Factors of the Fleeing Status of Victims Among Fatal Shootings in the US

NAXA: Alayna Binder, Natalie Ashton, Xiaoran Chen, Ayse Celebi

2023-12-11

Introduction and Data

Motivation and Research Question

Over 1,200 people in the United States have been fatally shot by the police in the last 12 months, amounting to over three fatal shootings per day, according to the Washington Post's Fatal Force Database. While devastatingly tragic, this is unfortunately not an uncommon occurrence. Since 2015, there have been over 9,000 fatal police shootings in the US tracked by the Post.

The motivation behind this research stems from the pressing need to analyze and understand the dynamics of these fatal police shootings in the US. In recent years, this subject has quickly become a very prevalent issue given the need for accountability within police departments – particularly in incidents involving people from minority backgrounds. In the wake of incidents like the 2014 shooting of Michael Brown in Ferguson, Missouri, the question of racial bias in fatal police shootings has drawn significant attention (VerBruggen). In addition to race, studies have also explored the role of mental health and firearm possession in these fatal encounters. Given the importance of this issue, this project aims to conduct a comprehensive analysis of fatal police shootings in the United States. By shedding light on this complex issue, this research contributes to the ultimate goal of fostering a more equitable and just society.

Our research question is as follows: Among fatal shootings in the US over the last 8 years, is there a relationship between whether a victim tried to flee the scene and other factors related to the incident, such as the victim's age, gender, race, if they were armed, what they were doing before the incident, and whether they have a history of mental health issues?

Data Description and Cleaning

We are utilizing The Washington Post's Fatal Force Database, specifically its enhanced v2 dataset, which is updated daily by researchers at the Post. Every record requires at least

two verified sources and is checked by Post editors before publishing. The dataset contains information about the death records for each victim, such as name, gender, age, race, and state. There is also information about the incident itself, such as if the victim fled, what the victim was doing leading up to the incident, if the victim was armed, and if they had a history of mental illness. The v2 version of the dataset was introduced in 2022 with the intention of enhancing data coverage and improving the quality and consistency of information. This is the version that we are using for our analysis. We scraped this data on October 26th, 2023, so the dates of our data range from January 2nd, 2015 to October 26th, 2023.

To prepare the data for fitting models, we had to take certain steps to clean and wrangle the data. Many of our categorical variables (specifically race, what the victim was armed with, and victim flee group) had several levels, so we collapsed these into broader categories. For example, for the data describing what the victim was armed with, there were quite a few values that included multiple weapons (ex. "other;blunt_object;knife", "knife;vehicle"). We combined all of these into one categorical level called multiple. The original response variable, flee status, included the method as to which the victim tried to flee, but we are only interested in whether or not they attempted to flee for the purposes of our data analysis. Therefore, we transformed this into a binary variable with categories did not flee and fled. Lastly, we grouped the states in which the shootings occurred into regions: North, South, West, and Midwest.

Another action we took was releveling each of our potential predictors, so that the level with the largest amount of observations would be the baseline in our model. This resulted in the following baselines: male for gender, White for race, gun for what the victim was armed with, no for if the incident was related to mental health, South for region, and shoot for the threat type.

Furthermore, we considered several methods for addressing the missing data. Using the mice package, we imputed the numerical variable, victim age, in a manner that does not change its distribution. This imputation filled the missing data in through a process called predictive mean matching (PMM). PMM selects a data point from the nonmissing data that has a predicted value similar to the predicted value of the missing sample. The closest N values are considered, and ultimately one of those N values is chosen at random. Originally, we imputed the categorical variables by filling in the missing values with the most frequent value that occurred for that variable (mode imputation); however, we realized that this didn't take into account that the missing data might actually be statistically significant and affect the results that the most frequent values had on our response variable in our logistic regression, so we tried an alternative imputation method. We filled the missing values with the string "missing" to examine the effect that they have on our model overall. Finally, we had to handle the missing data in our response variable. 14% of the values for the victim flee group were missing. We decided to check to see if the missing data was random. To do this, we ran the same regression from our chosen model against a response variable of whether or not there was a value. The output of the model can be seen in appendix Figure 1. We found that this model had some significant factors, so we took this into consideration when interpreting the results.

