

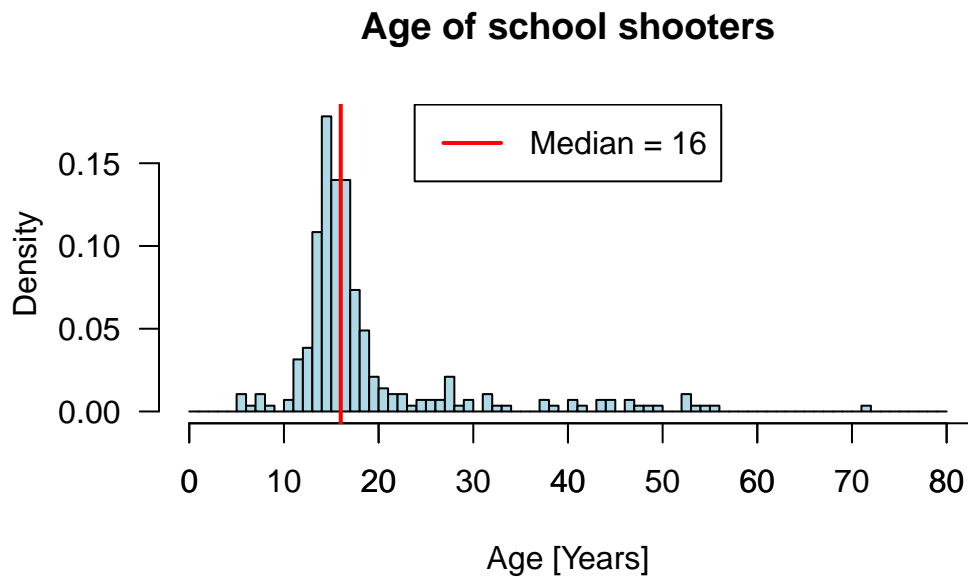
# **Semesterarbeit Explorative Datenanalyse**

Gamper Gian, Bratschi Jonas

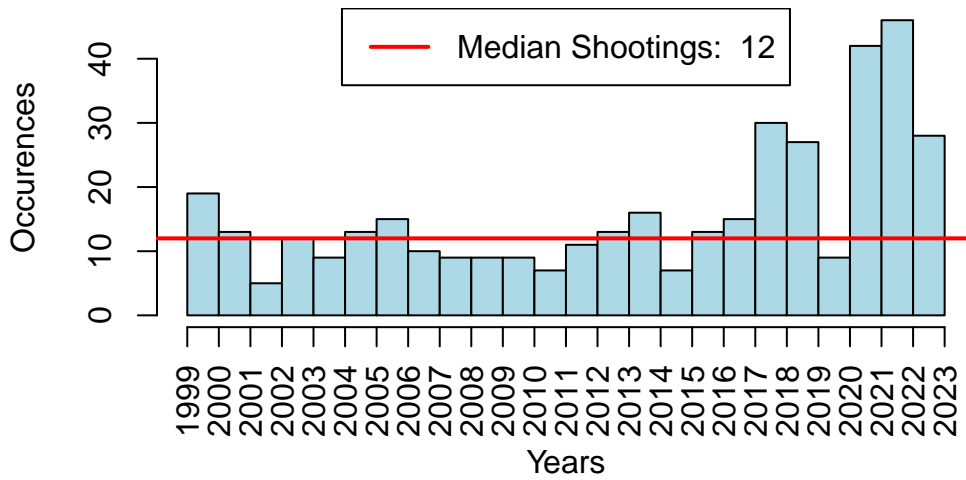
## Einleitung

In dieser Semesterarbeit analysieren wir Daten zu Schießereien der letzten drei Jahrzehnte an US-Amerikanischen Schulen. Der Datensatz umfasst, basierend auf Forschung von Pah/Amaral/Hagan und Wikipedia, eine Bandbreite von Vorfällen, von unbeabsichtigten Schüssen bis zu Massenmorden. Im Fokus steht die relative Gefährdung von Schülern durch Schusswaffen, unabhängig von den Tätermotiven. Den Datensatz stammt von der Website Kaggle (<https://www.kaggle.com/code/ecodan/us-school-shootings-dataset>).

