

**Global Terrorism Data Analysis using Tableau**

**Sumeet Shivgand**

**R00182850**

# Assignment 1

## Introduction:

Tableau is a Business Intelligence and analytics software that helps data scientists and researcher analyze data quickly, easily, and usefully. Tableau let you connect to data in the cloud (i.e. big data, SQL, spreadsheet, or Google Analytics and Salesforce), build calculations from existing data, create reference lines and forecasts, make trend analyses, regressions, and correlations, with his respective statistical summaries. This software does not require writing code and the power users can pivot, split, and manage metadata to optimize data sources. The main objective of this assignment is to show how the visualization can be implemented or used in Tableau.

Crime awareness is crucial in today's world and one must always be aware of the criminal activities within the area they live. The crimes committed can be measured against factors like years, crime types, area, bureau and so on. We are working on global terrorism dataset which consist of all crimes registered from 1970 to 2017 including fields such as date - date on which crime was reported, event id – the crime no, crime description such as types of attack/types of target/types of weapon, crime location such as city/country/region, latitude and longitude details.

## Exploratory Data Analysis (EDA):

The given dataset consists of around 1,80,000 observations and 135 columns. There are some NA or missing values which we are going to remove before proceeding with visualization. Also, we are going to remove the unnecessary columns which are not important for visualization. After doing EDA using R, we reduced the observations to around 127000 and columns to 22. Now our data is cleaned and ready for visualization using tableau. Before proceeding to tableau, we first summarize some columns to get insights of it.

region_txt		attacktype1_txt	
Middle East & North Africa	:39526	Bombing/Explosion	:58225
South Asia	:35375	Armed Assault	:31645
Sub-Saharan Africa	:13780	Assassination	:15060
South America	:10022	Hostage Taking (Kidnapping)	: 8747
Southeast Asia	: 9384	Facility/Infrastructure Attack	: 6530
Western Europe	: 7990	Unknown	: 5309
(Other)	:11256	(Other)	: 1817

Here, the region is nothing but the continents. The region which is most affected by terrorism is 'Middle East & North Africa' with 39526 records. Similarly, 'Bombing/Explosion' is the type of attack which holds the maximum number of records with 58225.

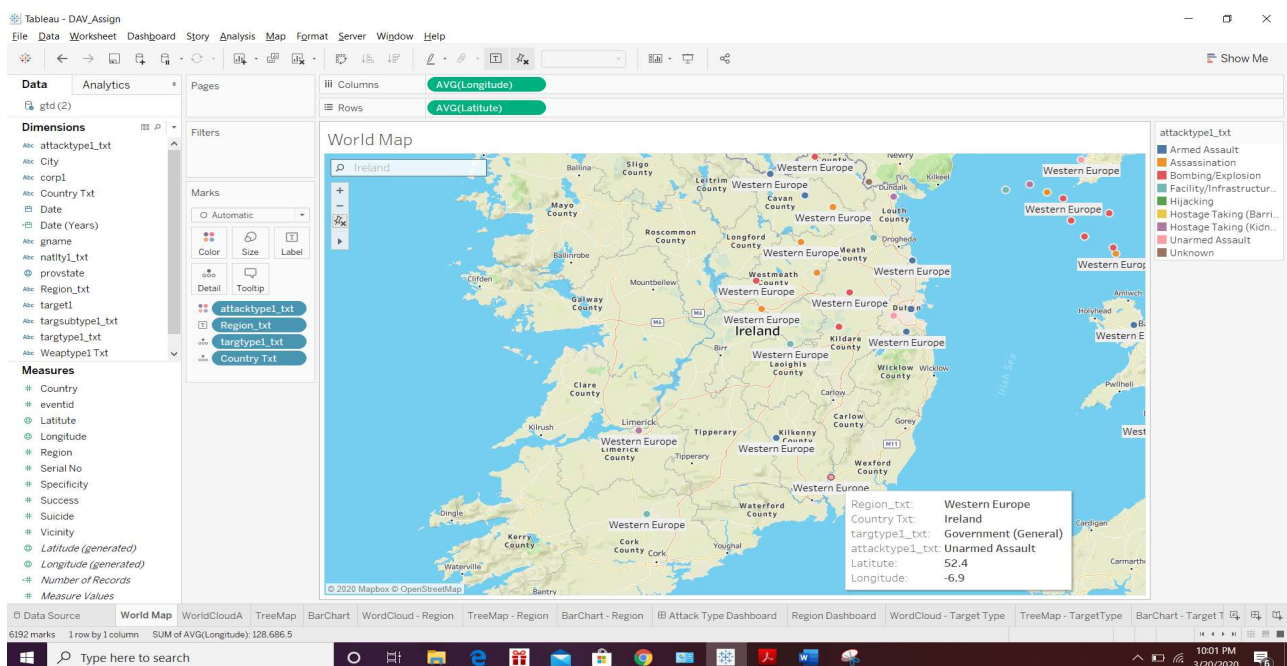
## Begins with Tableau:

