**Global Terrorism Analysis**

**Team:** AbracaData

**Introduction to the dataset:**

1. Event-level database of 180k+ terrorist attacks ranging from 1970 to 2017
2. 100+ variables that give an understanding of when and where the attacks occurred, motive, attack type, targets, perpetrators, etc.
3. Managed and maintained by START, the National Consortium for the Study of Terrorism and Responses to Terrorism, University of Maryland, College Park
4. Based almost entirely on unclassified information, mainly media reports

**Problems with our dataset:**

1. Month and Day columns have 0 as values which is an unacceptable representation of the exact date of the event.
2. Multiple Target and Weapon type/Subtype columns which are redundant.
3. “Unknown” values in each row will have to be dealt with.
4. The record of incidents for the year 1993 is missing in the dataset.
5. Group/Individual responsible and Motive for the attack is not known for many incidents.
6. This is a large dataset with 1,80,000+ records. (is this really a problem??)
7. For prop extent, the range is stupid. <=1mn ; 1mn to 1bn which is vast so it doesn’t really help in the analysis

**Data preprocessing and cleaning:**

1. Our data set had 180k entries and 136 columns
2. We decided to use data for 2001 to 2017
3. We created a subset of data 2001-2017 (the most recent)
4. At this stage or data set has 110,041 entries and 136 columns
5. Next, we investigated these variables thoroughly for missing values. The data set had “NA” and “Unknown” missing values and we dropped these missing values, the exercise was done using R studio.

**Questions from Dataset:**

The preprocessing gave us a good insight into data and helped us to raise specific questions worth answering from the data.

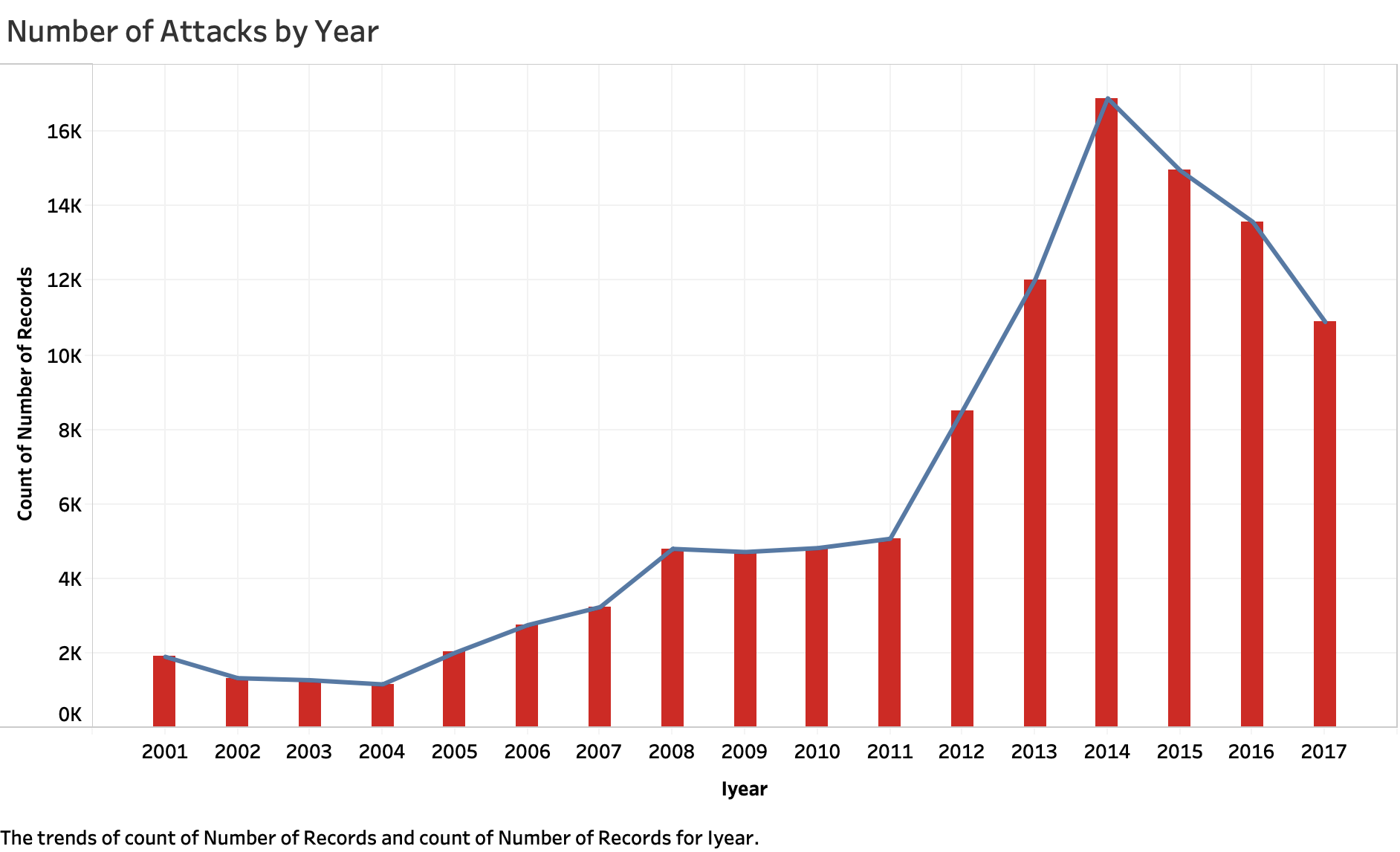
**Questions**

* How has terrorism activities changed over the years? Is it increasing or decreasing?
* Which regions/countries are the most affected by terrorism (most number of terror attacks?)
* Which terrorist groups have been most active?
* What is the most common attack type?
* What is the intensity of attacks by region for a particular year?
* What are the weapons used in these attacks?
* What are the most common target types?

**Analysis**

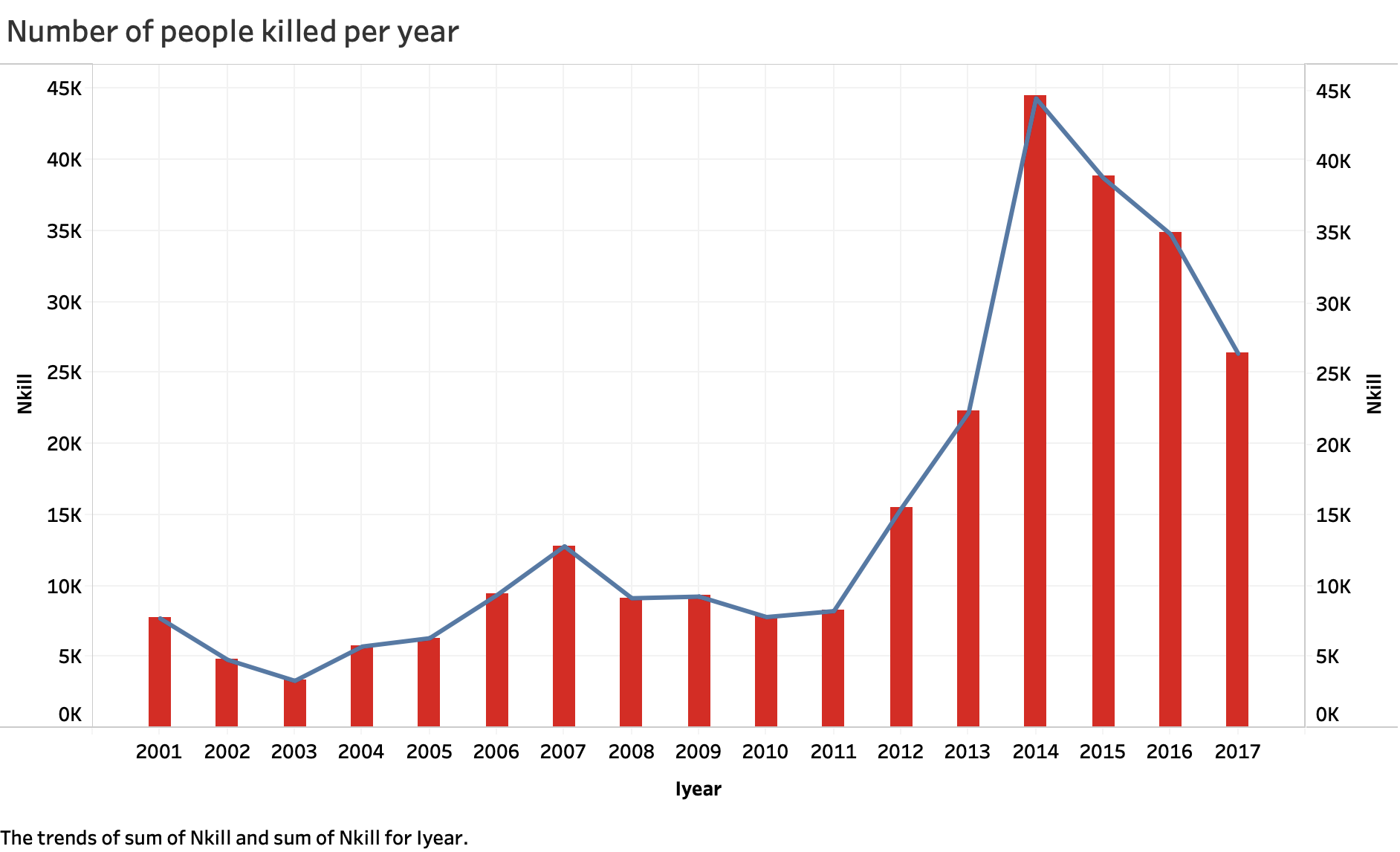
1. We used R and Tableau to perform our analysis.
2. Our first step was to filter out variables from our dataset which are relevant to our questions. For this, we took a deep dive into data and after a thorough investigation, we filtered the original 136 to 26 columns - this process was done using R-studio.
3. Next, we loaded our prepared data set into Tableau to perform visual analysis for any trends which would answer our questions.
4. Considering the nature of our data set and our questions, we could do most of our analysis in Tableau. The below graphs summarize our analysis using Tableau.

