.min())
```

**Time Difference of Calculation Methods**

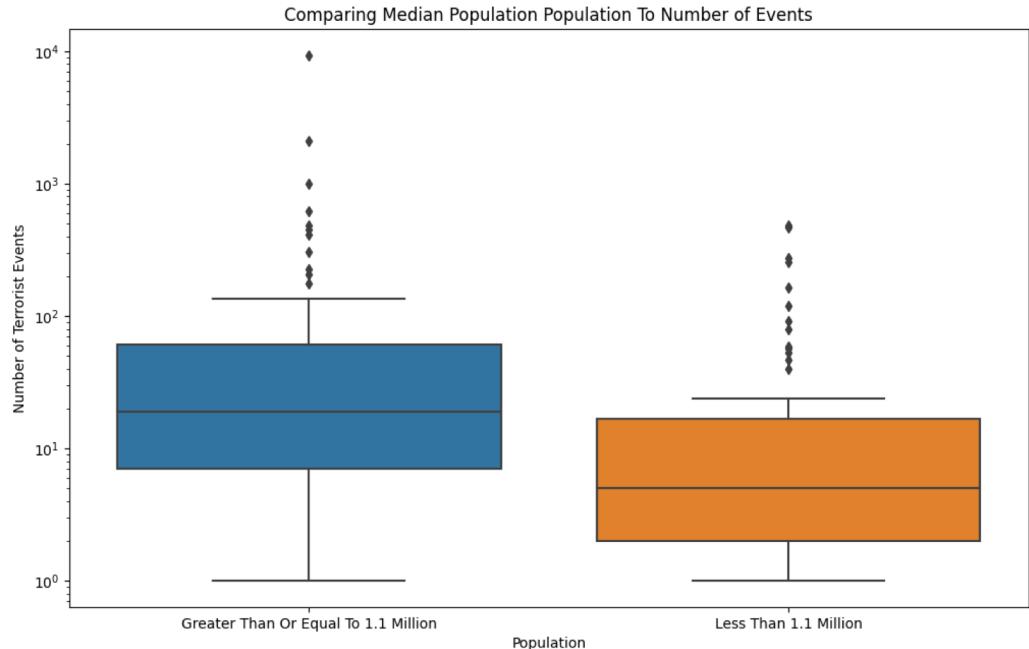
# of total calculations = ~27 million

Iterative Calculation Method = 2+ hours

Vectorized Calculations = 10 minutes

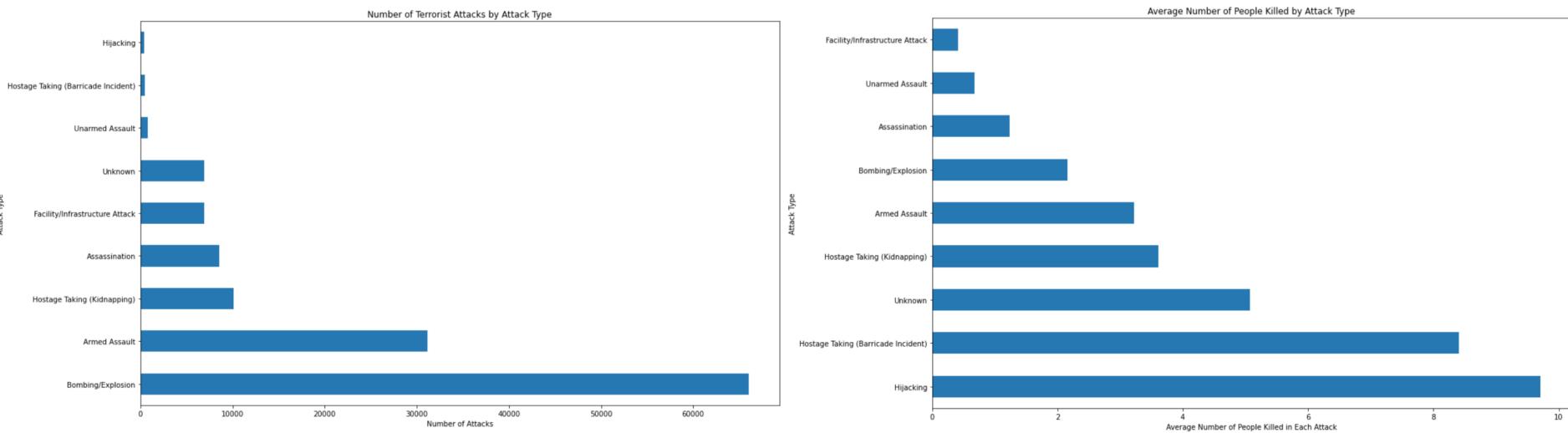
# Where: A Closer Look Into Capital Cities

- 15.2% of all terrorist events were within 20 miles of a capital city in the world.