In the below data source, we can see one table with the observations in the Rows and the variables in the Columns. This is nothing but the tableau framework.

#	gtd.csv	gtd.csv	gtd.csv	gtd.csv	gtd.csv	gtd.csv	gtd.csv	gtd.csv	gtd.csv	gtd.csv	gtd.csv	gtd.csv
	Specificity	Vicinity	Success	Suicide	Attack Type	Target Type	Target Subtype	Corporation	Target	Nationality	Group Name	Weapon Type
1	0	0	1	0	Armed Assault	Police	Police Building (...)	Cairo Police Dep...	Cairo Police Hea...	United States	Black Nationalists	Firearms
1	0	0	0	0	Assassination	Police	Police Security F...	Uruguayan Police	Juan Maria de Lu...	Uruguay	Tupamaros (Uru...	Firearms
1	0	1	0	0	Bombing/Explosi...	Utilities	Electricity	Pacific Gas & Ele...	Edes Substation	United States	Unknown	Explosive
1	0	1	0	0	Facility/Infrastru...	Military	Military Recruit...	R.O.T.C.	R.O.T.C. offices at...	United States	New Year's Gang	Incendiary
1	0	1	0	0	Facility/Infrastru...	Government (Ge...	Government Buil...	Selective Service	Selective Service...	United States	New Year's Gang	Incendiary
1	0	1	0	0	Facility/Infrastru...	Military	Military Recruit...	Army Recruiting ...	Army Recruiting ...	United States	Left-Wing Milita...	Incendiary
1	0	1	0	0	Hijacking	Airports & Aircraft	Aircraft (not at a...	Trans World Airli...	Flight 802 Boein...	United States	Unknown	Firearms
1	0	1	0	0	Facility/Infrastru...	Government (Ge...	Government Buil...	U.S. Government ...	Packard Properti...	United States	Left-Wing Milita...	Incendiary
1	0	1	0	0	Facility/Infrastru...	Business	Retail/Grocery/B...	American owned ...	Baker's Store	United States	Armed Command...	Incendiary
1	0	1	0	0	Bombing/Explosi...	Educational Insti...	School/Universit...	High School	James Madison ...	United States	Black Nationalists	Explosive

## 1. The Data Source Window

This is the first sheet which includes a world map where all terrorism take place. From this, we can easily come to know which part of world has recorded maximum terrorism attacks. As an example, we used the column names such as 'attack type', 'Region (Continent)', 'Target type' and 'Country'. Afterwards, all the data points of this columns get loaded into map and then if we hover the cursor over any points, we can see the details of that attack.