In light of this, we still decided to drop the missing values because there was a large number of initial observations (n = 8796).

Key Variable Definitions

- `age`: age of victim (numerical)
- `gender`: gender of victim (binary)
- `race`: race of victim (categorical)
- `armed_with_grp`: if the victim was armed and, if so, with what (categorical)
- `was_mental_illness_related`: whether the victim had a history of mental illness or was experiencing a crisis at the time of the incident (binary)
- `threat_type`: actions the victim took leading up to the shooting (categorical)
- `region`: region of the US the shooting occurred in (categorical)

Exploratory Data Analysis

Lets take an exploratory look into the data and variables pertaining to our research question.

Table 1: Frequency Table for flee_group

flee_group	Frequency	Percentage
Did Not Flee	4708	61.99631
Fled	2886	38.00369

Although the majority of victims did not flee (62%), a relatively large portion of victims did (38%). The number of victims who did not flee is approximately double those who did.
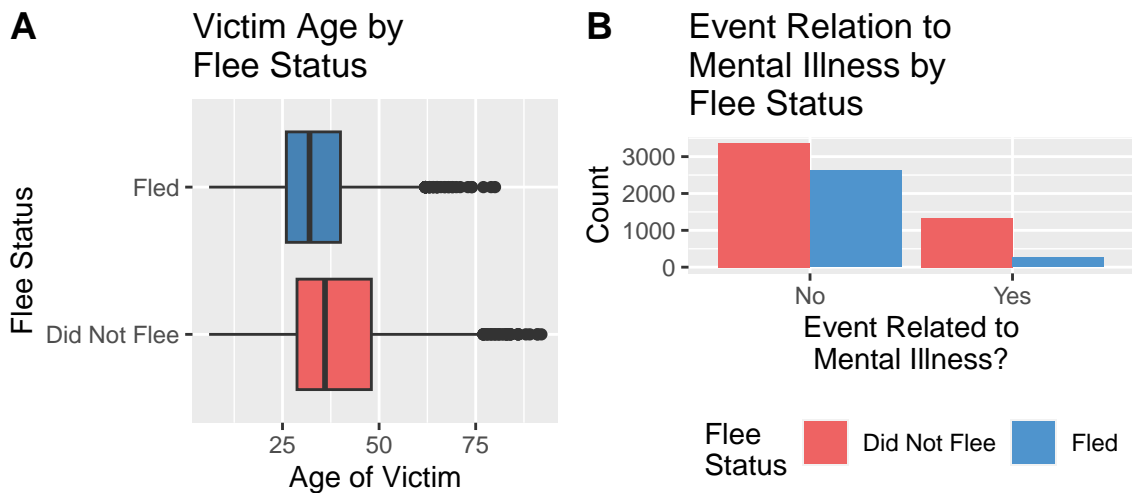
A — Victim Age by Flee Status

B — Event Relation to Mental Illness by Flee Status

Figure A The median age appears to be slightly lower for those who fled, indicating that younger victims were more likely to flee than older victims. The age range for victims that did not flee extends to be older than the age range of victims that fled. This could indicate that age alone is not sufficient to predict a victim's flee status.

Figure B More events were not related to mental illness than were. Furthermore, a higher percentage of people did not flee in events that were related to mental illness than in events that were not related to mental illness.