- The average distance of a Capital City Event is within 5 miles of the center of the city.
- The more populated a city is, the more likely an event is to occur.

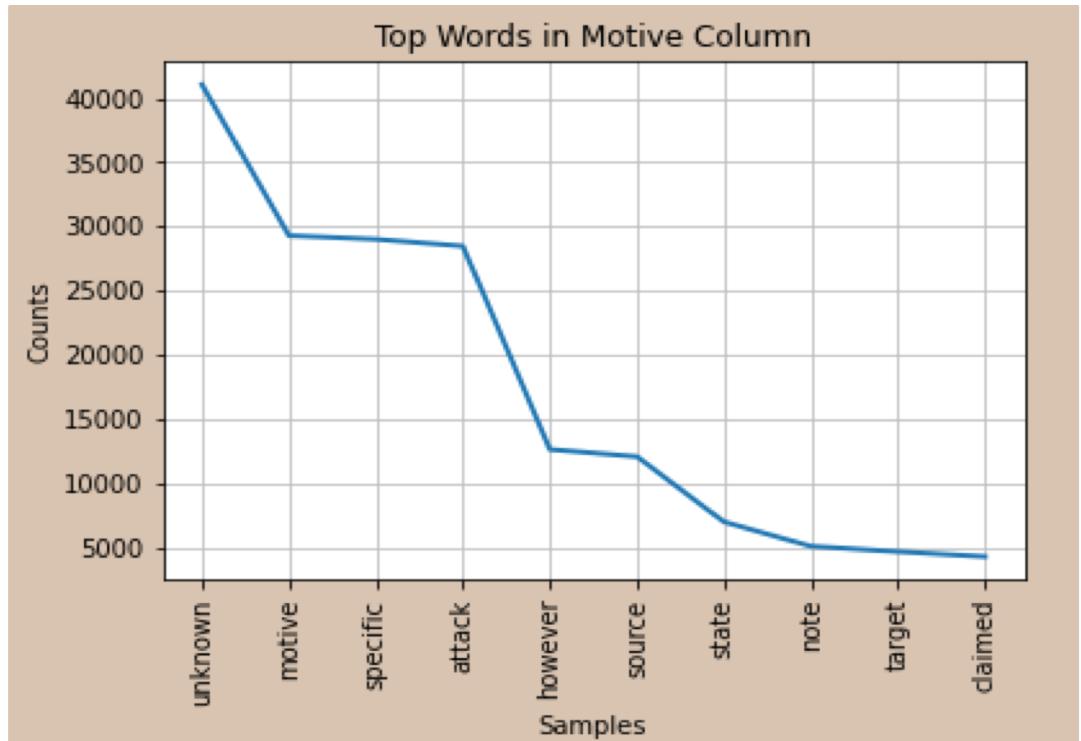


# How: Event Type and Fatalities

While Bombing/ Explosions were the most common attack type from 2000-2019, Hijacking attack types caused the most fatalities on average.



# How: Free Text Search in Unstructured Data



# Discussion and Conclusion

It is apparent that there are patterns commonalities when analyzing terrorist events. From day of the week, to motive behind the attacks can all play an important role to predict and prevent attacks in the future.

While this was extensive research there are still additional opportunities to dig deeper than our current state. The data is extremely detailed on attacks



[Link to image](#)

# Questions?