**How has terrorism activities changed over the years? Is it increasing or decreasing?**

****

**Fig 1.0**

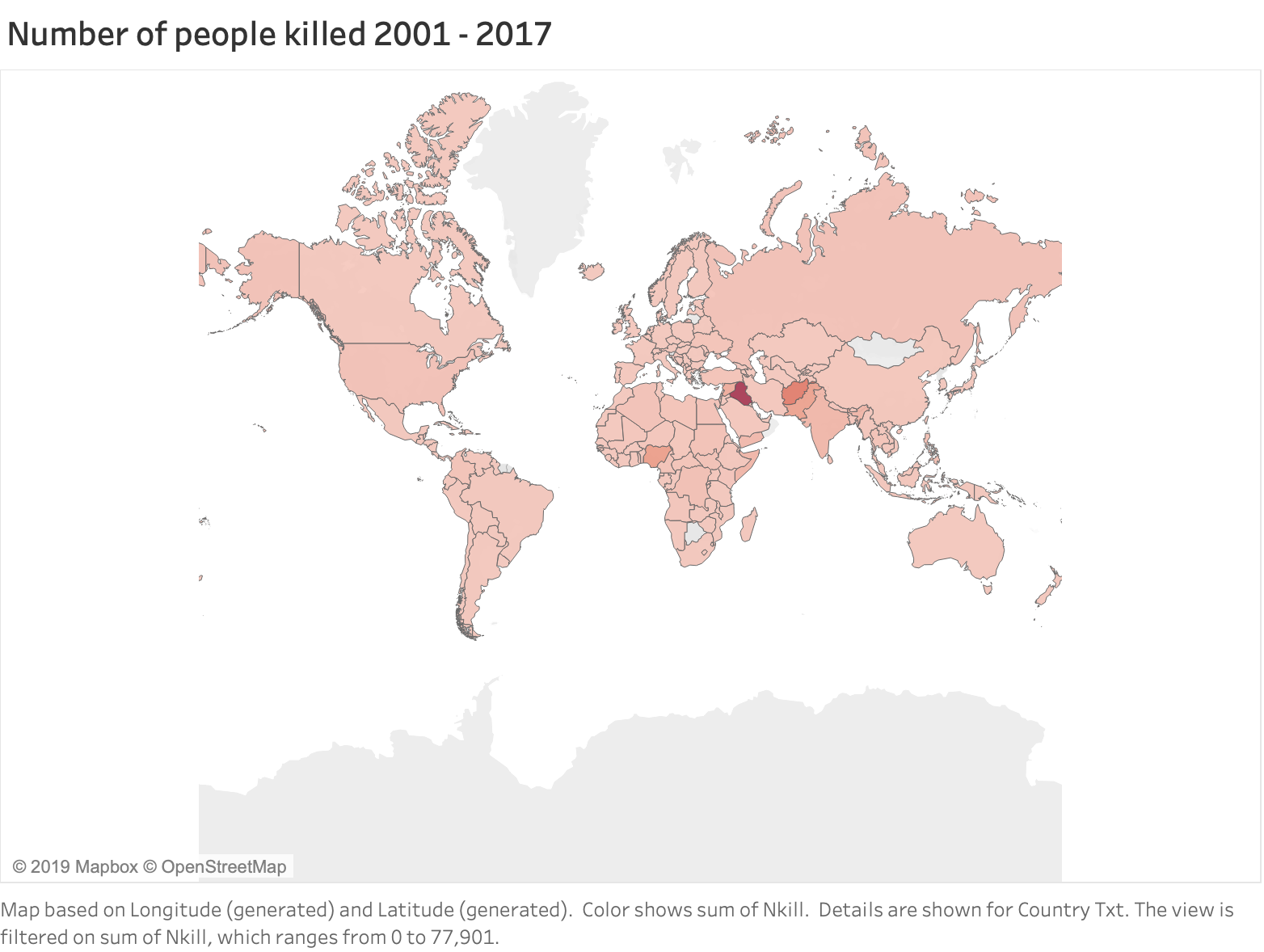
As seen in figure 1.0, the number of terrorist attacks are increasing over the years. There is a spike in the year 2014 where the number of attacks was more than 16,000.



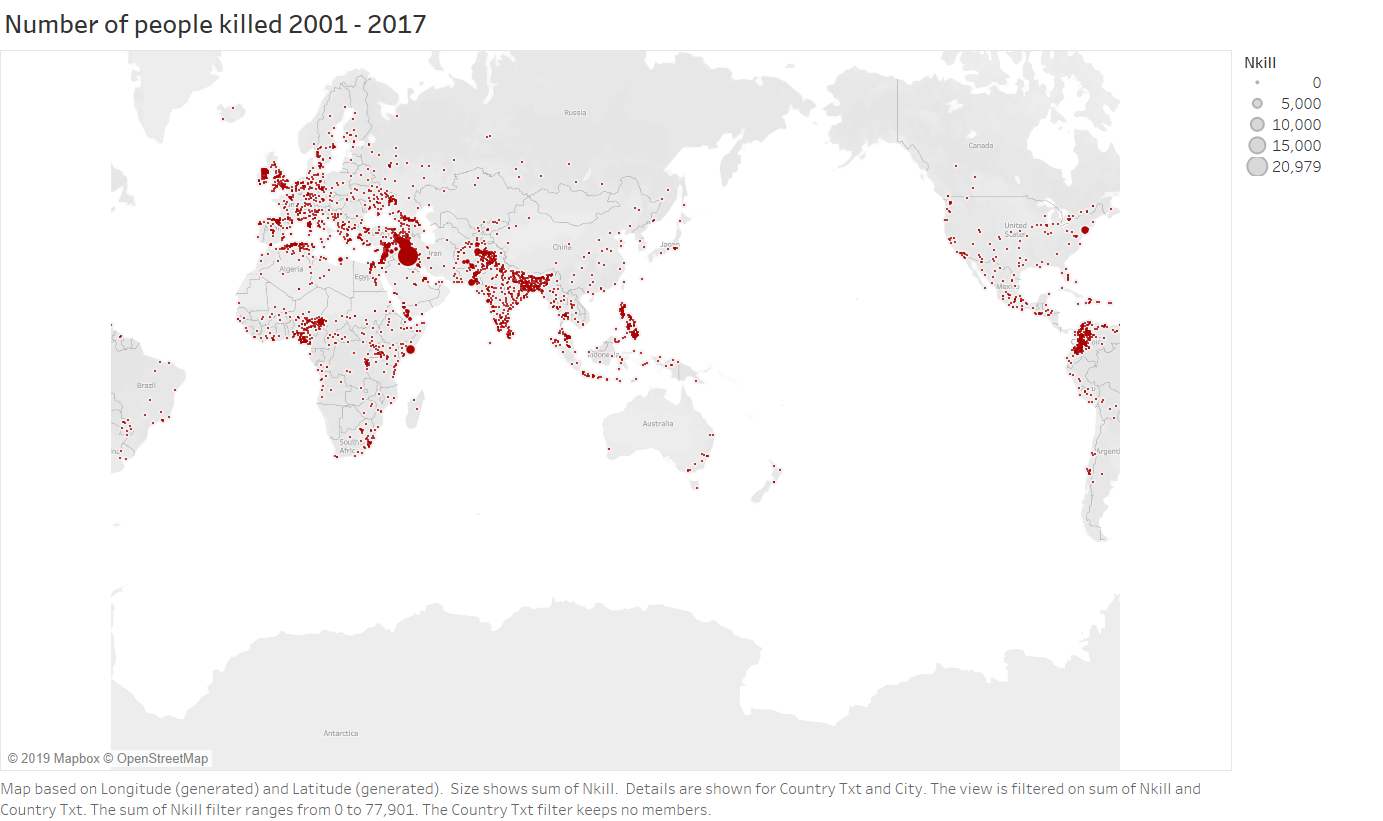
**Fig 2.0**

In coherence with the previous results from Fig 1.0, we can see that with the rise in attacks as there is a rise in the number of people killed.