C   Victim Age by Threat Type
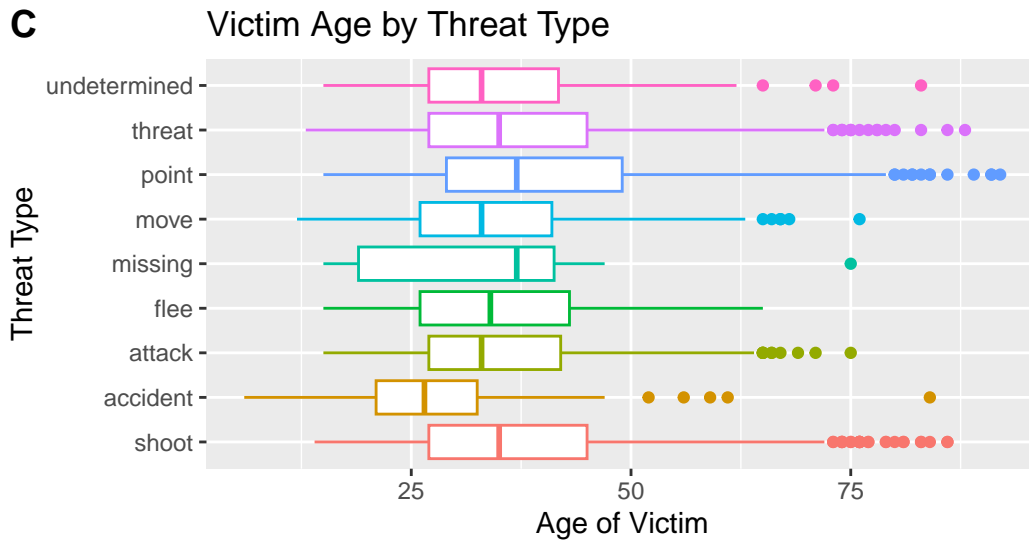
Figure C The median age differs slightly depending on the threat type, which describes the actions the victim took leading up to the fatal shooting. Victims that pointed a weapon at another individual, on average, were older than victims that made a threat of any other type. On the other hand, victims that did not make a threat (classified as an accident) were younger than other threat type groups.
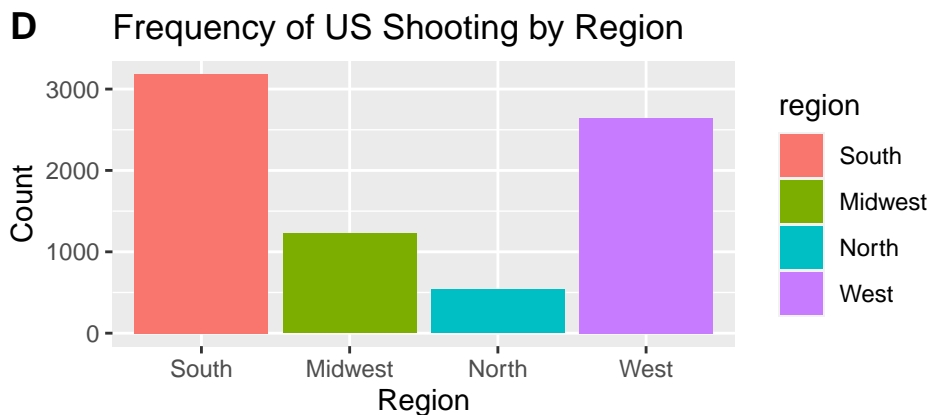
D   Frequency of US Shooting by Region

4

Figure D Out of all regions in the US, the most shootings occurred in the South, followed closely by the West. The Midwest had roughly half the number of shootings as the West, and the fewest occurred in the North.

Methodology

We want to look at whether a victim's age, gender, race, what weapon(s) they were armed with, what actions they took leading up to the fatal shooting, and if their behavior during the shooting was related to mental illness can help predict whether or not the victim attempted to flee the scene. We chose these predictors because since attempting to flee or not is a choice that each victim made, we thought that the different factors related to their police encounter, i.e. whether or not they were armed and what they were doing leading up the the encounter, as well as their characteristics, i.e. race, age, gender, relation to mental illness could all have influence on the decision they made during the fatality. To accomplish this, we wanted to consider a model with the victim flee group (did not flee or fled) as the response variable and the six aforementioned variables as predictors. We chose logistic regression because the response variable is binary. Additionally, based on the findings in our EDA, we considered two other models that included all of the predictors from the first model as well as the region the fatality occurred in for the second model and an interaction between victim age and the action the victim took leading up to the fatality for the third model.

