Exploratory Analysis of the Factors Related to Gun Mortality

Gabrielle Ilenstine
Evelyn Fox
Stephanie Mendoza

California Polytechnic State University, San Luis Obispo

Introduction

Every year, tens of thousands of Americans are killed in a gun related accident, suicide, or homicide[1]. This tragically high number of deaths, along with the idea that many of these incidents could have been prevented, has put the topic of gun and firearm reform in the spotlight of political and social debate. However, defunding of publicly funded gun-related research[2] has led to inadequate knowledge for creating an effective prevention plan. It is clear that local, state, and federal policy leaders, gun owners, and most importantly: the general public, would benefit from a better understanding of this public health crisis so that the high rate of firearm-related deaths and injuries can be reduced.

Goal

Our main goal is to inform policy-makers on factors that relate to gun mortalities with the goal of creating more efficient prevention methods. We plan on approaching this objective in a few different ways. First, we will identify the variables that are most predictive of whether or not a death is gun related. However, since policy makers would need vastly different prevention tactics for the three contrasting types of gun mortalities, accident, suicide, and homicide, we also want to delve further into each manner of death and break down which variables have the most influence on each manner of death. In order to analyze these factors thoroughly, we will use various statistical procedures such as logistic regression, random forests procedure, chi-square test for independence, and multiple graphs that present the primarily categorical data in a meaningful way. With our statistical analysis, we aim to provide foundational knowledge that can be used as a jumping off point for further investigation on the influences behind gun-related deaths.

Data

The Centers for Disease Control and Prevention (CDC) releases detailed data every year on all deaths that occur in the United States. The dataset is reported under the National Vital Statistics Systems, and includes details about the cause and manner of death along with demographic information of the deceased. The dataset that we used was for all deaths that occurred in the United States in 2014[3]. The data is commonly used for determining life expectancy, comparing mortality trends within the nation and between other countries, and analyzing characteristics of the various death occurrences.

We chose to focus on only gun-related mortality, so in most of our analyses we have reduced the 2,626,418 total registered deaths in 2014 to a subset of 20,215 deaths caused by firearms. A few factors that were not entirely relevant to gun-related deaths were eliminated, while variables such as age, race, marital status, education, and manner of death were retained, as they were more pertinent to the specificity and focus of our analysis.

Data Cleaning

Recode Adjustments:

Variables such as education, race, and age contained inconsistencies in classifications, so the multiple re-codes of these variables were sorted through and grouped into appropriate categories that allowed for all observations to be included and consistent. For example, in the original data set, some observations contained education reported in terms of years, while others were reported in terms of stages (i.e. “3 years of college” or “4 years of college” vs. “Bachelor’s degree”). We created broader categories for education, such as “some college and completed college,” that would allow us to blend the observations that had education in terms of years with those in terms of stages without making potentially incorrect assumptions. Our updated classifications of recoded variables can be viewed in Table 1 along with the code descriptions of the other indicator variables that were used.

Because age of the decedent is the only variable in the data set with the ability to remain quantitative, we chose not to group the observations into age groups but rather leave the ages in terms of years. For infant deaths where age is recorded in terms of either minutes, hours, days, or months (all of which were calculated to be less than a year), the age was recorded as 0, meaning they didn’t live to be 1 year old.

Subsetting To Gun-Related Deaths:
The process of narrowing down almost 20,000 unique causes of death to only those that were gun-related was executed using the INDEX function and searching for the keywords “gun”, “firearm”, or “rifle”. This step reduced the data set to 20,215 observations among 10 unique gun-related reasons for death.