**Which Countries are the most affected by terrorism (most people killed?) What is the intensity of attacks by region for a particular year?**

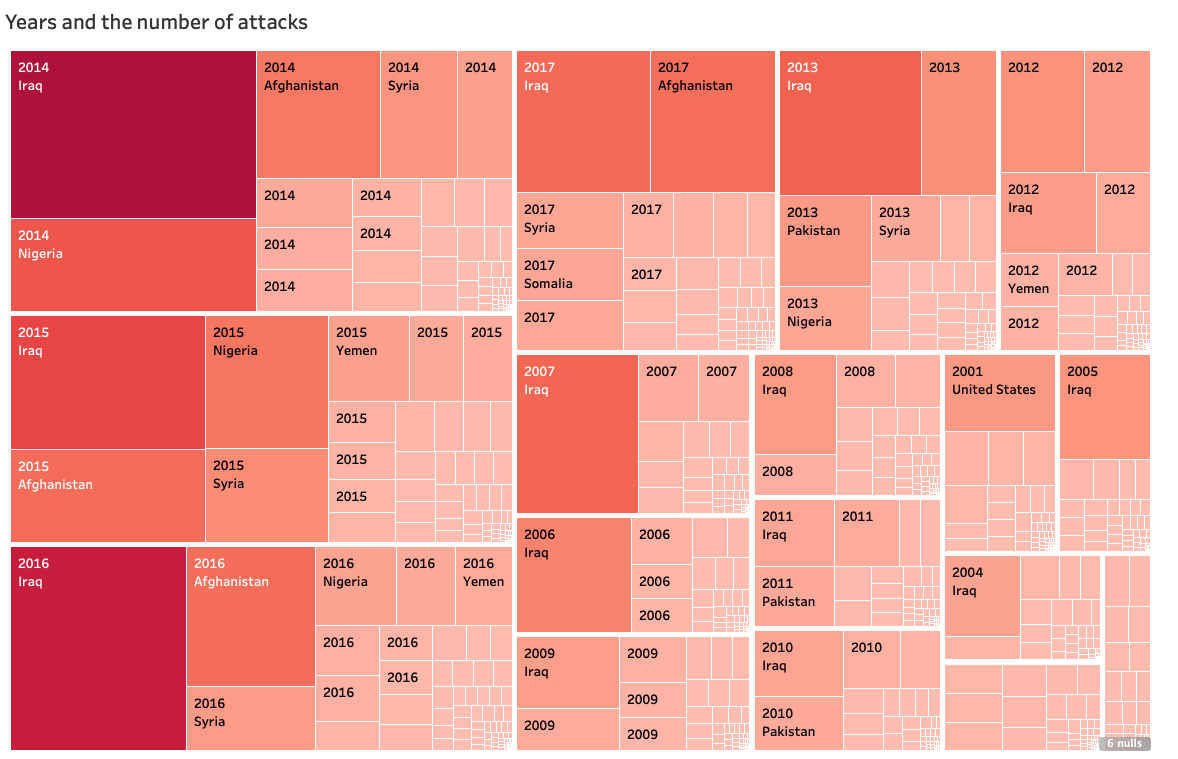


**Fig 3.0**

Figure 3.0 above showcases the heat map for countries with the number of people killed between 2001 and 2017. Figure 4.0 is an added visualization at the city level.

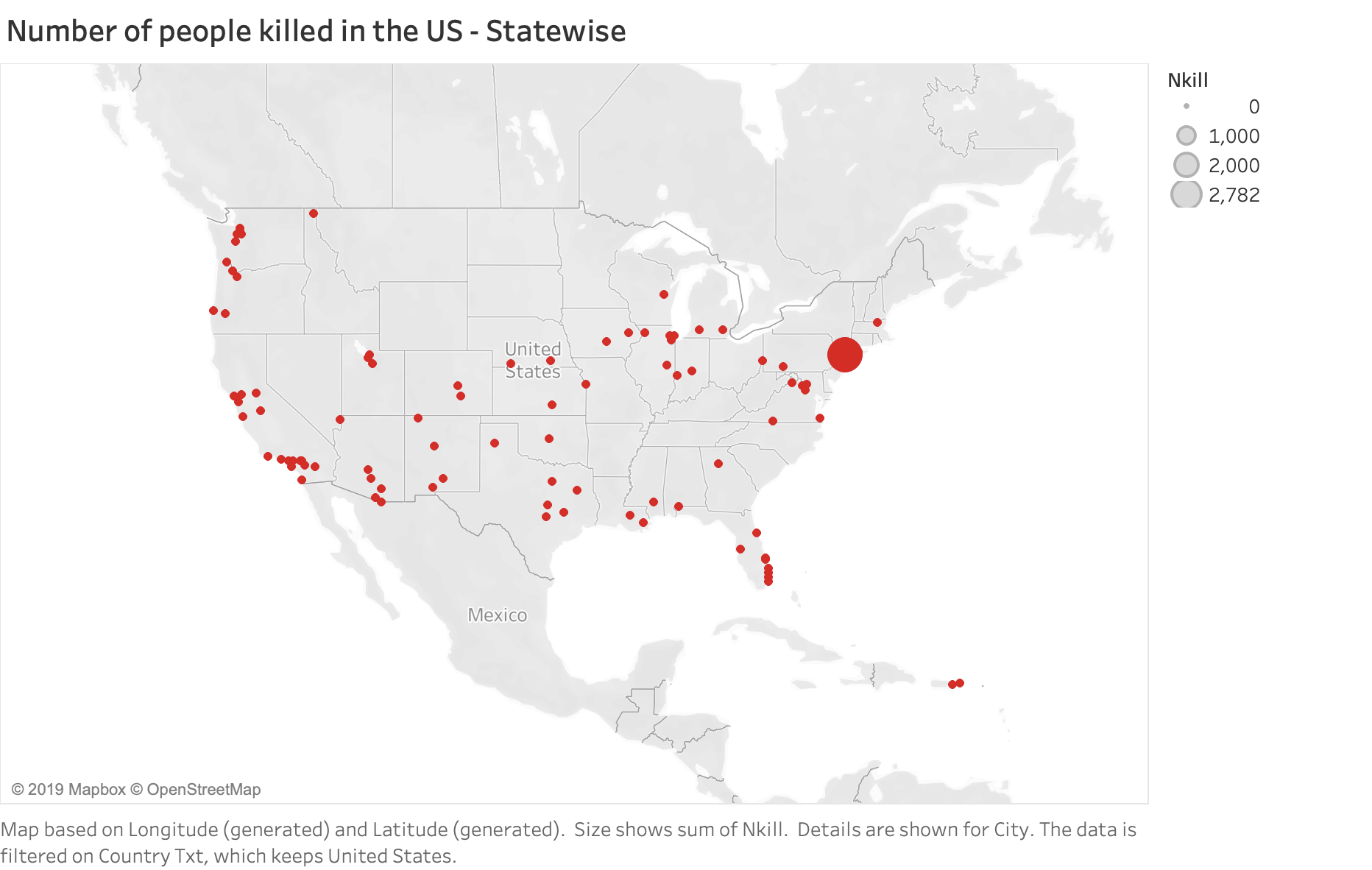
**Fig. 4.0**

Fig 5.0 showcases the heat map wherein it depicts the temporal, the region-wise intensity of attacks across the countries. It can be seen that Iraq has a significantly high density.