## Tabelle mit den wichtigsten Variabeln und ersten Werten

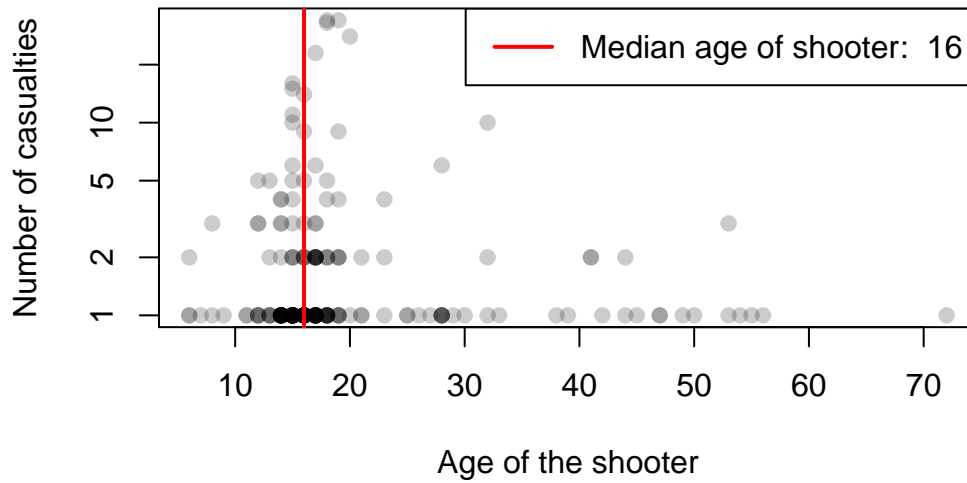


## Numbers of school shootings in the US since 1999

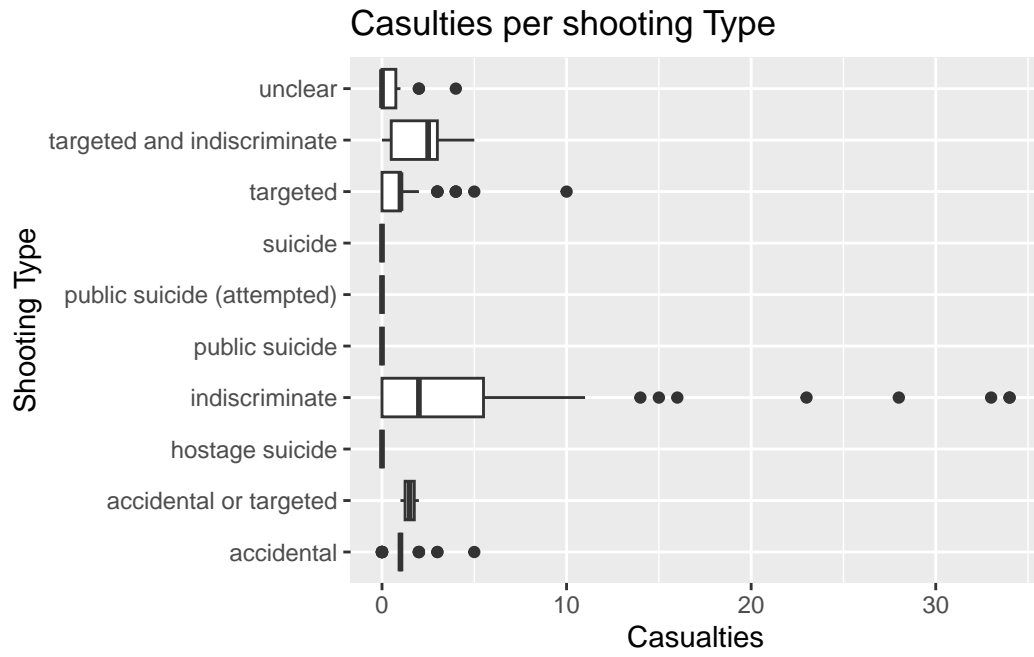


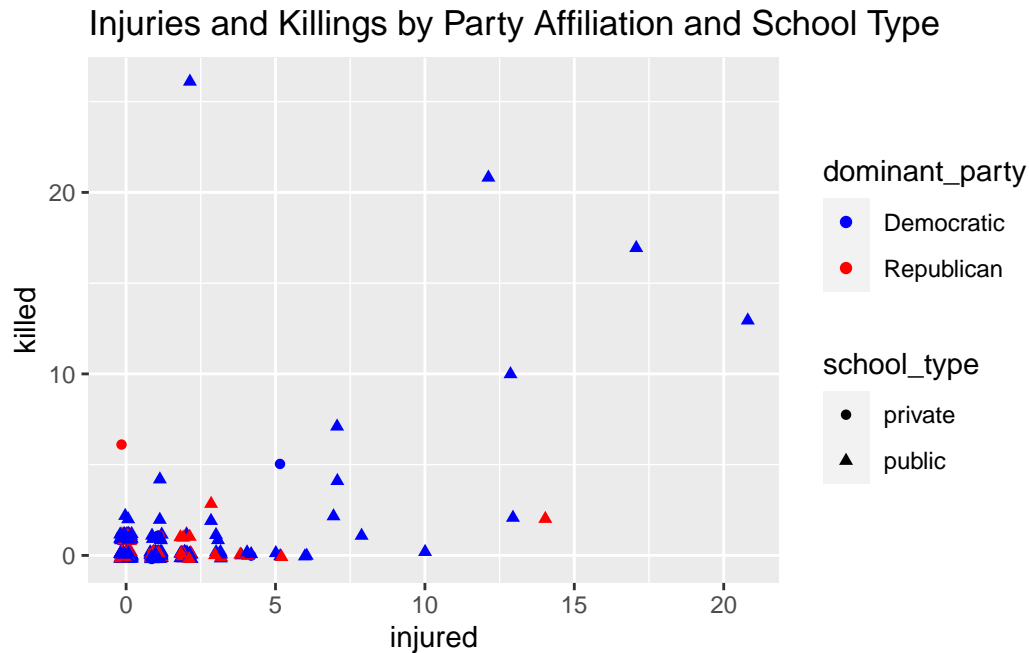
Dieses Histogramm visualisiert die Anzahl an Schiessereien an Schulen für jedes Jahr, seit 1999. Der Median seit 1999 liegt hierbei bei 12 vorfällen pro Jahr. Über die Aussage des Medians in dieser Grafik lässt sich streiten. Nichtsdestotrotz ist aber eine klare Zunahme der Fälle seit 2017 zu erkennen.

## Shooter Age / Number of casualties



In diesem Scatterplot sind auf der X-Achse die Altersangaben der Schützen und auf der Y-Achse die Anzahl der “Casualties” dargestellt, also der Toten und Verwundeten. Der Median des Schützenalters liegt bei 16 Jahren. Deutlich wird in der Grafik, dass diejenigen, die erheblichen Schaden verursacht haben, größtenteils im Altersbereich von 14 bis 20 Jahren anzutreffen sind. Dies legt nahe, dass diese Schützen möglicherweise klare Absichten hatten und entschlossen handelten, wobei wir annehmen, dass das Tatmotiv in dieser Altersgruppe wahrscheinlich auf einen Amoklauf zurückzuführen ist. Andererseits fällt auf, dass es sowohl sehr junge als auch sehr alte Schützen gibt, bei denen wir davon ausgehen, dass es sich hauptsächlich um Unfälle handelt.