We followed the logistic regression workflow. After splitting the data into training and testing sets, we created recipes for each of the three potential models and then conducted 10-fold cross validation on them.

After creating each of the recipes, we had to check for any instances of multicollinearity in the models. We did this by calculating the VIF for all levels of the predictor variables. If any of these are over ten, then it means that there is a correlation between that predictor and another in the model, which we don't want. We found that there was one high VIF for the flee threat type level within the third model, however we eventually selected the first model, so this does not have any impact our inference. The rest of the VIFs for all three model were under five. The extended output for each of the three models can be seen in appendix Figure 3.

We calculated the metrics associated with the cross validation (accuracy and ROC) as well as using drop in deviance tests and AIC and BIC values in order to aid us in the model selection process. The accuracy was the highest for Model 1 and the ROC is fairly similar for all three models with it being the highest for Model 2. We decided to look at more model fit statistics because these metrics were not enough for us to conclude which model is the best fit for our data.

5

Table 2: Results from Drop in Deviance Tests

Comparison	Test Statistic	p_value
Model 1 vs. Model 2	6.458	0.091
Model 1 vs. Model 3	26.524	0.001

We used two drop-in-deviance tests (output shown above) to evaluate whether the extra predictors in our second and third models did not have coefficients of 0 (i.e. are important). We compared the second and third nested models to the first model because all the predictors of the first model were in the second and third. The test between the first and second model did not have a significant p-value, according to a threshold of 0.05, so we determined that we should not add region to our model. The p-value for the test between the first and third model was very small, indicating that the interaction terms were significant, however other statistics that we will talk about later led us to the conclusion that Model 1 was a better fit for our data and the conclusion from this test can be negated by the fact that we also value parsimony and thus prefer more simplistic models.

Table 3: AIC and BIC for Potential Models

Model	AIC	BIC
Model 1	6351.929	6544.702
Model 2	6351.471	6564.186
Model 3	6351.471	6587.356

After calculating the AIC and BIC for all three models (shown above), the AIC for the Model 3 was the highest and the AIC was almost identical for Model 1 and Model 2. The BIC for Model 1, however, was lower than both the second and third models, which we expected given that both the second and third models had an additional predictor with multiple levels and BIC has a harsher penalty for additional terms. Taking all of these tests and calculations into consideration, we decided that Model 1 was the best fit for our data.

Results

After the model selection process, we decided that the first model with our initial predictor variables was the best out of the three models that we considered. Now, we need to check the model performance for our chosen model.

Table 4: Confusion Matrix

flee_group	Did Not Flee	Fled
Did Not Flee	992	182
Fled	411	314

We created a confusion matrix (shown above) to calculate the misclassification rate of our model, as well as an ROC curve with the respective AUC value (seen below) to evaluate the performance of our model. Based on the misclassification rate, 31.23% of the data was incorrectly predicted by our model, meaning that 68.77% was correctly predicted.
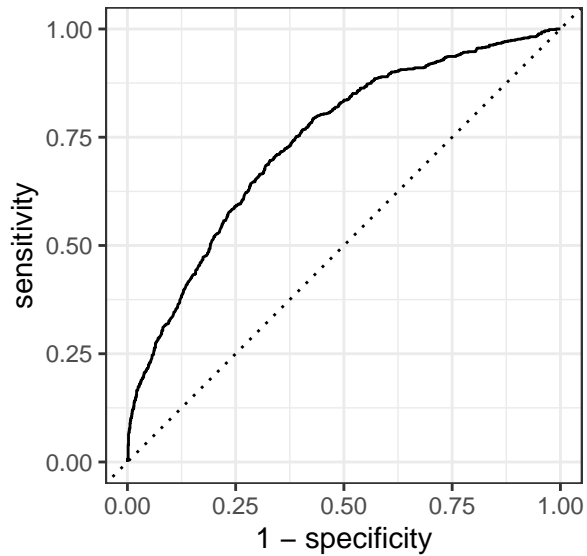
Table 5: AUC Value from Curve

.metric	.estimator	.estimate
roc_auc	binary	0.741

The ROC curve and AUC of 0.74 for our chosen model are displayed above. This is a fairly good AUC, both in general, and in relation to our other two models. This indicates that our model is efficient and performs well when distinguishing between positive and negative classes. Overall, we can conclude our model is a fairly good fit for the data; however, with more time and/or data we could make improvements, which we will expand upon in the discussion.