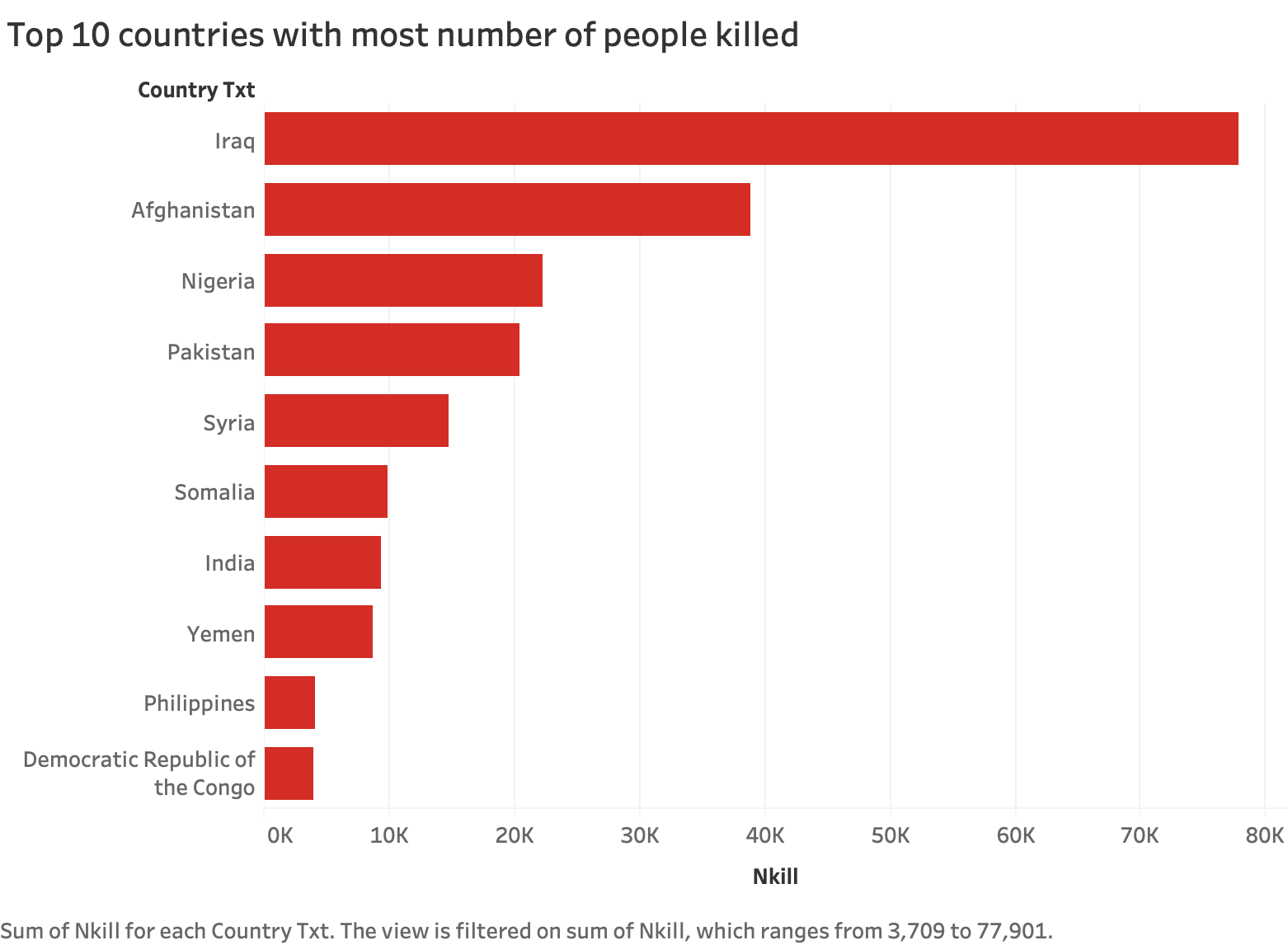
**Fig 5.0**

In Fig. 6.0 below we zoomed in further to the USA, and see a higher number of attacks on the East Coast.

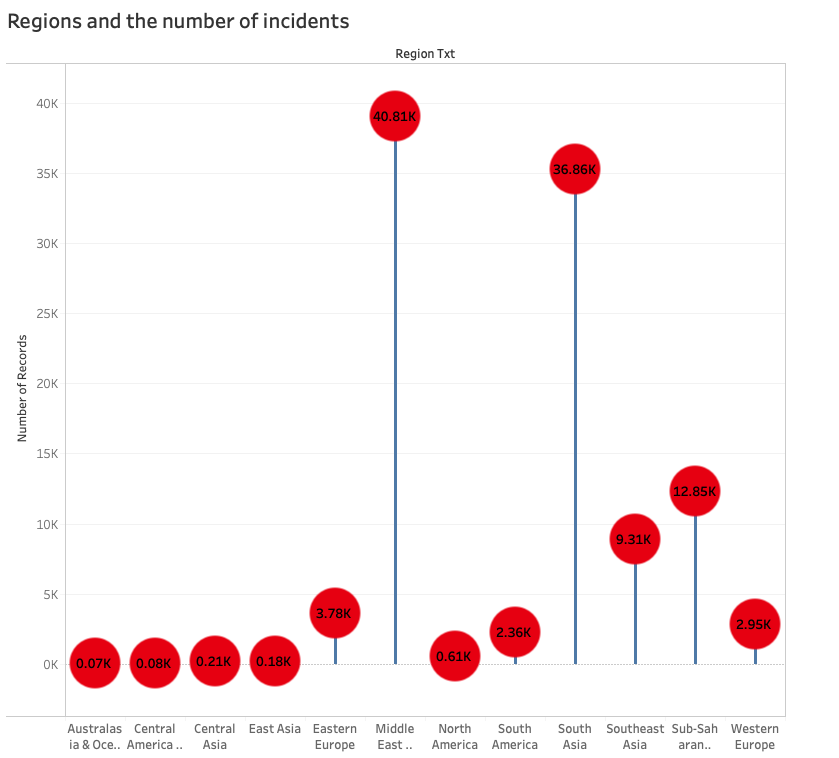


**Fig 6.0**

Figure 7.0 and 8.0 below highlights the top countries and regions impacted by terrorist activities in terms of people killed. We see Iraq as the most affected country, The surge in 2014 was largely due to increased activity of groups like the Islamic State in Iraq and Boko Haram in Nigeria. The civil war in Syria also spurred worldwide terrorist attacks.

