

Analysis of School Shootings in the United States

Paranormal Distribution

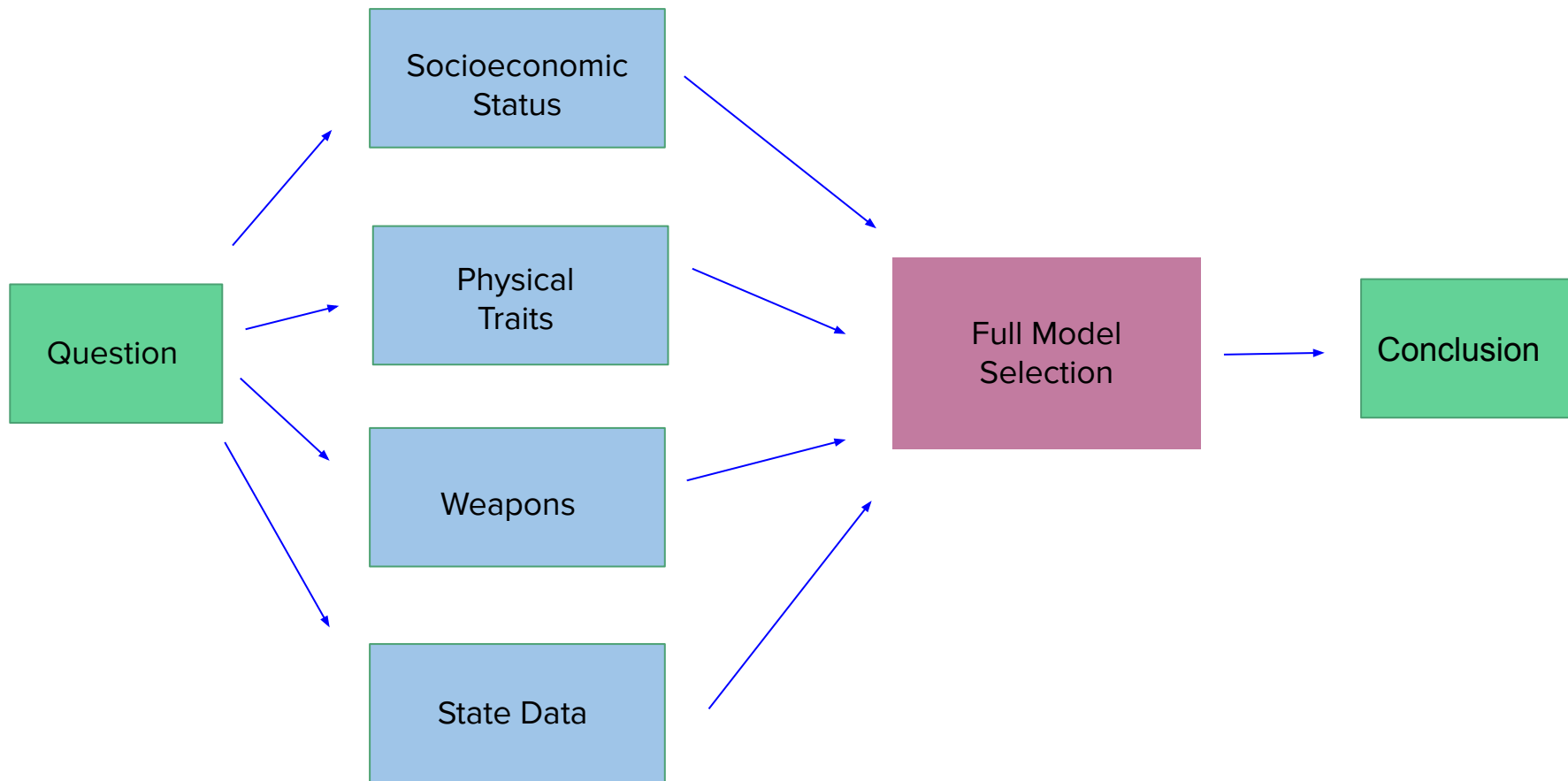
Research Question

What features of people and communities characterize high casualty school shootings?