Finally, there are three conditions that the model needs to meet in order for us to continue with inference: independence, randomness, and linearity.

The independence condition is met because the dataset that we are using looks at individual cases of fatal police shootings in the US, so none of the observations are influenced by each other.

Based on our dataset, the observations are not random compared to the average population. This makes sense because the observations represent people who were shot in police encounters, which is not necessarily something that would happen at random to anyone. However, because we are not interested in using this model for prediction, but just to gain insight into previous patterns, this is not an issue. Nonetheless, one way that this data could be considered random is if the population that we are considering is people who have encounters with police, then we can conclude that the people who got shot were random because police should not be specifically targeting victims. There are caveats to this in modern day, but for the sake of our project we will consider that to be true.

For linearity, we need to check whether or not each level of our categorical response variable has a linear relationship with the numerical predictor variable age. To do this, we created two empirical logit plots, which can be seen in Figure 4 of the appendix. Based on both plots there seems to be a linear relationship between the two levels of the response variable and age, so the linearity condition is met.

Because all three conditions are met, we are able to use the model for inference. We fit the model, and the output is shown below.

term	estimate	std.error	statistic	p.value
(Intercept)	-0.260	0.067	-3.893	0.000
age	-0.032	0.003	-12.246	0.000
gender_female	-0.438	0.169	-2.587	0.010
gender_missing	1.127	0.721	1.563	0.118
race_Asian	-0.752	0.279	-2.695	0.007
race_Black	0.027	0.080	0.342	0.732
race_Hispanic	0.018	0.091	0.198	0.843
race_Missing	0.240	0.096	2.509	0.012
race_Native.American	0.413	0.266	1.552	0.121
race_Other	-0.044	0.587	-0.076	0.940
armed_with_grp_blunt_object	-1.246	0.261	-4.778	0.000
armed_with_grp_knife	-1.257	0.120	-10.503	0.000
armed_with_grp_missing	1.384	0.211	6.569	0.000
armed_with_grp_multiple	1.035	0.258	4.006	0.000
armed_with_grp_other	-1.009	0.350	-2.880	0.004
armed_with_grp_replica	-0.717	0.191	-3.751	0.000
armed_with_grp_unarmed	-0.064	0.157	-0.408	0.683
armed_with_grp_undetermined	0.120	0.233	0.515	0.606
armed_with_grp_unknown	-0.309	0.271	-1.137	0.255
armed_with_grp_vehicle	1.672	0.206	8.133	0.000

term	estimate	std.error	statistic	p.value
was_mental_illness_related_Yes	-1.050	0.091	-11.546	0.000
threat_type_accident	-0.376	0.406	-0.925	0.355
threat_type_attack	0.037	0.133	0.281	0.779
threat_type_flee	5.327	1.013	5.257	0.000
threat_type_missing	13.521	232.953	0.058	0.954
threat_type_move	0.316	0.159	1.981	0.048
threat_type_point	-0.211	0.088	-2.396	0.017
threat_type_threat	-0.081	0.093	-0.873	0.383
threat_type_undetermined	0.035	0.228	0.154	0.878

