# Bayes Project

*Bianca, Heather, Casey*

*3/5/2018*

**Intro**

School shootings have long been unique in America for their frequency and intensity. We have obtained a dataset on all US school shootings since 1766 through UMass GRID in order to begin research on an issue that has had very little scientific study. By examining this data-set in a statistically rigorous way, we hope to gain a better understanding of a tragic phenomenon that does not seem to be in decline.

**Data**

```
csv.data <- read.csv("data.csv")
head(colnames(csv.data),10)
```

```
##  [1] "Notes.1"
##  [2] "Date"
##  [3] "City"
##  [4] "County"
##  [5] "State"
##  [6] "School.s.Name"
##  [7] "Is.Campus.a.Gun.Free.Zone..Gun.Restricting.Zone..or.Gun.Allowing.Zone."
##  [8] "Public.or.Private.Institution"
##  [9] "School.s.Classification"
## [10] "Type.of.Education.Instituion"
```

Potential model: Let

$$y_i = \text{ length of sentence of the ith school shooting perpetrator}$$

$$r_i = \text{ race of the perpetrator}$$

$$g_i = \text{ gun used by perpetrator}$$

$$s_i = \text{ type of institution}$$

$$\alpha_j[i] = \text{ random intercept accounting for geographic region}$$

$$Y_i \sim Pois(\lambda_i)$$

$$log(\lambda_i) = \alpha_{j[i]} + \beta_1 * r_i + \beta_2 * g_i + \beta_3 * t_i + \beta_4 * s_i$$

We will also look at interactions and possible transformations of covariates as needed.

Questions:
- Would a time series component work for this model?
- Should institution type be a hierarchical structure?