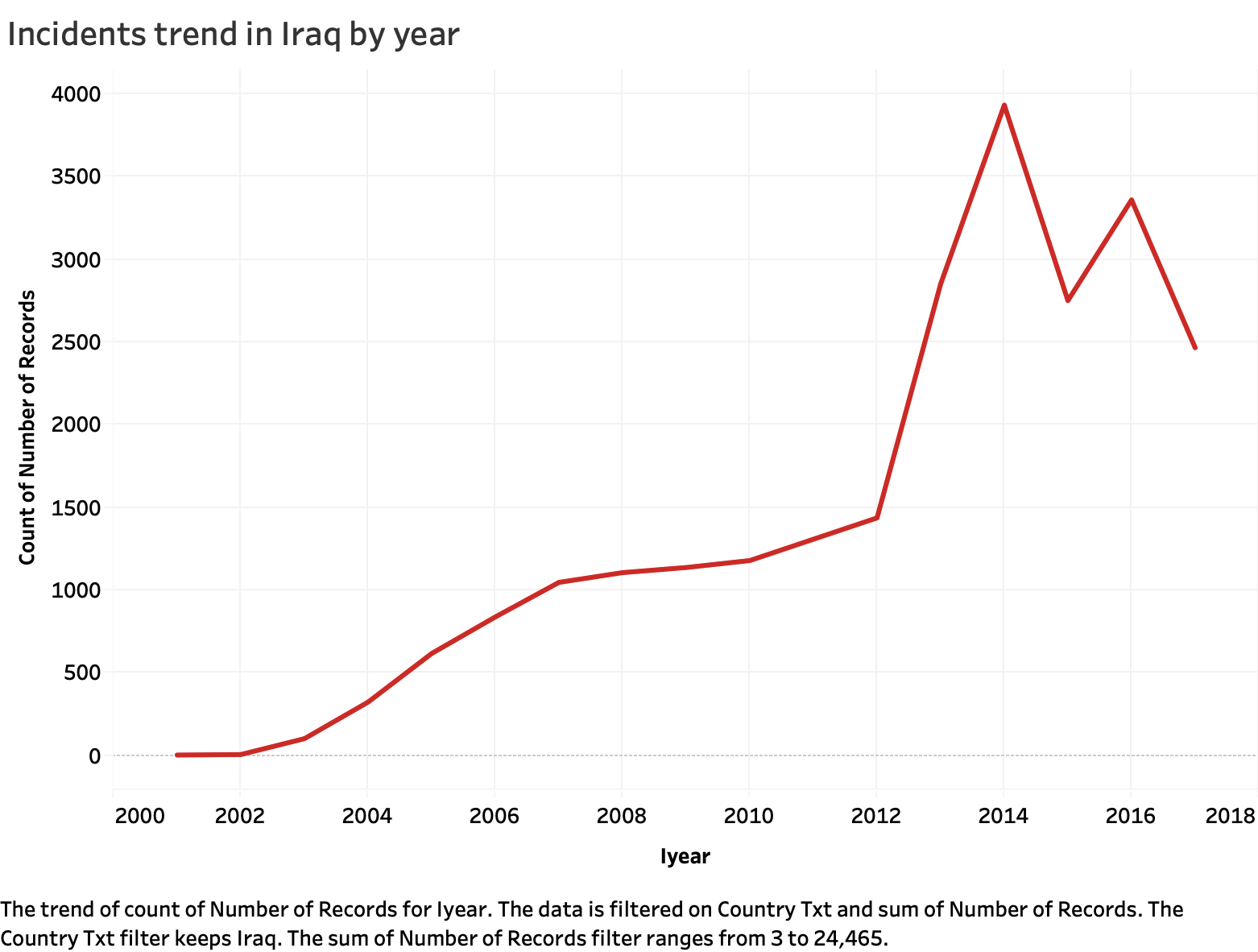


**Fig 7.0**

****

**Fig 8.0**

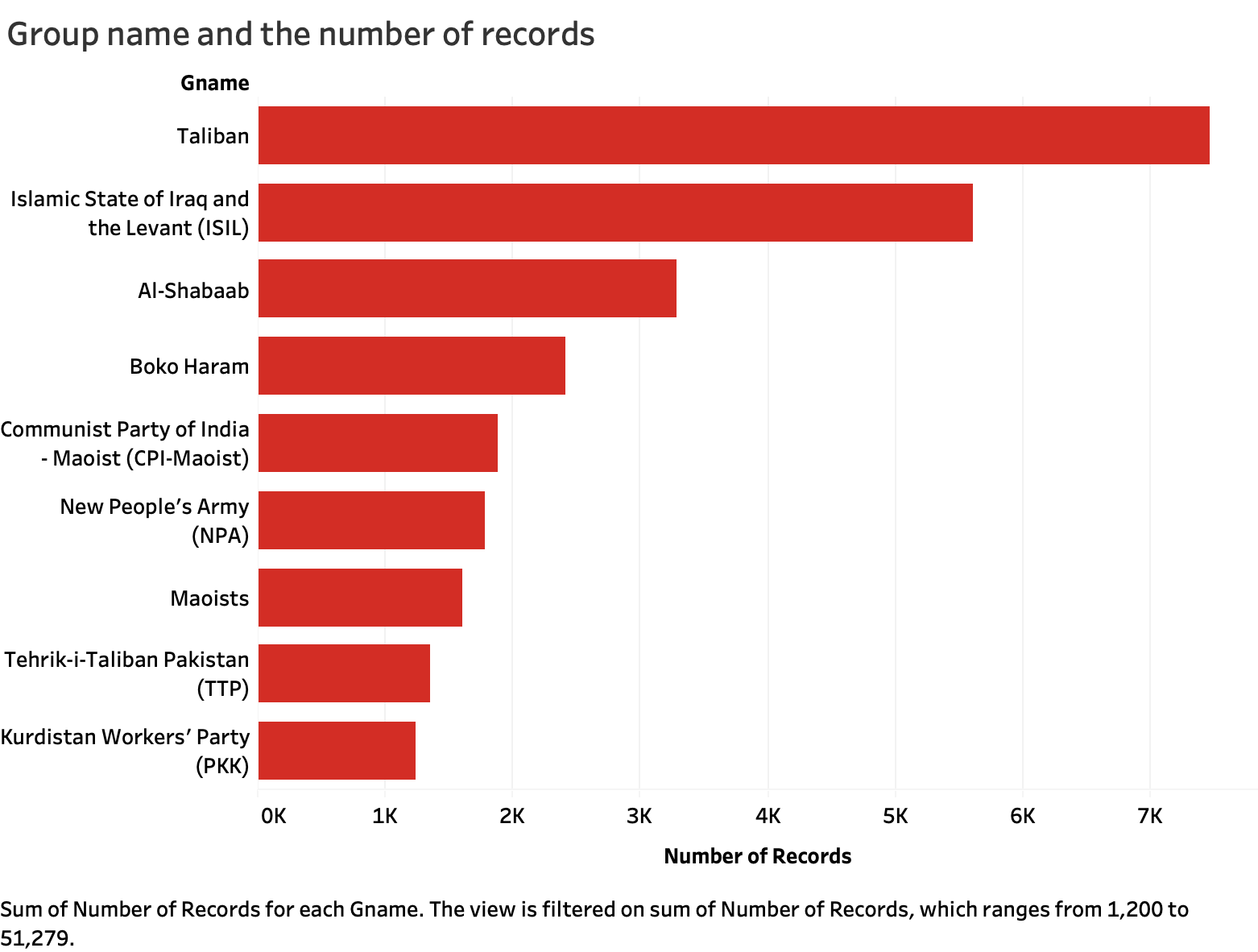
The Figure 9.0 we zoomed in to one of the top impacted region, Iraq and we could perhaps associate it with Iraqi Civil War which continued till 2017.



**Fig 9.0**

**Which terrorist groups have been most active?**

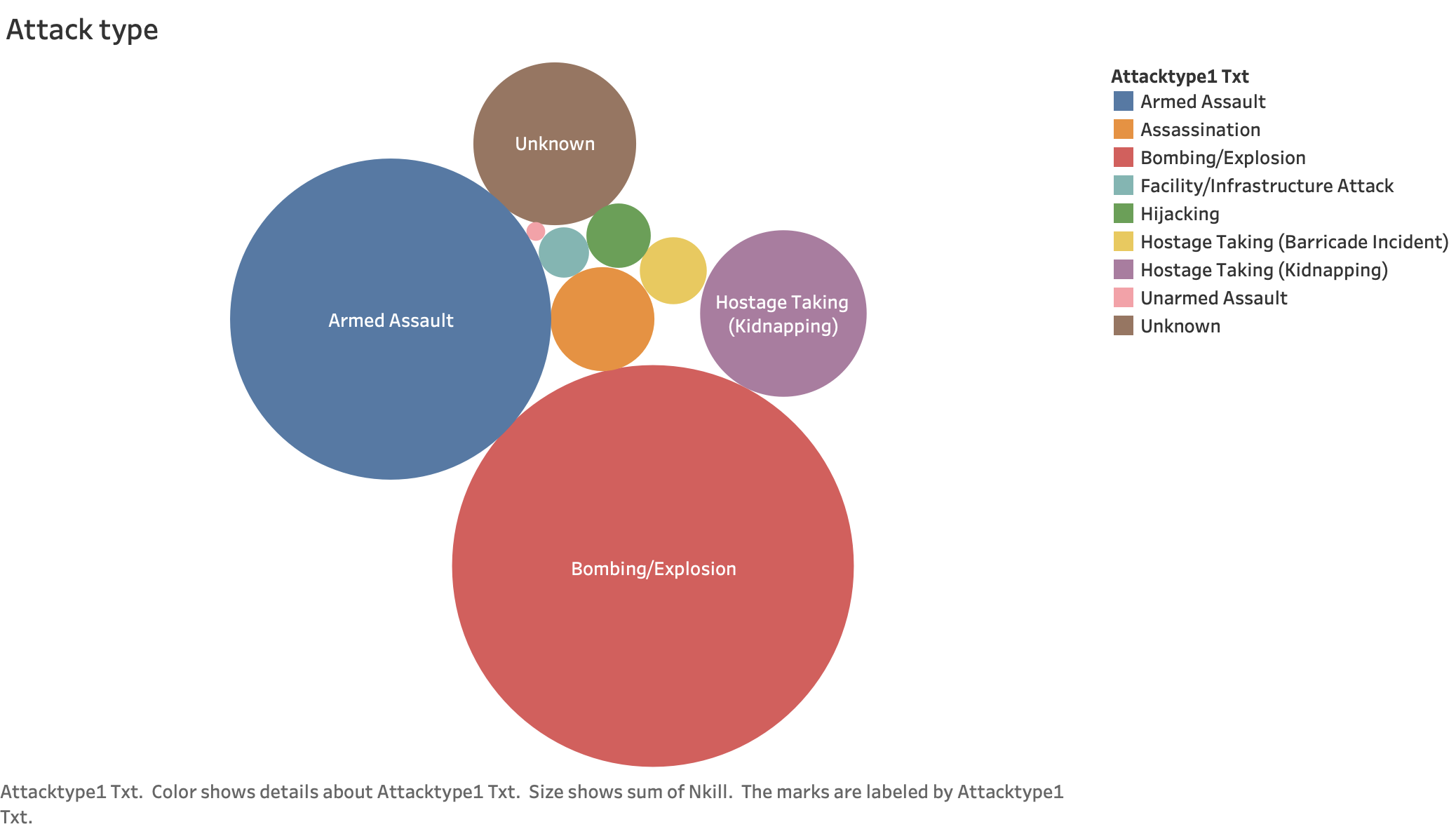
Figure 10.0 depicts the most active terrorist groups or organizations. These groups have taken part in and caused a large number of attacks.