Data Set

- Washington Post's GitHub
 - Data was collected using Nexis, news articles, open-source databases, law enforcement reports, school websites, and calls to schools and police departments
 - Observations are those that happened before, during, or just after classes
 - Each observation is a single shooting
- Gallup Dailies
 - Data was collected through sampling of the U.S. populations
 - Asked questions relating to topics like the economy, personal health, etc.
- Kalesan, B., Villarreal, M. D., Keyes, K. M., & Galea, S. (2015). Gun ownership and social gun culture. *Injury Prevention*
 - Sampled for gun ownership proportions in different U.S. states





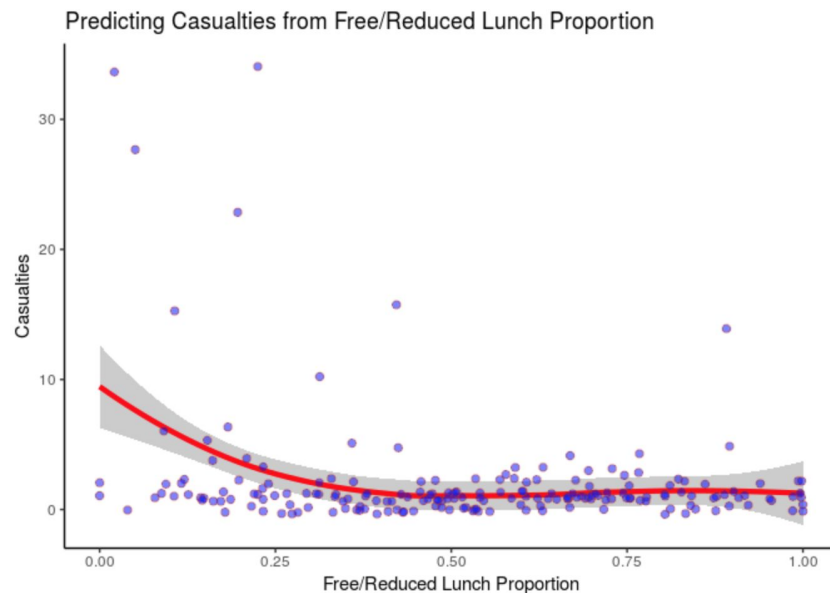
Socioeconomic Status

What tests were used?

- Made column for proportion of students in school with free/reduced lunch
- Model to predict casualties from proportion of free/reduced lunch
- Bootstrap distribution for 95% confidence interval of slope for free/reduced lunch that predicts casualties

What did this show?

There is a negative correlation between proportion of free/reduced lunch and number of casualties, as the confidence interval is entirely below 0.



Physical Traits

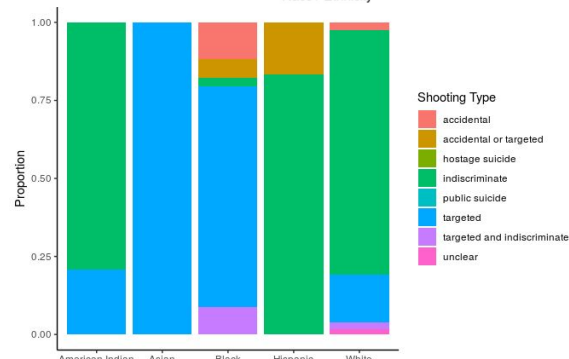
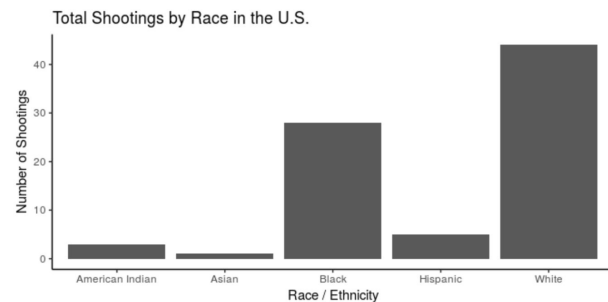
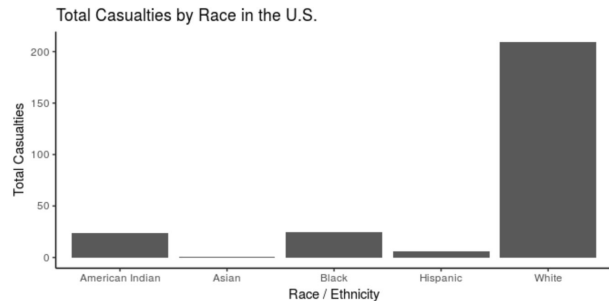
What tests were used?

Plotted total casualties by race, total shootings by race, shooting types by race, and a linear model based only on age, race, and gender.

What did this show?