## 2. World Map Window

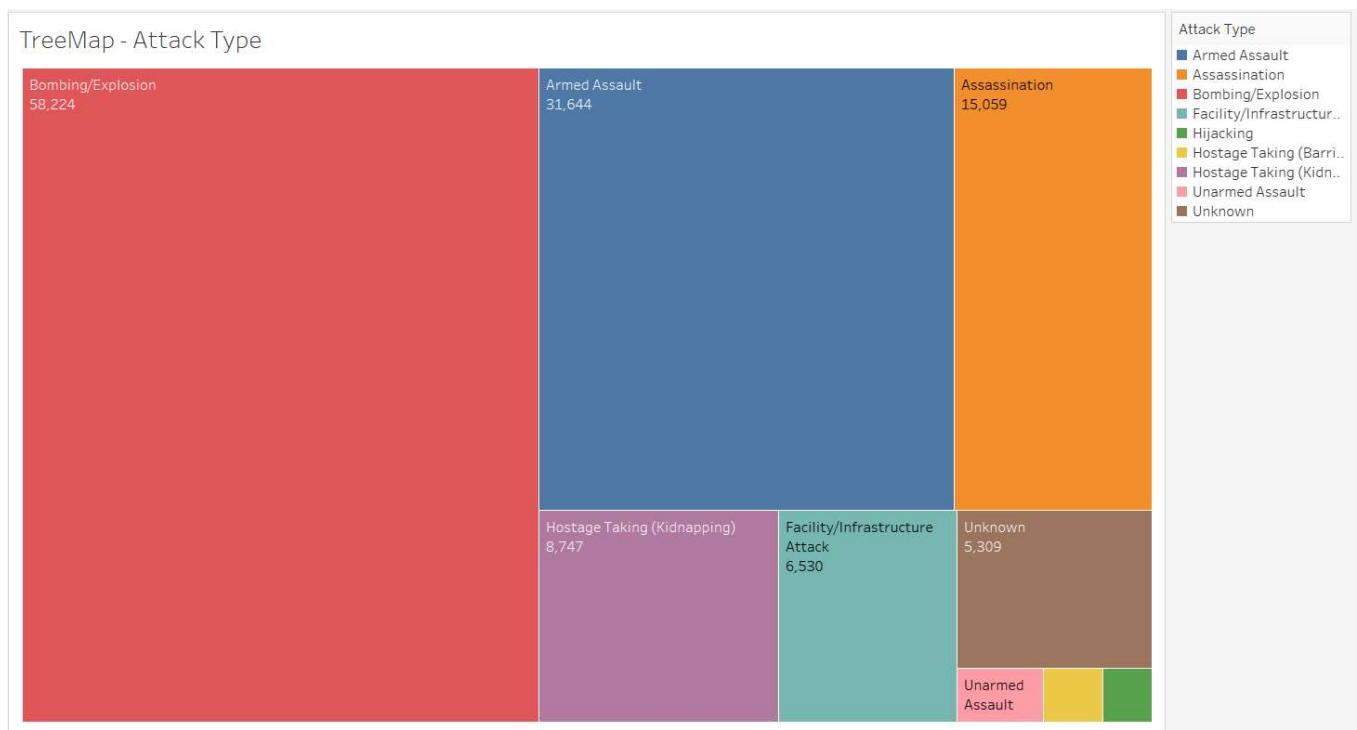
Here, we are using word cloud for ‘attack type’ columns. Word clouds are a method for visually presenting text data. They are popular for text analysis because they make it easy to spot word frequencies. The more frequent the word is used, the larger and bolder it is displayed.

