188.195 DOPP 2019W
Exercise 3 – Executive Summary – Global Terrorism

Introduction

Over the past months our group focused on global terrorism and its causes. With extensive pre-processing and modelling we managed to answer four questions about global terrorism, its predictors as well as causes.

Questions

We adapted our initial questions and formulated four independent and scientific questions:
- Does unemployment correlate with the terrorism within a country?
- Does GDP per capita correlate with terrorism within a country?
- Which economic and social factors influence terrorism?
- What data is necessary to predict terrorism in the next year?

Data

To answer the questions above diverse data was necessary. All data was available for free from either www.kaggle.com or data.worldbank.org. In total five different datasets were used, containing information on:
- Terrorism
- Gross Domestic Product (GDP)
- GDP per capita
- GDP per capita growth in %
- Unemployment
- Life expectancy at birth

Pre-processing

During pre-processing we cleaned our data and restricted ourselves to the years 1992 to 2017 as data was mostly complete for this timeframe. Furthermore, most countries took their current shapes in 1992 as the Soviet Regime fell. Missing values were imputed with linear interpolation. Outliers in the terrorism dataset got removed, as it would be impossible for any model to predict disasters like 9/11. Due to an error the global terrorism data of the year 1993 was completely unavailable and was imputed by backwards interpolation.

After cleaning, all datasets were merged. During this process scaling was applied where necessary. In addition to known variables, we computed information on previous terrorism for every country. This data would later be included for answering the questions previously defined. The next step after generating our complete dataset was to visualize the data and answer the question to the data.

Conclusions

Does unemployment correlate with the terrorism within a country?

A correlation plot and test were used to identify dependencies. One can clearly see that there is no significant correlation between unemployment and terrorism, therefore this assumption is discarded. The correlation is also slightly negative which would be quite unintuitive to say the least.

Does GDP per capita correlate with terrorism within a country?
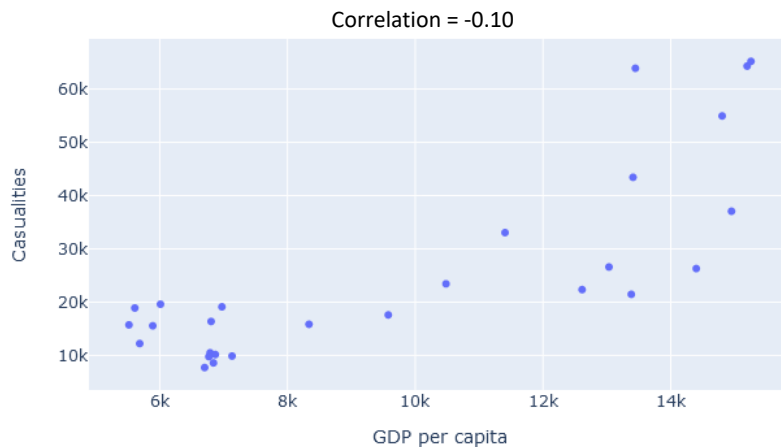
Once again there is a slight negative trend, but the correlation test was quite clear and shows that the correlation of the GDP per capita and terrorism is also non-significant.

Which economic and social factors influence terrorism?

Two different models were used to measure the influence of economic and social factors. A complex decision tree model (Random Forest) and a special regression model (Ridge Regression). The highest influence is caused by the life expectancy of the respective country, followed by the unemployment rate and the total GDP. These influences stand for both models, which validates their outcomes. This concludes that social factors influence terrorism, but so do economics. Especially since economic and social factors are often correlated with each other.

What data is necessary to predict terrorism in the next year?

In this case both models also agreed on the most significant variables to predict future terrorism. The most important information is by far the number of casualties of the last year. Followed by the GDP, which is this time slightly more influential than the social factors unemployment and life expectancy. The casualties of last year are obviously positive correlated to future casualties, as well as unemployment, while GDP and life expectancy have a negative influence of the predicted amount.