- White shooters have committed a total of 44 shootings while black shooters have committed 28 shootings.
- The number of casualties resulted from black shooters is at a proportional amount of 25, meaning black shooters are most likely to commit targeted shootings.
- However, for the case of white shooters, the casualties are not proportional to the number of shootings committed with an amount of 209 casualties.
- The linear model consisting of age, race, and gender backed up this observation by showing the direct correlation each race had to the casualty count, which matched the predictions based on the charts.

term	estimate
<chr>	<dbl>
1 (Intercept)	-8.49
2 age_shooter1	0.678
3 race_ethnicity_shooter1ai	6.55
4 race_ethnicity_shooter1b	-1.95
5 race_ethnicity_shooter1h	-0.613
6 race_ethnicity_shooter1w	3.11



Physical Traits Model

casualties-hat = -8.49

+ (*shooter age*) x 0.678

+ (American Indian, True[1] or False[0]) x 6.55

+ (Black, True[1] or False[0]) x -1.95

+ (Hispanic, True[1] or False[0]) x -0.613

+ (White, True[1] or False[0]) x 3.11

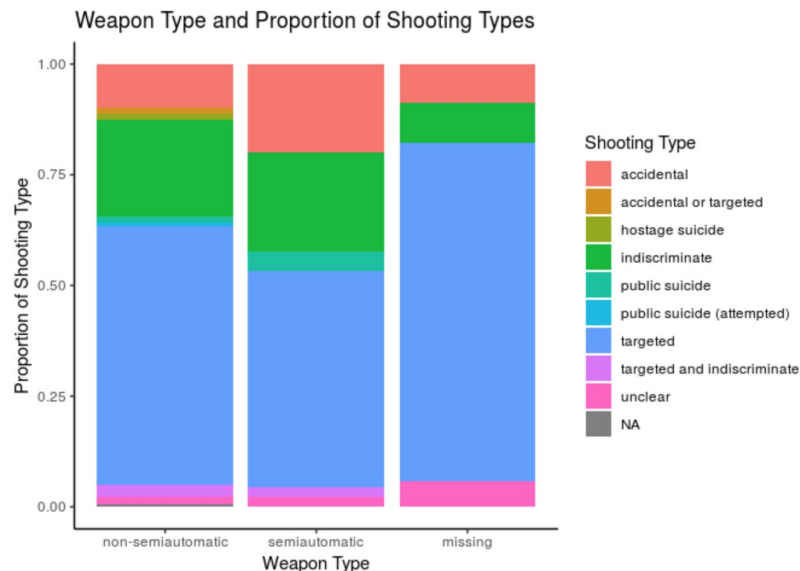
Weapons

What tests were used?

- Create weapon type and number of weapons column
- Linear model to predict casualties from these variables
- Plot proportion of weapon type used for each shooting type

What did this show?

Semi-automatic weapons have a higher proportion of accidental shootings. When the weapon is missing after the fact, there is a higher proportion of targeted shootings. With all else held constant, semi-automatic weapons contribute to a higher number of casualties. Likewise, with all else held constant, having multiple weapons significantly increases the number of casualties.



Weapons linear model

$$\begin{aligned}\text{casualties-hat} &= 1.452 \\ &+ 0.983 \times \text{weapon_typesemiautomatic} \\ &- 0.629 \times \text{weapon_typemissing} \\ &+ 6.475 \times \text{num_weaponsmultiple}\end{aligned}$$

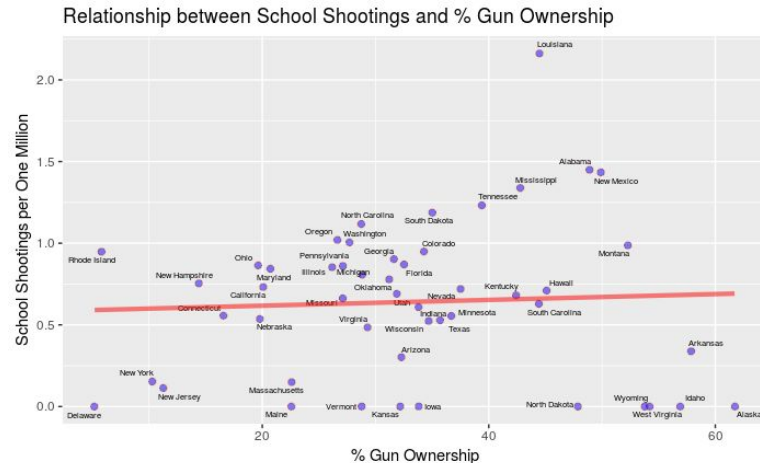
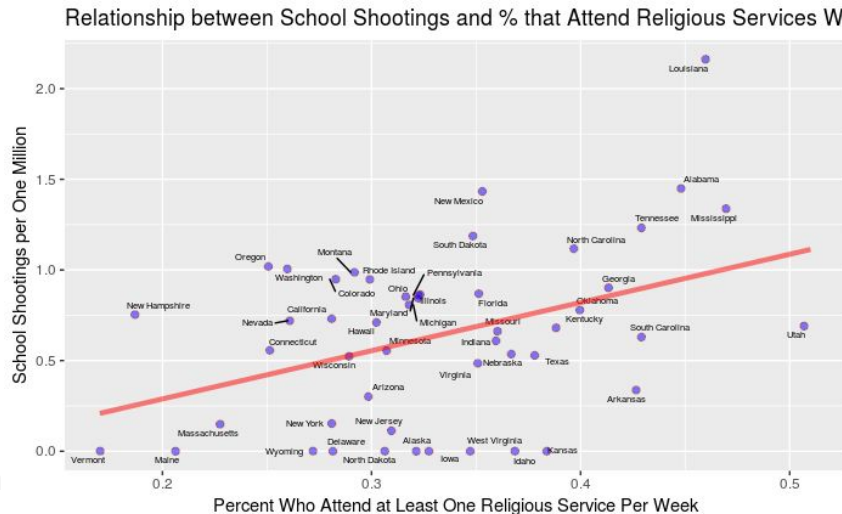
State Data