WordCloud - Attack Type



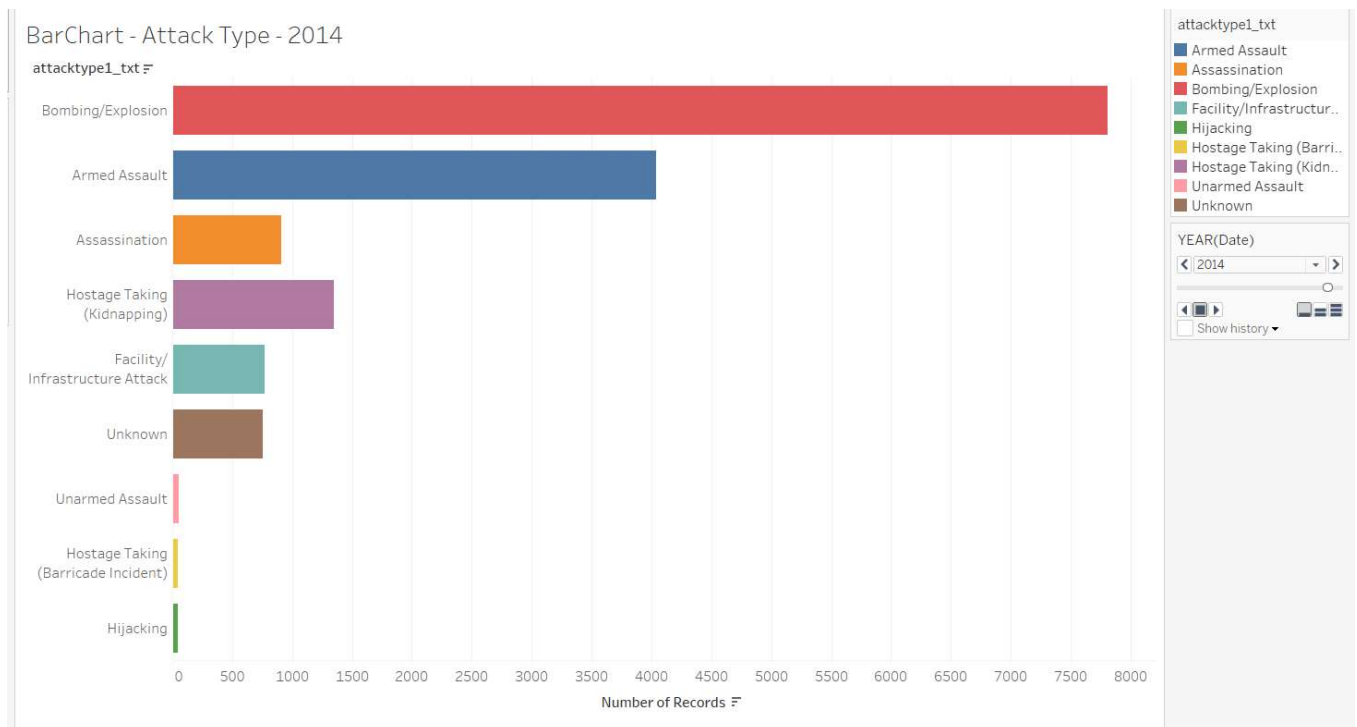
### 3. Word Cloud – Attack type

Treemapping is a data visualization technique that is used to display hierarchical data using nested rectangles. The given below tree map compares the ‘attack type’ for the crime dataset for the year 1970 to 2017. This tree map displays two measures and one dimension at the same time. We can clearly see from the below tree map which ‘attack type’ holds most of the records.



