Kubra Iqbal
Jimi George
Khizra Masood
Diksha Joshi
Preethi Prakash

Non-Technical Report

Data Explorers

The goal of this project is to determine if a potential male shooter can be predicted using different variables. We will analyze various variables such as race and location to see if there is a pattern to help us get a better understanding of our data.

We used gender as our dependent variable and looked at US crime rates specifically focusing on males. The reason we decided to concentrate on males was because the data showed that the males committed higher frequency of crime compared to females (There were 290 male shooters and 5 female shooters)*(table 1). The independent variables used from the data set were region, race and mental health status. With this, we could compare crime rates in different regions and how they are related to the race of the male shooter. The five team members each chose a random race and region to compare that are listed below:

Male shooters in the West that are of African American descent. (Khizra Masood)
Male shooters in the Midwest that are of Latino American descent.(Jimi George)
Male shooters in the Southwest that are of Caucasian American descent.(Kubra Iqbal)
Male shooters in the Northeast that are of Caucasian American descent. (Diksha Joshi)
Male shooters in the Southeast that are of Asian American descent. (Preethi Prakash)

With our analysis, the questions we are trying to analyze with these comparisons are

- Predict if we can use race and region to determine if a male living in a specific region and belonging to a specific race is likely to be a shooter?
- Predict if there is a correlation between mental health status and the number of victims killed by the shooter?
- Predict if a male belonging to a specific race and suffering from a mental health issue most likely to be a shooter?
- Predict if a male belonging to a specific race, region and suffering from a mental health issue most likely to be a shooter?

Equation :

Final Fitted Model :
LogP(d_males =1)/1-P(d_males = 0) = -0.2125 + 1.6431 d_white + 14.5629 d_black + 2.1609 fatalities + 1.8052 total_victims + 1.8530 injured + e

Where d_white = 1 when u_gender = white(White American or European American; otherwise = 0
Where d_black = 1 when u_gender = black(Black American or African American); otherwise = 0

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	166.693	116.268
SC	170.345	138.183
-2 Log L	164.693	104.268

R-Square	0.1910	Max-rescaled R-Square	0.4353

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	60.4242	5	<.0001
Score	52.4807	5	<.0001
Wald	24.4538	5	0.0002

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate
Intercept	1	-0.2125	0.5833	0.1327	0.7157	
d_White	1	1.6431	0.5596	8.6203	0.0033	0.4525
d_Black	1	14.5629	241.0	0.0037	0.9518	3.5108
Fatalities	1	2.1609	0.7449	8.4153	0.0037	6.7868
Total_victims	1	-1.8052	0.7593	5.6520	0.0174	-35.4547
Injured	1	1.8530	0.7746	5.7220	0.0168	32.4404

Conclusion:

According to our analysis, our final model showed that the only significant predictors were: race (white and black) and fatalities. This shows us that these predictors are the only ones that can be used to determine if a male is more likely to be

a shooter. Although the model showed these results our R^2 value was only 0.19 which shows that this model is not the best depiction for future predictions to determine if a male is likely to be a shooter or not.

Table1:

Frequency Percent Row Pct Col Pct	Table of U_Gender by Region							
	Region(Region)							
U_Gender(U_Gender)	MidAtlantic	Midwest	Northeast	Southeast	Southwest	Unknown	West	Total
Both	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 0.31 20.00 0.98	0 0.00 0.00 0.00	0 0.00 0.00 0.00	4 1.25 80.00 5.41	5 1.56
Female	0 0.00 0.00 0.00	1 0.31 20.00 1.69	0 0.00 0.00 0.00	1 0.31 20.00 0.98	0 0.00 0.00 0.00	0 0.00 0.00 0.00	3 0.94 60.00 4.05	5 1.56
Male	1 0.31 0.34 50.00	54 16.88 18.62 91.53	34 10.63 11.72 94.44	93 29.06 32.07 91.18	41 12.81 14.14 97.62	4 1.25 1.38 80.00	63 19.69 21.72 85.14	290 90.63
Unknown	1 0.31 5.00 50.00	4 1.25 20.00 6.78	2 0.63 10.00 5.56	7 2.19 35.00 6.86	1 0.31 5.00 2.38	1 0.31 5.00 20.00	4 1.25 20.00 5.41	20 6.25
Total	2 0.63	59 18.44	36 11.25	102 31.88	42 13.13	5 1.56	74 23.13	320 100.00