Design Progress Report 1

13 April 2018
Jacob Jasper, Amerens Jongsma, Julia Jelgerhuis, Lennert Jansen, Simon Kemmere

Topic
An analysis of mass shootings in the contiguous United States of America over the past 50 years. In light of recent developments regarding Trump's discussions with the NRA on US gun legislation following the Stoneman Douglas High School shooting, we found this dataset to be of sufficient societal relevance for this project.

Description of dataset
This dataset shows the mass shootings in the United States from the past 50 years. There were 398 mass shootings between 1966 and 2017, which resulted in 1996 deaths and 2488 injured. Based on this dataset, the average number of mass shootings in the US per year is 7, which translates into 39 deaths and 48 injured per year.

https://www.kaggle.com/zusmani/us-mass-shootings-last-50-years
398 observations
13 variables

Specification of dataset
This dataset consists of:

- Location of the shooting + type of location (open / closed)
- Age, gender, and race of the shooter
- Number of fatalities + injured
- Total victims
- Whether any policeman were killed
- Cause of the attack (religious, hate crime, terrorism, no reason)

Aim
Create a heatmap of the contiguous United States of America depicting the distribution of mass shootings and various aspects thereof.

Approach
First, we will explore the dataset further and find out what relevant data we can use. Accordingly, we will be able to decide what variables we want to use for the scope of this project and perhaps decide to introduce other data sets as well. In our analysis, we will map the US mass shootings over time and try to find possible connection between location and the aim of the shooters. If there is time, we will use other datasets to put the number of deaths through mass shootings into perspective, by visualizing data about e.g. traffic fatalities and other gun violence deaths.