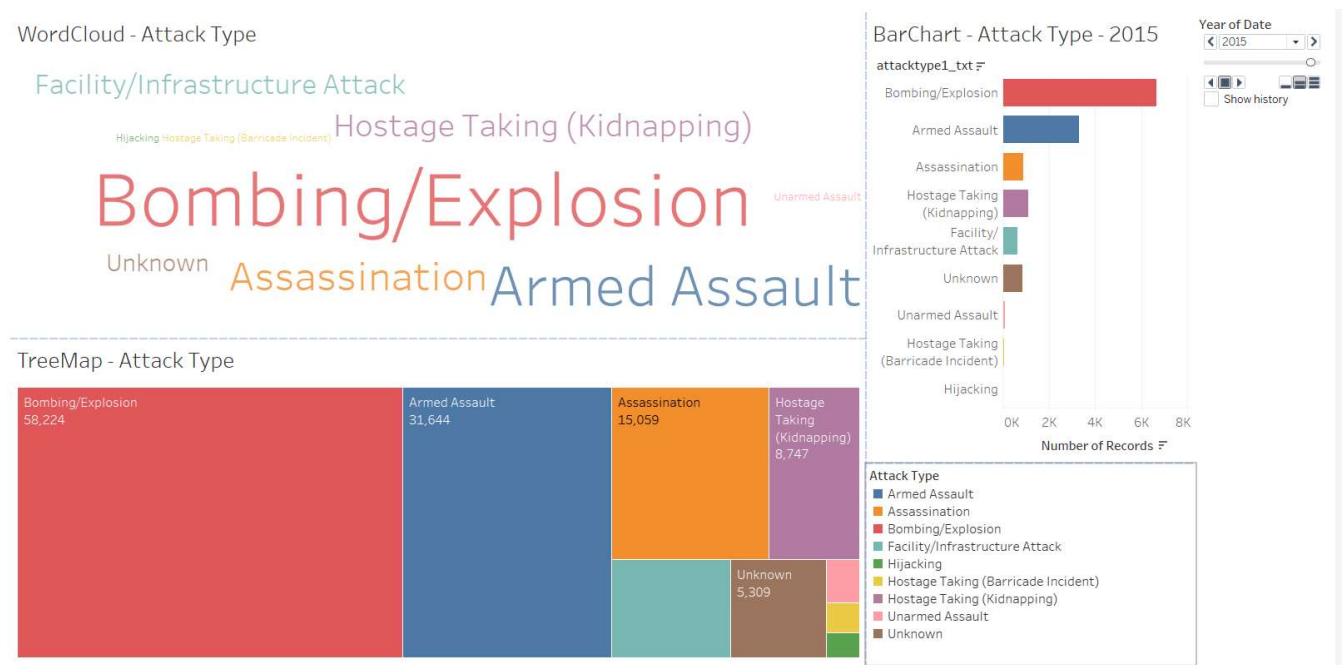
### 4. Tree Map – Attack Type

This given bar chart displays crime records within each 'attack type' from 1970 to 2017. For example, in the year 2014, we can clearly interpret in the given bar chart 'Bombing/Explosion' holds the highest number of records whereas 'Hijacking' holds the minimum number of records.



## 5. Bar Chart – Attack Type

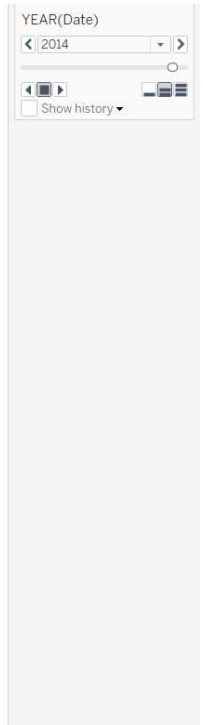
The given below is the dashboard for 'attack type'. This is an easy way to see as all the graph can be visualize at the same time rather than seeing it differently. Also, we had set a filter on year so if we change the year then we get a corresponding record for that year. Here, we merge the word cloud, tree map and bar chart.



## 6. Dashboard – Attack Type

The bubble chart shows the number of crimes committed per region or continents categories for each year. There will be different bubble charts for every year i.e. from 1970 to 2017.

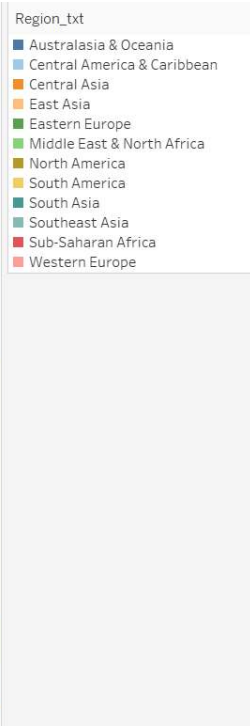
Bubblegraph - Region - 2014



## 7. Bubble graph – Region/Continents

The given below tree map compares the ‘region/continents’ for the crime dataset for the year 1970 to 2017. This tree map displays two measures and one dimension at the same time. We can clearly see from the below tree map which ‘region/continents’ holds most of the records. ‘Middle East & North Africa’ has the maximum number of cases compared to all other regions/continents.