****

**Fig 10.0**

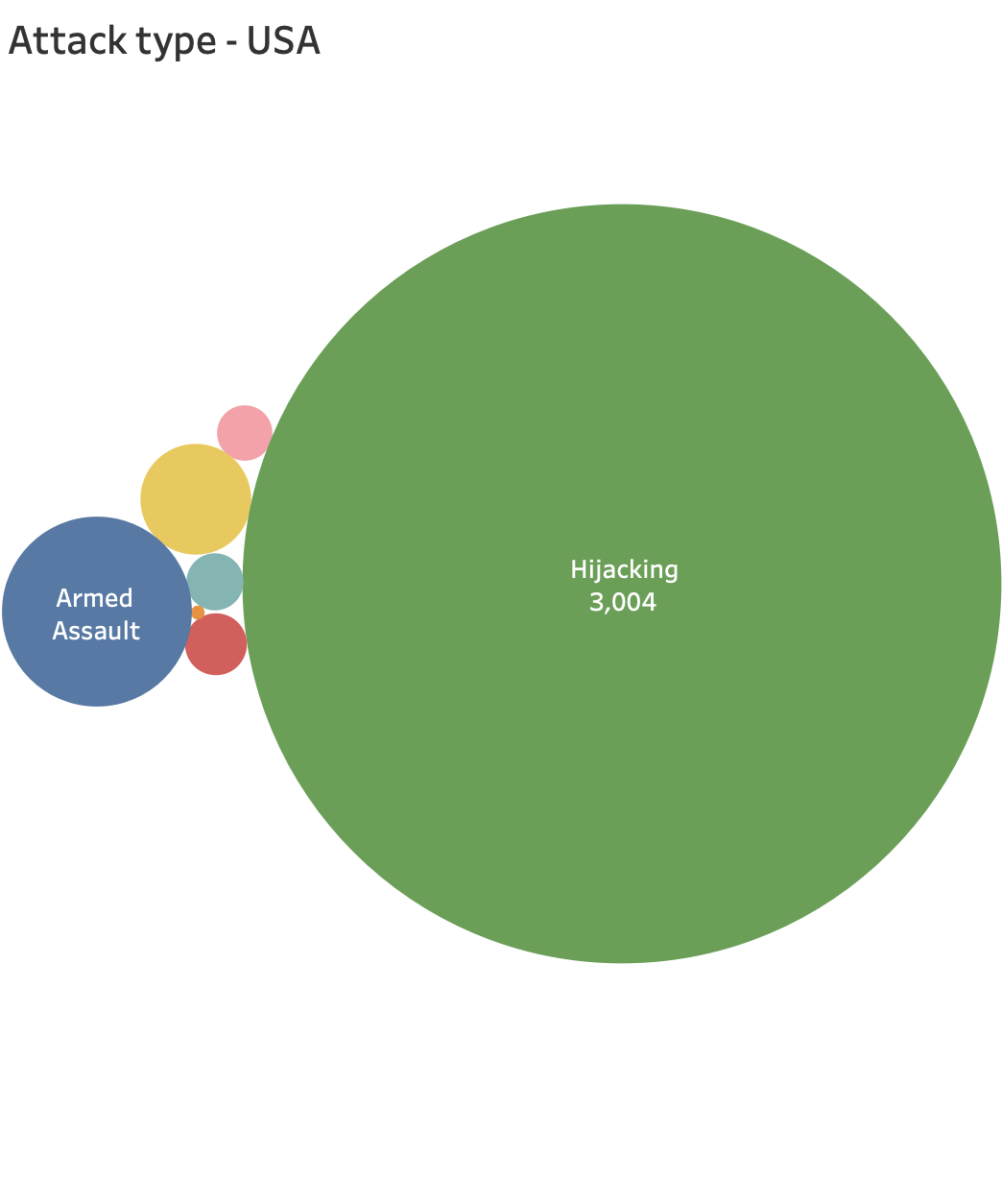
**What is the most common attack type?**

Figure11.0 below depicts the most common attack types. As seen from the figure, bombing/explosion is the most common attack type.



**Fig 11.0**

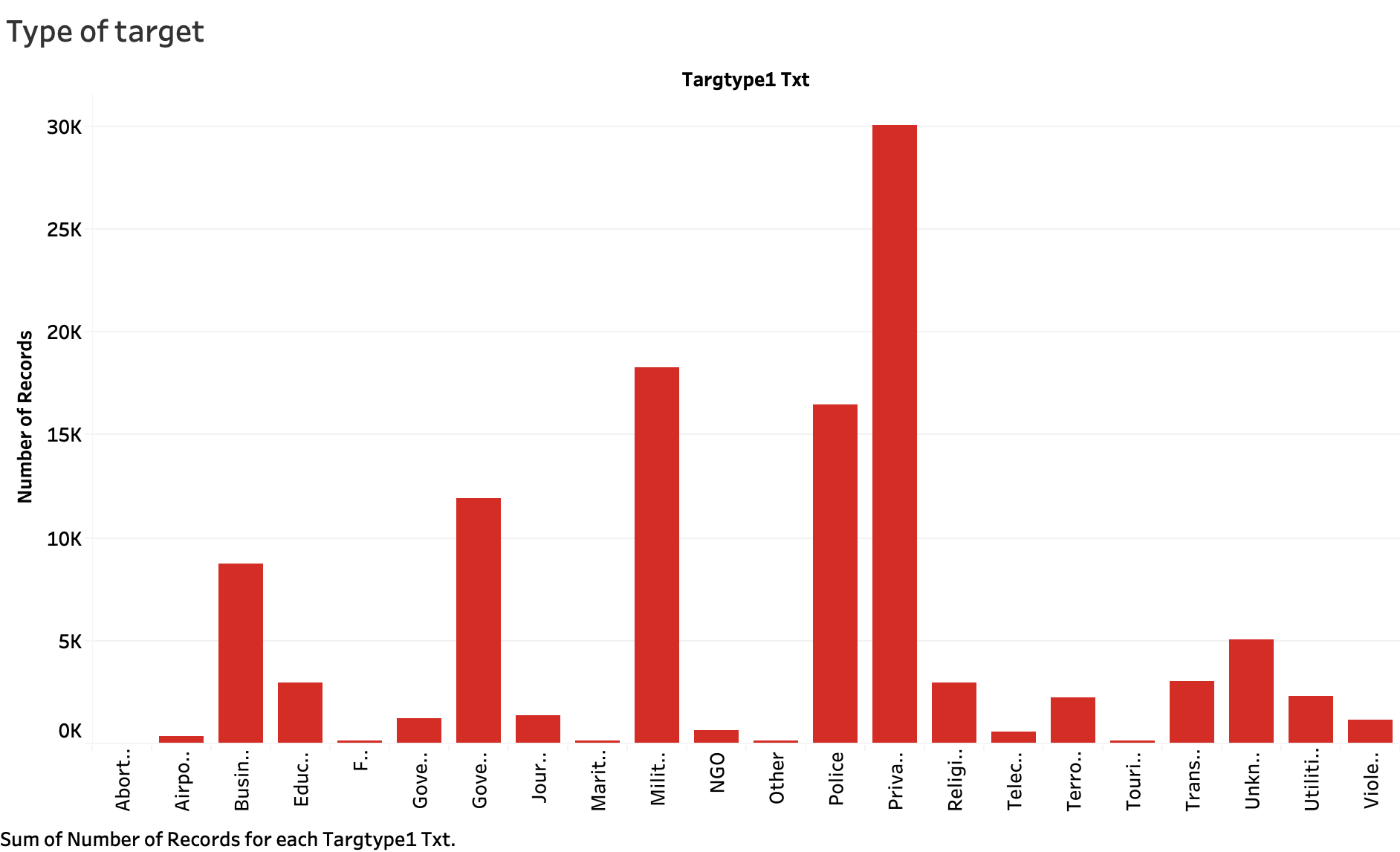
Figure 12.0 below depicts the most common attack types in the USA. Hijacking is the most common type of attack as can be seen.



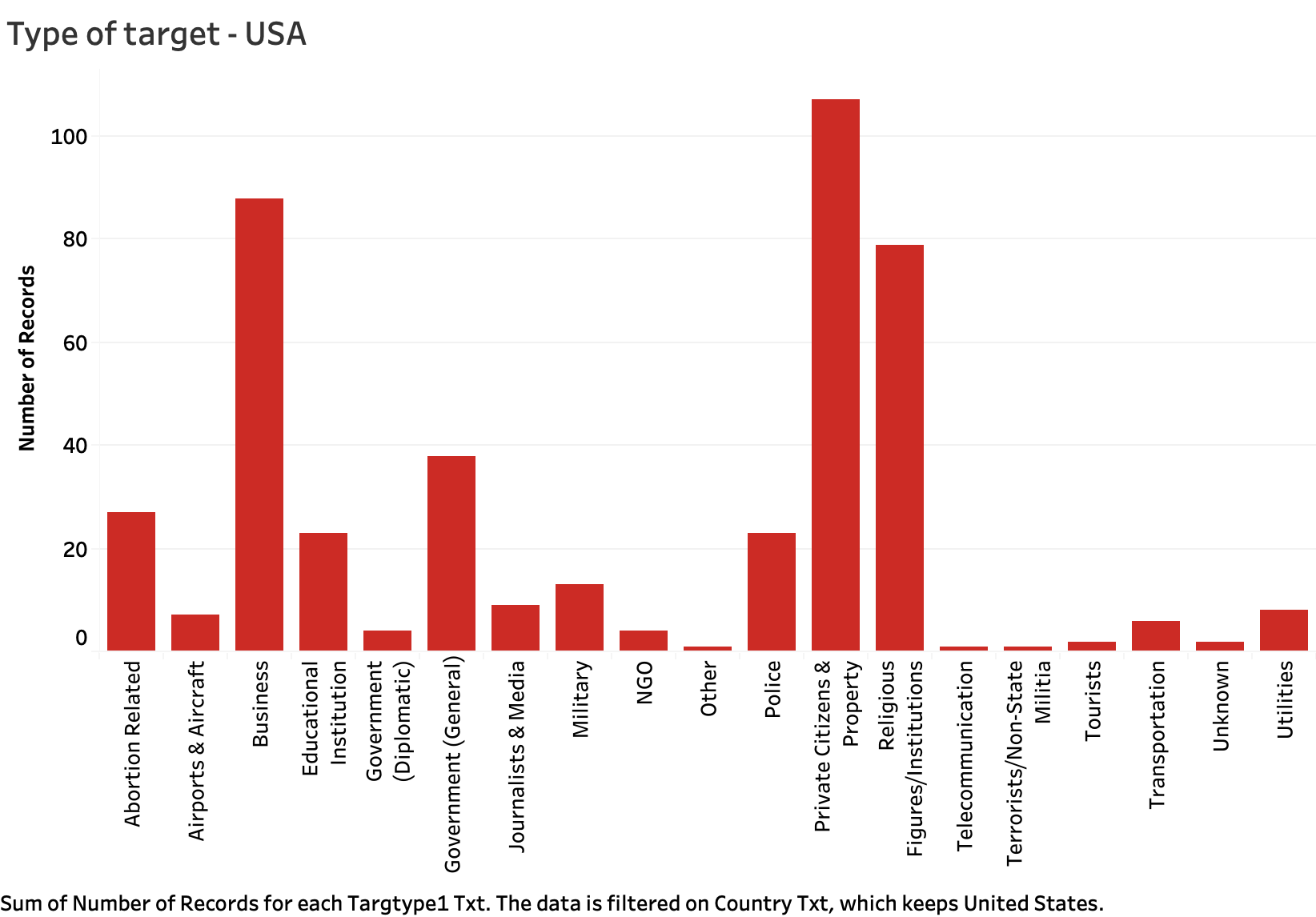
**Figure 12.0**

**What are the most common target types?**

Figures 13.0, 14.0 and 15.0 depict the most common target types in the world and in the USA with private citizens and religious figures showcasing the peak.