What tests were used?

- Counted the number of shootings in each state
- Factored in statewide characteristics: density, depressive rates, weekly religious attendance rates, gun ownership rates and per capita income
- Created linear model to visualize relationship between these factors and number of state shooting events per one million people
- Used backwards selection to identify relevant state characteristics in a linear model

What did this show?

Contrary to initial thought, the gun ownership levels had a negative correlation with the number of shootings. The strongest factor was the percentage of the population who attended weekly religious services.



Final State Model

$$\begin{aligned} \text{shootings-per-one-million-hat} &= 0.07 \\ &+ \text{gun ownership \%} \times -0.014 \\ &+ \% \text{ that attend religious services once per week} \times 0.03 \\ &+ \text{state density} \times -0.001 \\ &+ (\text{gun ownership \%} \times \text{density}) \times 0.00005 \end{aligned}$$

This model accounts for 26.85% of the variability in school shootings between states.

Full Model Selection

What tests were used?

AIC backward selection based on age, race, gender, proportion of students who qualify for free or reduced lunch, weapon type used, number of weapons used, gun ownership rate of state, depression rate among population, population density, religious population of state, and per capita income of state.

What did this show?

The full model showed that the most significant variables to the number of casualties a school shooter can cause were age, socio-economic status of the school, number and type of weapons used, and gun ownership rate of that state. Generally, as the age increased, the casualty count increased. As the proportion of students that qualified for free or reduced lunch increased, the casualty count dropped, and Semi-automatic weapons caused more casualties than a non semi-automatic weapon type.

	term <chr>	estimate <dbl>
1	(Intercept)	-26.0
2	age_shooter1	0.453
3	prop_lunch	-7.02
4	weapon_typenon-semiautomatic	3.36
5	weapon_typesemiautomatic	5.84
6	num_weaponsmultiple	5.34
7	ownership_rate	0.239
8	density	0.0113
9	per_capita_income	0.000277

Full Model

casualties-hat = -26

+ (*shooter age*) x 0.453

+ (% *school population eligible free/reduced lunch*) x -0.702

+ (non semi-automatic, True[1] or False[0]) x 3.36

+ (semi-automatic, True[1] or False[0]) x 5.84

+ (multiple weapons, True[1] or False[0]) x 5.34

+ (gun ownership %) x 0.239

+ (population density) x 0.011

+ (per capita income) x 0.0002

Conclusions

What features of people and communities characterize high casualty school shootings?

1. Socioeconomic Status: schools with higher proportion of students who qualify for free lunch tend to have had lower casualty rates (negative relationship)
2. State Data: when looking solely at shooting rates in states, the factors that contributed, in descending order of importance, are religious attendance % (+), gun ownership rate (-), density (+) and ownership rate * density (-)
3. Physical Traits: when looking solely at the physical characteristics of a shooter, age and race have the biggest influence over the casualty count
4. Weapons: shootings with multiple and semiautomatic weapons tended to have higher casualty rates
5. Full Model: The full model showed that the most significant factors contributing to casualties from school shooters were age, socio-economic status of the school, number and type of weapons used, and gun ownership rate within that state.



Works Cited

1. Washington Post. (2019). Data on School Shootings. Retrieved from <https://github.com/washingtonpost/data-school-shootings/blob/master/school-shootings-data.csv>
2. Gallup Analytics. (2013). Dailies survey data for 50 States. Retrieved from <https://analyticscampus.gallup.com>.
3. Kalesan, B., Villarreal, M. D., Keyes, K. M., & Galea, S. (2015). Gun ownership and social gun culture. *Injury Prevention*, 22(3), 216-220. doi:10.1136/injuryprev-2015-041586
4. United States Census Bureau. (2014). Median Household Income (In 2013 Inflation-adjusted Dollars) by State Ranked from Highest to Lowest Using 3-Year Average: 2011-2013. Retrieved from https://www2.census.gov/programs-surveys/demo/tables/p60/249/stateonline_13.xls.