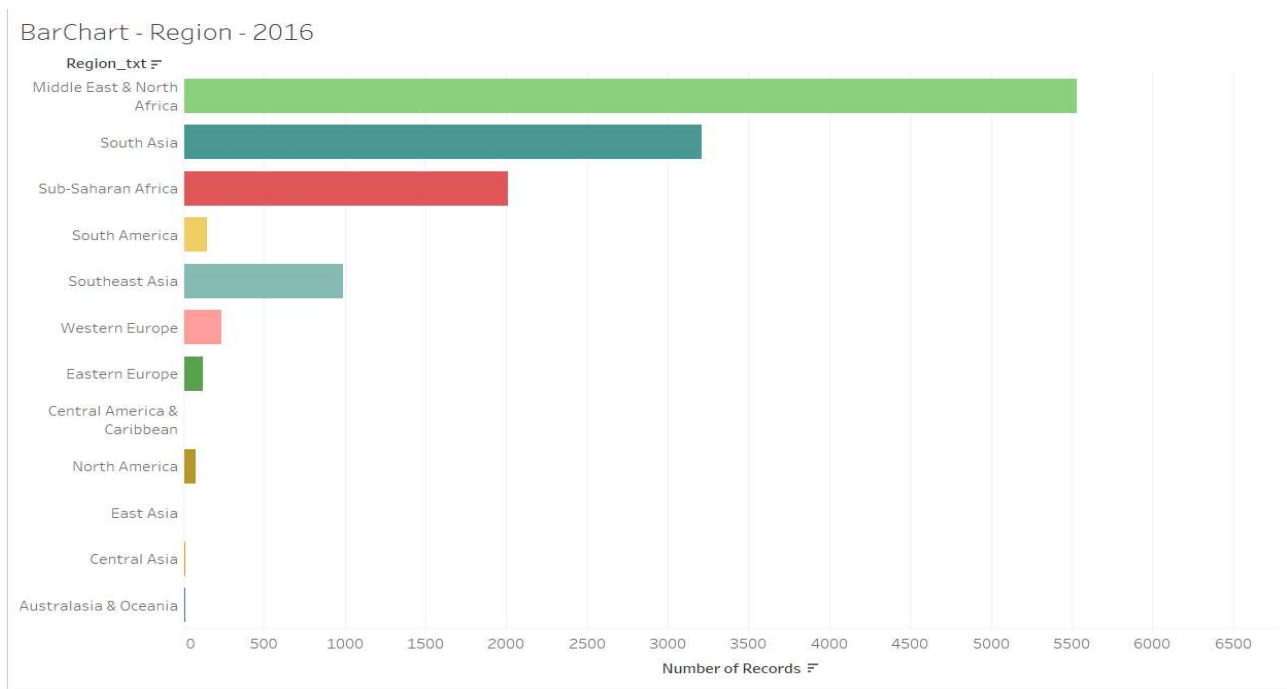
TreeMap - Region



## 8. Tree Map – Region

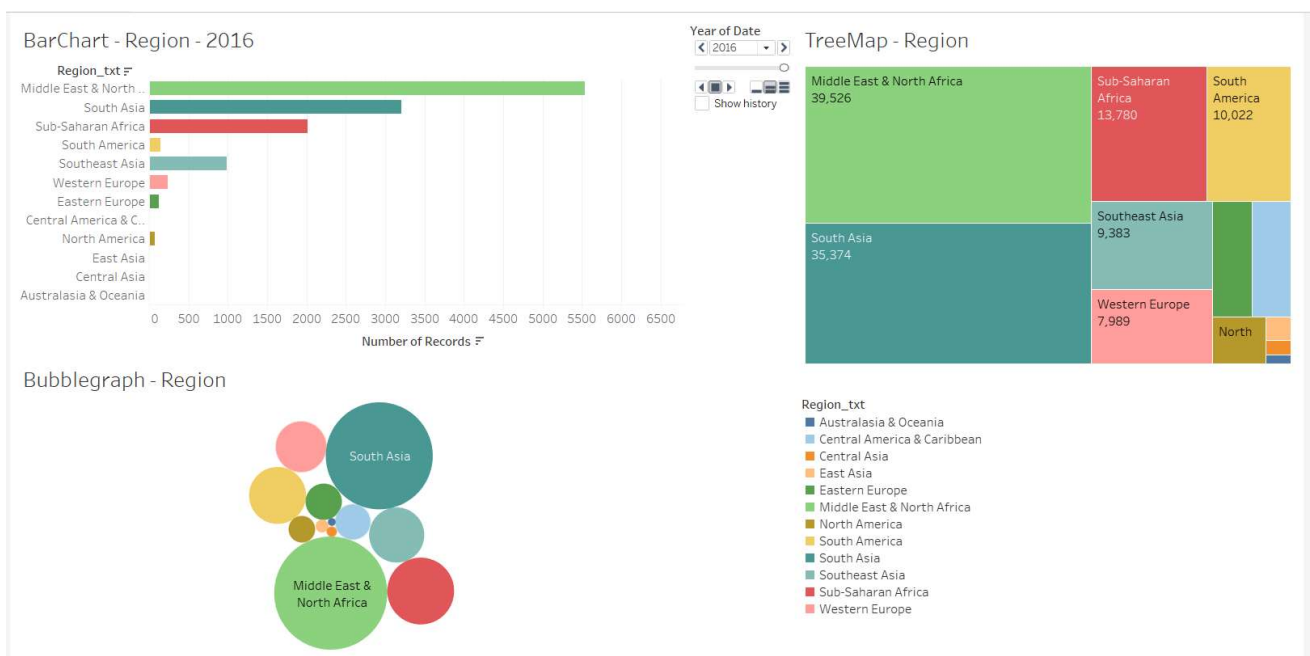


This given bar chart shows crime records within each ‘region/continents’ for overall years. For example, we can clearly interpret in the given bar chart ‘Middle East & North Africa’ holds the greatest number of records whereas ‘Australasia & Oceania’ holds the least number of records.



