A BLOOD AND FIRE DATA STORY

1 Eras of terrorism

We expect to see different kinds of terrorism in function of the period and localisation. We are focusing on the targets, the weapons and the types of attacks all around the world.

1.1 Type of weapons, targets and type of attacks

There is not a lot of significant results for this part. We will not continue the analysis in this direction.
However, some results are worth to explore further like the kills per attacks and the material damages informations.

1.2 Geographical context

The terrorism attacks can give us a lot of informations about the geopolitical context of a region. In this perspective, we retrieve the information with the help of the textual data that we transform into a word map per year and per region.

2 History through terrorism

Attacks mostly occur in capitals because they can reach a lot of people. Is it possible to link the importance of a city with the number of attacks they suffered from?

We tried to filter the attacks with two methods in order to find this link but with no concrete results. In the first method, we filtered the attacks with only the terrorist groups that organized attacks in more than 3 different countries. In the second, we filtered the attacks by keeping the attacks with a number of death or injured people higher than 30. Comparing the dataframes of terrorist attacks resulting from these two methods with some indices showing the importance of a country/city did not give satisfactory results.

As we do not think that the results we found are relevant, we are not going to use them in our Notebook .

3 Evaluate the impact of a terror attack

With a few thousands of terror attack each year, it seems obivous that not all of them have the same impact on our globalized word. From a few symbolic and well discussed attack such as 9/11, Bataclan's attack, Munich massacre to numerous hostage taking situation, kidnapping and else. We would like to have a better understanding of what makes a terror attack impactful. In order to do so we planned on studying media coverage (if possible) and then see if we can predict an impact coefficient based on the attack features (types of target, localization, number of killed, ...).

3.1 Number of killed and number of attacks evolution

Firstly, we decided to study the evolution of the number terrorist attacks and of their number of victims over the years. After studying this, we wanted to understand the distribution of victims over the attacks. We have found that on average 2 peoples die at each terror attacks and the maximum number of death during one event is around 1'400 peoples. This look like a power distribution with a lot of non deadly attacks and only 4 attacks with over 500 resulting deaths.

3.2 Evaluate the media coverage of an attack

Then, we would like to study how we can determine an impact coefficient for a terror attack. At first we thought of using the media coverage or to use google trends. We found that using google trend would not give us very interesting results. As it only gives us, relative search trend and not an absolute number of search that we could use, we cannot use it. The google trend dataset can also not go further back than 2004, we could also only use it for the most recent attacks then. Hence, why we decided to use the GELP project to retrieve a dataset of published articles in the week following a terror event. We are still experimenting on this, we do not have results to show yet.

3.3 Predict the impact of an attack based on a relevant selection of features

In this domain we have not done anything yet. We need the previous section to be working in order to study this.

4 Correlations with societies

Since the terrorism objective is to create a political, social or religous rupture in the society, we expected to see some correlations in these areas. Our analysis is based purely on correlation and not on causation.

We explored political aspects via the raise of far right parties, economical via the tourism growth and GDP and social via the happiness index which contains different features such as life expectancy, trust in government, trust in the future,... We tried to link these different aspects to basic informations on the attacks such as the target type and the type of attacks. This analysis is done per regions

We found slight correlations but we are not really sure about their significance. For the next weeks, we are going to make some interactive visualisations to better understand our datas but we are not going to explore further correlations part.

Conclusion

We made an extensive analysis of datasets, unfortunately we did not manage to get relevant results in all the parts. For this reason, we decided to focus our project on Parts 1 and 3 mainly.