**Figure 13.0**



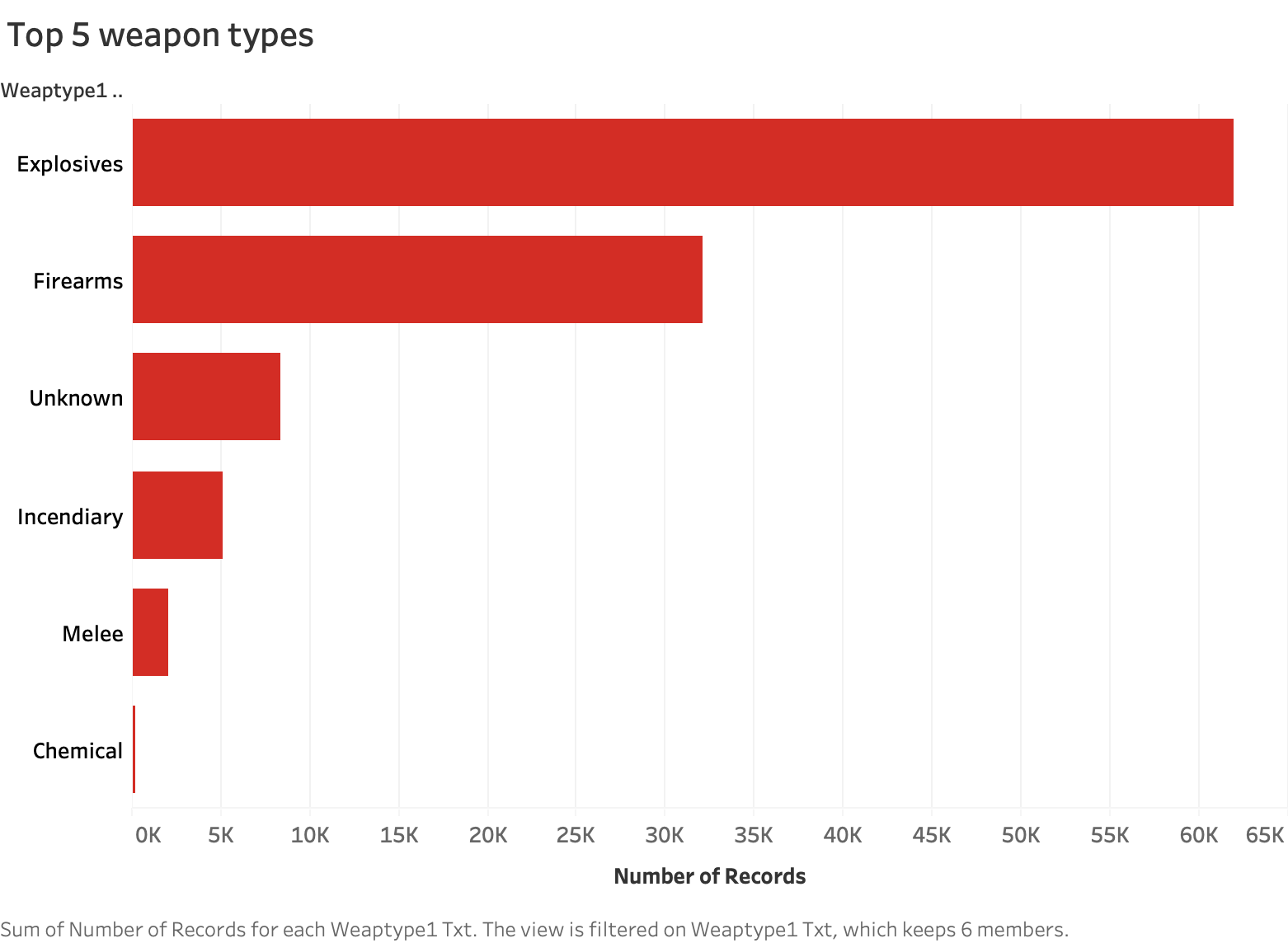
**Figure 14.0**



**Figure 15.0**

What are the weapons used in these attacks?

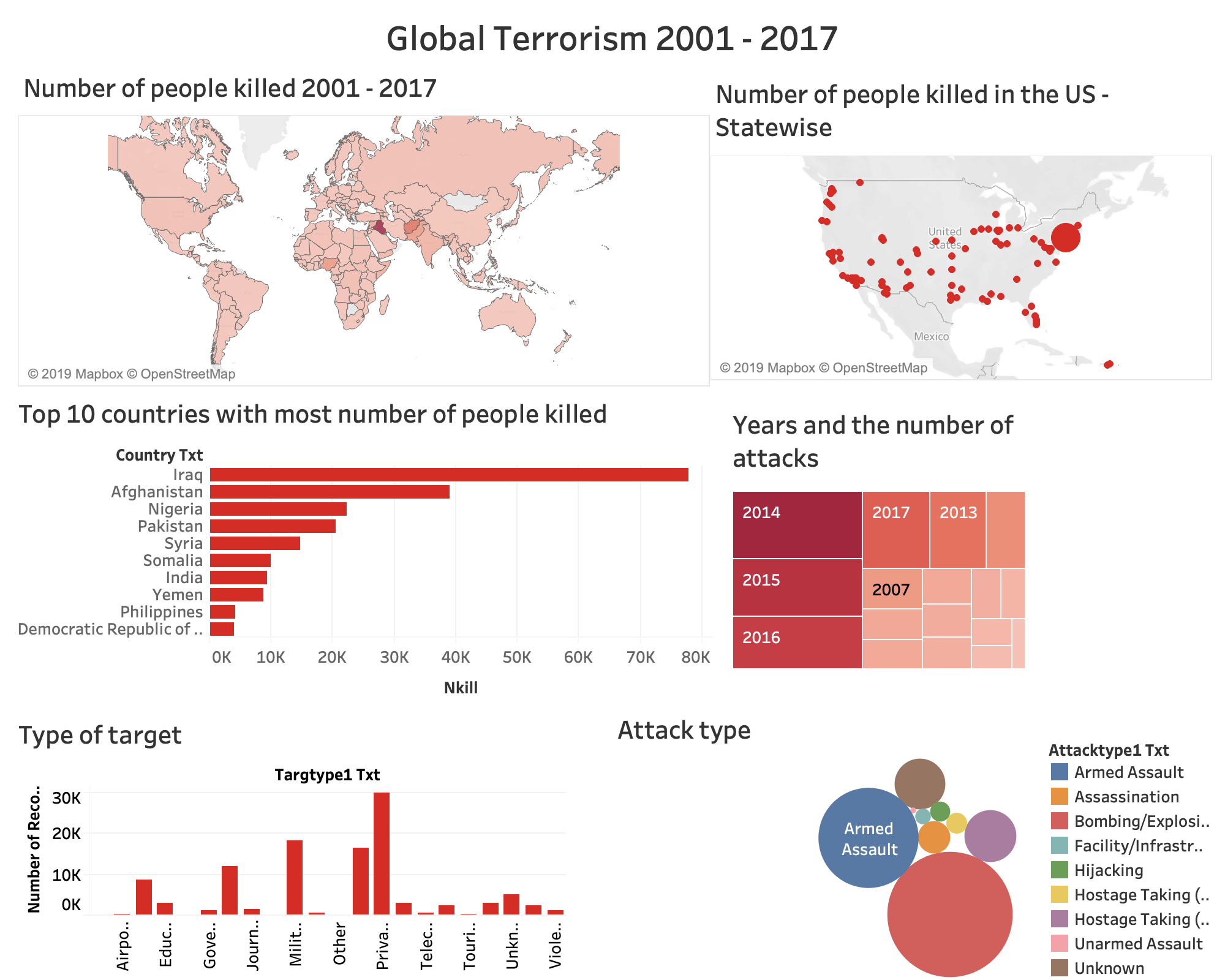
Figure 16.0 shows us in a bar graph the top 5 most common weapons used in the terrorist attacks.



**Figure 17.0**

The Fig 17.0 showcases one the dashboard of our analysis in Tableau. It must be noted that the dashboard is generic and we would ideally want to tailor it based on the specific stakeholders.