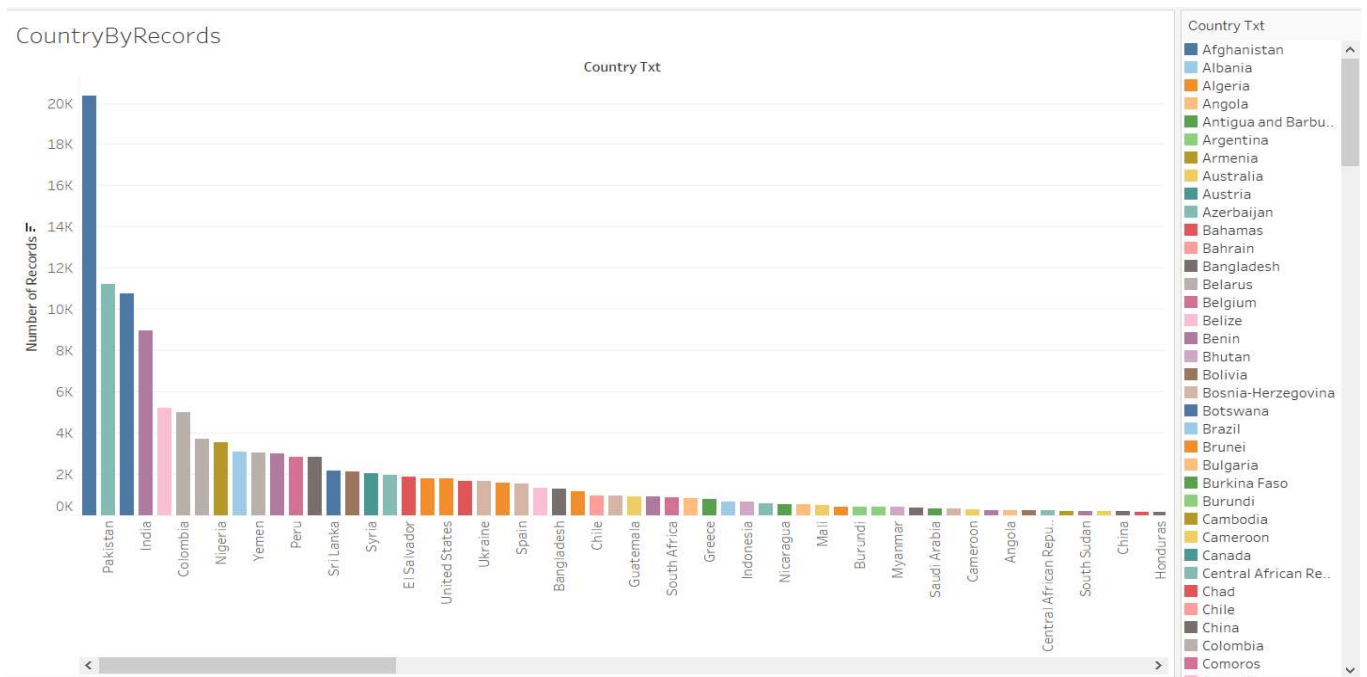
## 9. Bar Chart – Region

This second dashboard consists of crime records based on categories of ‘region /continent’. This dashboard is like previous dashboard, but we had use bubble chart instead of word cloud and rest is kept as same. In this dashboard action a filter is set on the bubble chart and bar chart except for the tree map. Now when we choose any year in the bubble chart, the corresponding value of that crime records will appear in the other charts.