The model output revealed several interesting findings. The coefficient for age is statistically significant; for each additional year in age, the odds that a victim attempts to flee is expected to multiply by a factor of 0.969, holding all else constant. The coefficient for gender is also statistically significant; the odds that a female victim attempts to flee is expected to be 0.645 times the odds of a male victim, holding all else constant. The only statistically significant level for race was Asian; the odds that an Asian victim attempts to flee is expected to be 0.471 times the odds of a White victim, holding all else constant. With a victim having a gun as the baseline for the predictor variable of what the victim was armed with, almost all other levels of this predictor are statistically significant. The model suggests that individuals armed with a blunt object, knife, vehicle, multiple items, a replica weapon, or an other non-firearm weapon have different odds of fleeing compared to those armed with guns, specifically the odds of them fleeing are expected to multiply by a factor of 0.288, 0.285, 5.323, 2.815, 0.488, and 0.365 respectively, holding all else constant. The threat type level of point is also statistically significant; the odds that a victim pointing a gun at someone when the police arrived on the scene attempts to flee is expected to be 0.81 times the odds of a victim who shot someone, holding all else constant. Finally, whether or not the accident was related to mental illness is statistically significant; the odds that a victim with an event related to mental illness attempts to flee is 0.35 times the odds of a victim with an event unrelated to mental illness, holding all else constant.

Discussion

Overall, from our model, we concluded the following: The odds of a victim attempting to flee the scene are expected to increase when the victim is armed with multiple objects or has a vehicle. Both of these conclusions are compared to the baseline and holding all else constant. On the other hand, the odds of a victim attempting to flee the scene are expected to decrease if the victim is younger, female, Asian, if they were armed with a knife, blunt object, a replica weapon, or other non-firearm weapon, if the event was related to mental illness, and finally if they were pointing a weapon when the officer arrived on the scene. Again, this is

all compared to the baseline and holding all other factors constant. The specific statistics behind these conclusions are discussed in the results section above. Even though we came to these conclusions, there were many limitations of our model, so these conclusions may not be completely accurate.

The first limitation is the fact that we had to drop 14% of the response variable. As mentioned before, when checking to see if the missing data was random, we found that there were statistically significant variables. The model output revealed that some of the missing variables for applicable categorical predictors were also significant in our model and the coefficients were the same sign, which means that the model may be undercounting these specific predictors when classifying whether or not a victim will flee.

The dataset is also continually changing because observations are being added everyday. Thus, the conclusions that we drew from the model will also likely be changing, especially as time elapses. A positive of this is that over time, more and more data will be collected, which could help make the model more accurate or give us the ability to improve our model.

Finally, there are many other factors that could potentially have an effect on whether or not a victim would attempt to flee that are not included in the data, such as how many police encounters a victim had before the incident, their socioeconomic status, or even their feelings toward/level of trust in police officers (qualitative). The model that we used to draw our conclusions does not include any of these factors, since we did not have data on them, so these are likely not the only statistically significant factors for determining whether of not a victim attempted to flee the scene of a fatal police shooting.

An idea for future work could be looking to see if these conclusions are generalizable to all police encounters, including those that do not result in a fatality. This would be interesting to look at, as less than .002% of police encounters result in fatalities, which is an extremely small subset of all encounters. This may prove to be difficult because such a dataset will likely not be as readily publicly available like the Fatal Force Database put together by the Washington Post, which would require creating our own dataset. However, we believe this would yield interesting results and is worth researching further.

References

Pipis, George. "How to Impute Missing Values in R: R-Bloggers." R-Bloggers, 18 Apr. 2020, www.r-bloggers.com/2020/04/how-to-impute-missing-values-in-r/.

The Washington Post. "Every Fatal Police Shooting since 2015." *The Washington Post*, 2022, www.washingtonpost.com/graphics/investigations/police-shootings-database/.

VerBruggen, Robert. "Fatal Police Shootings and Race: A Review of the Evidence and Suggestions for Future Research." *Manhattan Institute*, 9 Mar. 2022, manhattan.institute/article/fatal-police-shootings-and-race-a-review-of-the-evidence-and-suggestions-for-future-research.