In diesem Diagramm sind auf der X-Achse die Anzahl der verletzten Personen und auf der Y-Achse die Anzahl der getöteten Personen pro Vorfall dargestellt. Zusätzlich haben wir unseren Datensatz durch Informationen zur politischen Ausrichtung jedes Bundesstaates erweitert, die aus dem Wahljahr 2020 stammen. Die Symbole der Datenpunkte geben zudem Auskunft darüber, ob es sich um eine öffentliche oder private Schule handelt. Bemerkenswert ist, dass Schießereien anscheinend überwiegend an öffentlichen Schulen auftreten. Des Weiteren fällt auf, dass Schießereien mit vielen Toten und Verletzten hauptsächlich in demokratischen Staaten stattzufinden scheinen. Wir gehen davon aus, dass diese Korrelation auf eine Vielzahl von Faktoren zurückzuführen ist. Hierbei könnten die Hauptfaktoren unter anderem die Strenge der Waffengesetze, der allgemeine Zugang zu Waffen, sozioökonomische Unterschiede sowie die generelle politische Lage in den Bundesstaaten sein.

Wichtig ist aber, dass man anhand von dieser Grafik keine Kausalität ableitet. Und auch das Simpsonsche Paradoxon ist hier nicht zu vernachlässigen, könnte es zum Beispiel sein, dass in demokratischen Staaten mehr Schulen sind oder allgemein die Bevölkerung jünger ist und somit mehr Personen zur Schule gehen?

Gun Ownership per Capita vs. Killings by Party Affiliation