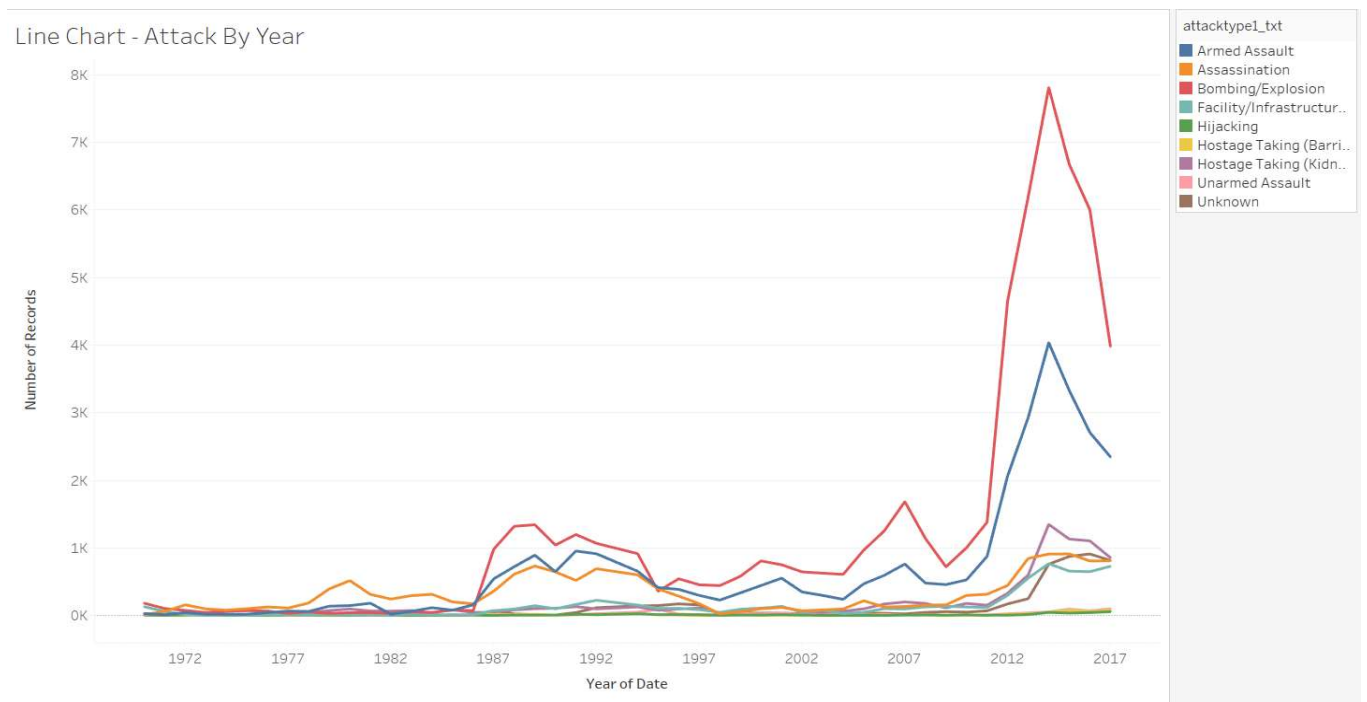
## 10. Dashboard – Region/Continent

Now, we had created a country wise bar chart. So, we get an idea which country has majorly affected by global terrorism. For example, from the below bar chart we see that 'Iran' was the country which is majorly affected by terrorism and after that 'Pakistan' is at 2<sup>nd</sup> number to hold highest terrorism.



## 11. Country by Records

An analysis made on the number of crimes over the years shows us whether the crimes have increased or decreased in number. A line graph has been used for the visual representation and it shows us the variation in crimes from 1970 to 2017. The total number of crimes was the lowest in the year 1971 around 260, there was an increased in 2007 and in 2014 there was a significant increased around 16k but in 2017 it was the lowest around 10K.



## 12. Line Chart – Attack by Year



**Conclusion:**

So, in this assignment we had studied tableau and how to use tableau. Firstly, we clean the data using R and import the cleaned data into tableau. Tableau can handle millions of rows of data with ease. Different types of visualization have been created with a large amount of data without impacting the performance of the dashboards. The dataset has been analysed in such a manner that we get to know the total number of crimes committed in each year, under each region/continent, the types of attacks committed and finally the status of these crimes that have been reported. These analysis helps us to understand how their respective governments have functioned over the years in order to create a safe environment for us.