Appendix

Figure 1: Model for Checking if Response Variable is Missing at Random

term	estimate	std.error	statistic	p.value
(Intercept)	2.403	0.726	3.310	0.001
age	-0.009	0.003	-2.905	0.004
gender_male	-0.144	0.217	-0.665	0.506
gender_missing	-2.104	0.753	-2.794	0.005
gender_non.binary	10.020	196.968	0.051	0.959
race_Black	-0.277	0.383	-0.723	0.470
race_Hispanic	-0.037	0.392	-0.095	0.925
race_Missing	-1.204	0.383	-3.140	0.002
race_Native.American	-0.835	0.476	-1.753	0.080
race_Other	-0.822	0.783	-1.050	0.294
race_White	-0.445	0.377	-1.180	0.238
armed_with_grp_gun	-0.651	0.332	-1.958	0.050
armed_with_grp_knife	-0.330	0.334	-0.988	0.323
armed_with_grp_missing	0.660	0.504	1.309	0.191
armed_with_grp_multiple	0.325	0.609	0.534	0.593
armed_with_grp_other	-0.401	0.496	-0.810	0.418
armed_with_grp_replica	-0.310	0.416	-0.745	0.456
armed_with_grp_unarmed	-0.093	0.383	-0.242	0.809
armed_with_grp_undetermined	-1.153	0.388	-2.971	0.003
armed_with_grp_unknown	-1.371	0.402	-3.411	0.001
armed_with_grp_vehicle	0.097	0.437	0.222	0.824
was_mental_illness_related_Yes	0.376	0.110	3.413	0.001
threat_type_attack	0.755	0.535	1.412	0.158
threat_type_flee	3.411	1.128	3.024	0.002
threat_type_missing	-1.330	0.698	-1.905	0.057
threat_type_move	0.674	0.547	1.233	0.218
threat_type_point	1.085	0.542	2.001	0.045
threat_type_shoot	0.894	0.538	1.661	0.097
threat_type_threat	0.343	0.532	0.645	0.519
threat_type_undetermined	-0.525	0.542	-0.969	0.333

Figure 2: Cross Validation for Models 1-3

Table 8: Cross Validation Metrics from Model 1

.metric	.estimator	mean	n	std_err	.config
accuracy	binary	0.699	10	0.005	Preprocessor1_Model1
roc_auc	binary	0.754	10	0.007	Preprocessor1_Model1

Table 9: Cross Validation Metrics from Model 2

.metric	.estimator	mean	n	std_err	.config
accuracy	binary	0.697	10	0.007	Preprocessor1_Model1
roc_auc	binary	0.754	10	0.007	Preprocessor1_Model1

Table 10: Cross Validation Metrics from Model 3

.metric	.estimator	mean	n	std_err	.config
accuracy	binary	0.697	10	0.006	Preprocessor1_Model1
roc_auc	binary	0.752	10	0.006	Preprocessor1_Model1

Figure 3: Check for Multicollinearity

Table 11: VIF Levels for Model 1

	x
age	1.088
gender_female	1.060
gender_missing	1.009
race_Asian	1.026
race_Black	1.286
race_Hispanic	1.216
race_Missing	1.172
race_Native.American	1.028
race_Other	1.009
armed_with_grp_blunt_object	1.092
armed_with_grp_knife	1.507
armed_with_grp_missing	1.148
armed_with_grp_multiple	1.059
armed_with_grp_other	1.052
armed_with_grp_replica	1.048

	x
armed__with__grp__unarmed	1.599
armed__with__grp__undetermined	1.267
armed__with__grp__unknown	1.064
armed__with__grp__vehicle	1.342
was__mental__illness__related__Yes	1.027
threat__type__accident	1.186
threat__type__attack	2.238
threat__type__flee	1.012
threat__type__missing	1.000
threat__type__move	1.408
threat__type__point	1.321
threat__type__threat	1.711
threat__type__undetermined	1.346