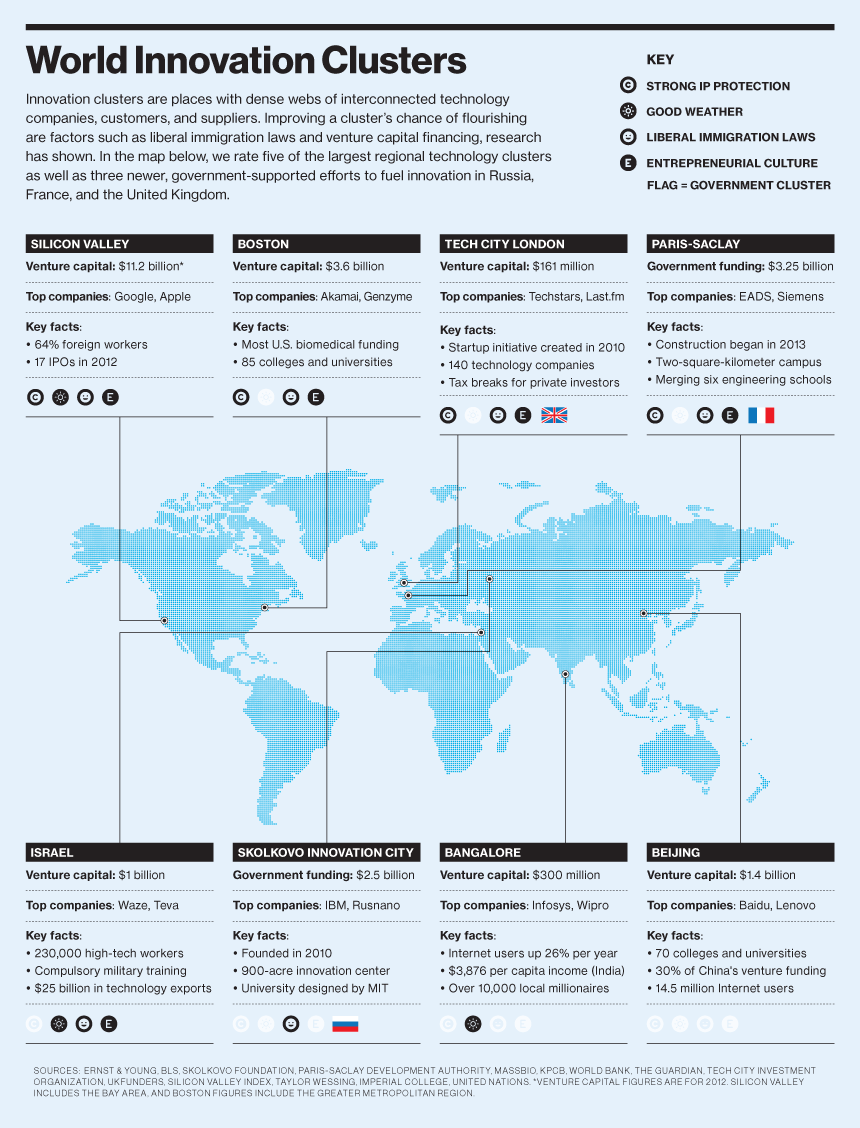
**Summary Dashboard**



**The implication of findings to specific stakeholders (example of the use cases)**

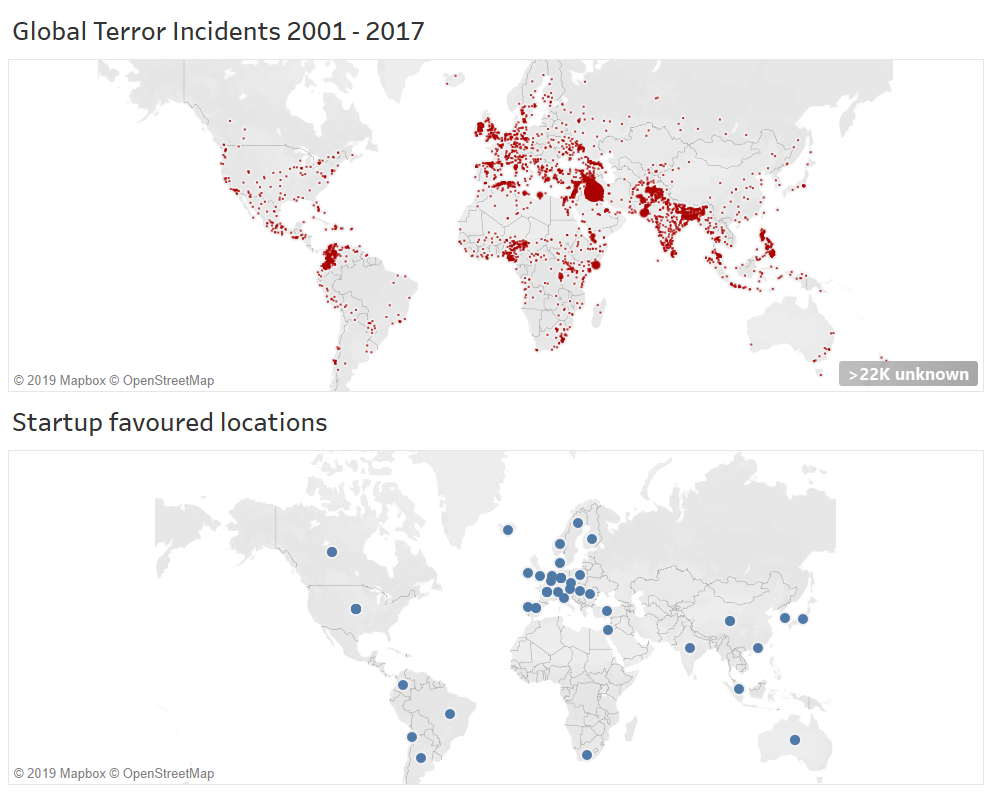
Individuals interested in the risk associated with moving their business to an international location could refer to the database to run a risk analysis for the region of interest.

Below is an infographic by MIT Technology Review, 2013 showing the concentration of innovation clusters at that time.

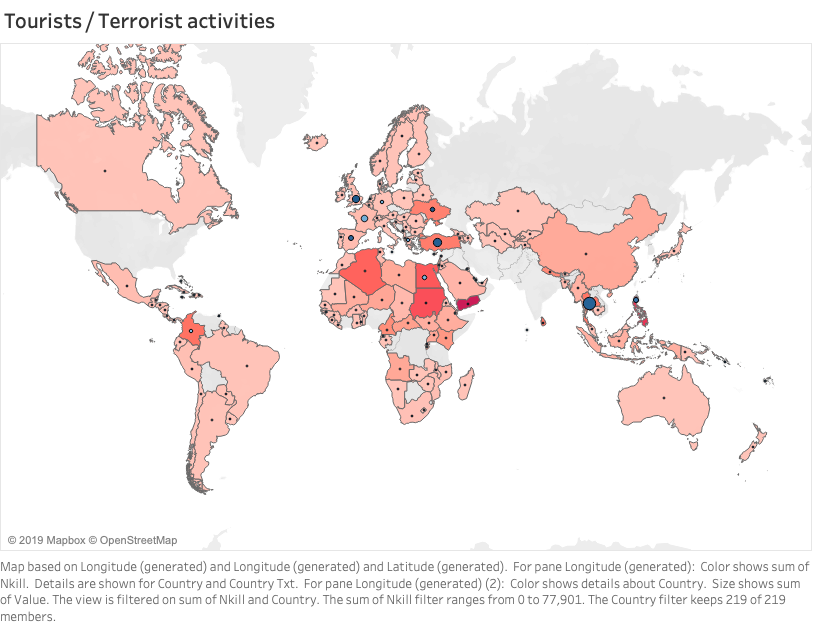


Reference: <https://www.technologyreview.com/s/517626/infographic-the-worlds-technology-hubs/>

The below map displays the location of the 50 best startup cities in 2019 as suggested by [Valuer](https://valuer.ai/blog/top-50-best-startup-cities/) . We mapped the data points from the article (50 best startup locations) in excel and performed visual analysis using Tableau. The Fig\_ below showcases a dashboard we created to showcase correlations between two data sets, and we can see that the 50 best startup cities fall in regions with comparatively low terrorist activities (limitation: we have limited resolution data for startup - only 50 cities, and it has lead to a situation where a country as big as the USA has only one data point for startup favored location visualization, as compared to many data points for region like Europe where there are comparatively smaller countries)



For our **2nd use case** we tried to look into tourist flows (we used Tourist/visitor arrivals and tourism expenditure data from UN data <https://data.un.org/>) and used Tableau to match it with our Terrorist data. Fig. below showcases our analysis where heat map shows the no. of people killed by terrorist activities and circles shows the number of tourists to these countries. Given the scope of this assignment, the below analysis is very crude (for instance we have not filtered by years, or zoom in to city-level), however, this exercise exemplifies the breadth of the use cases and corresponding stakeholders.



Conclusion:

Data shows terrorism activities peaked in 2014. There is a clear link to countries like Iraq - civil war. Individuals interested in the risk associated with their business could refer to the database to run a risk analysis for the region of interest.