Table 12: VIF Levels for Model 2

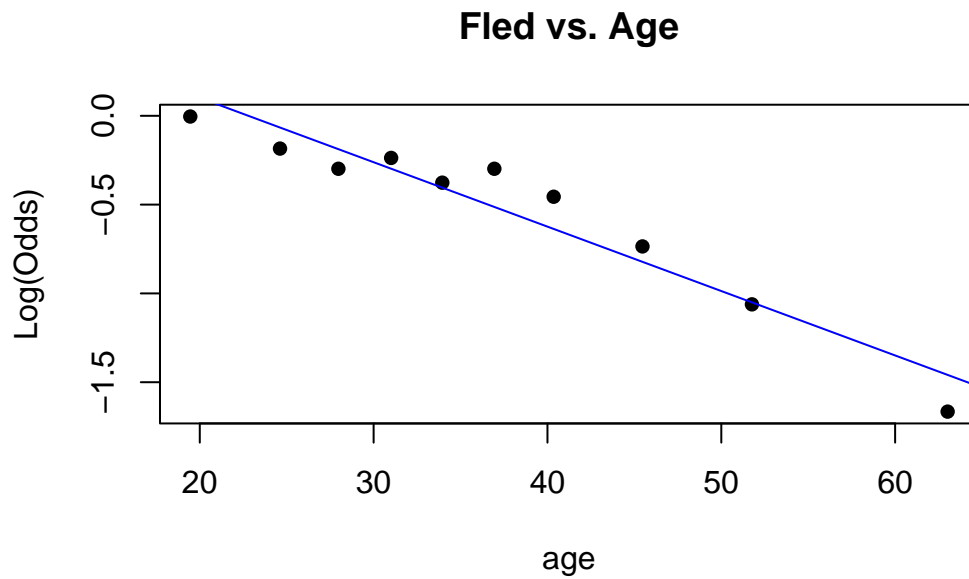|	x
age	1.090
gender__female	1.059
gender__missing	1.010
race__Asian	1.032
race__Black	1.309
race__Hispanic	1.318
race__Missing	1.197
race__Native.American	1.035
race__Other	1.009
armed__with__grp__blunt__object	1.093
armed__with__grp__knife	1.512
armed__with__grp__missing	1.150
armed__with__grp__multiple	1.060
armed__with__grp__other	1.053
armed__with__grp__replica	1.052
armed__with__grp__unarmed	1.601
armed__with__grp__undetermined	1.268
armed__with__grp__unknown	1.065
armed__with__grp__vehicle	1.343
was__mental__illness__related__Yes	1.030
threat__type__accident	1.186
threat__type__attack	2.242
threat__type__flee	1.012

	x
threat_type_missing	1.000
threat_type_move	1.412
threat_type_point	1.324
threat_type_threat	1.718
threat_type_undetermined	1.347
region_Midwest	1.178
region_North	1.091
region_West	1.353

Table 13: VIF Values for Model 3

	x
age	2.991
age_x_threat_typeaccident	1.840
age_x_threat_typeattack	1.410
age_x_threat_typeflee	1897.763
age_x_threat_typemissing	1.213
age_x_threat_typemove	1.219
age_x_threat_typepoint	1.653
age_x_threat_typethreat	1.714
age_x_threat_typeundetermined	1.090
gender_female	1.064
gender_missing	1.009
race_Asian	1.027
race_Black	1.289
race_Hispanic	1.217
race_Missing	1.172
race_Native.American	1.028
race_Other	1.010
armed_with_grp_blunt_object	1.093
armed_with_grp_knife	1.513
armed_with_grp_missing	1.150
armed_with_grp_multiple	1.062
armed_with_grp_other	1.055
armed_with_grp_replica	1.053
armed_with_grp_unarmed	1.604
armed_with_grp_undetermined	1.270
armed_with_grp_unknown	1.067
armed_with_grp_vehicle	1.346
was_mental_illness_related_Yes	1.028

	x
threat_type_accident	1.948
threat_type_attack	2.327
threat_type_flee	1897.763
threat_type_missing	1.213
threat_type_move	1.453
threat_type_point	1.325
threat_type_threat	1.741
threat_type_undetermined	1.356

Figure 4: Empirical Logit Plots

Fled vs. Age

Did Not Flee vs. Age