Analysis

I. Logistic Regression exploring whether or not deaths are gun related
In our first analysis, we want to explore what factors are most important in determining whether or not a death is gun-related. Since the desired response is binary and categorical, logistic regression is an appropriate model.[4] We ran a backwards elimination logistic regression reflecting the following formula:

$$\text{logit}[P(Y=1)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_{25} + \beta_{26} x_1 x_2 + \ldots$$

In this equation, Y is 1 when a death is gun-related and 0 if it is not. We used twenty-four categorical and one quantitative predictor variables, and various relevant interaction terms. We did not include three-way interaction terms because our main goal is to interpret our results to a non-technical audience of policymakers. We set the significance level to 0.01 to narrow down the many different variables included in the model. Although this logistic regression could serve as a predictive model, our primary intent is to identify variables that have the greatest influence on gun mortality. Therefore, we will focus on the significant variables included in the final model. As shown in Table 2, almost all the variables included in the logistic regression are significant (at the 0.01 level) in predicting whether a death is gun-related. Therefore, we will use these variables in our additional analysis investigating the individual manners of death.

II. Variable selection analysis within each type of gun related of death
Our next analysis aims to find variables that are best at predicting each manner of gun related death: accident, suicide, and homicide. We accomplished this by creating three separate random forests, one for each manner of death, using the HPFOREST procedure[6]. We decided to use the random forest model because it naturally handles categorical predictor variables, contains no formal distribution assumptions, and its automated variable selection will ultimately help us determine variable importance[7]. The random forest method uses training data to create many decision trees to determine variable importance and create a predictive model for a target variable, in this case, manner of death. A visual representation of a single tree in a forest is shown in Figure 1. To improve the generalization error of the model, we decreased the proportion of training observations to form a tree to .3[6]. We also changed the alpha level to .1 to account for the large number of observations in the training data. We included all variables that were significant in the logistic regression model into each random forest.

Random forest for accident-related gun deaths
The output for variable importance in the accident-related gun deaths random forest is shown in Table 3. Since the random forest model generated by accident-related gun deaths yielded a smaller baseline misclassification rate (0.017) than the out-of-bag misclassification rate, which is about 0.08 for a large number of trees, we conclude that the model may not be a good fit[6]. Therefore, we will not suggest that the predictive model be used, and we will not focus further analysis based on the random forest’s output of variable importance.

Random forests for suicide-related gun deaths
The misclassification rate for the suicide related gun deaths random forest (0.418) is slightly larger than the out-of-bag misclassification rate (0.395 for a large number of trees), revealing that the model has weak predictive ability. Therefore, we will again use this random forest as a variable reduction technique instead of a predictive model, where variables of greatest importance in creating the random forest Table 4 will be focused on in further analysis. Race, marital status, sex, age, and education have the largest out-of-bag Gini measures, so we will explore these predictors more thoroughly in determining whether a gun related death will be a suicide.

Random forest for homicide-related gun deaths
The random forest created for homicide related gun deaths is similar to the forest created for suicide related gun deaths. Since the misclassification rate (0.435) is so similar to the out-of-bag misclassification rate (0.401 for a large number of trees), we are not able to use the random forest as a predictive model due to low generalization ability. However, we will continue to explore the variables with the highest out-of-bag Gini measure to perform further analysis (Table 5). The variables of highest importance are identical to the results of the random forest for suicide related gun deaths: race, marital status, sex, age, and education.

III. Looking into significant variables for gun-administered suicide

Race
From our findings in the random forests procedure, we found that race was the most significant variable in determining whether a gun-related death was a suicide. We looked deeper into the trends of gun-administered suicides within race compared to all types of suicides within race. The table of these counts and proportions are shown in Table 6. A chi-squared test was used to determine if the proportions of suicides that were gun-administered were significantly different among the races.

H_0: Whether one's suicide is gun-related or not is independent of race.
H_a: Whether one's suicide is gun-related or not is dependent on race.

The test result table of the chi-squared test can be seen in Table 7. A p-value of <.0001 indicates that the test was significant (<.05), so at least one race had a significantly different proportion of gun-administered suicides from the others. We are able to conclude that there is a relationship between race and whether or not one's suicide is gun-related or not; these variables are not independent. By referencing Figure 2 in the appendices, one can visualize the differences in these proportions. With an overall percentage of 19.58% of suicides that were gun-administered, it's interesting to note that 21% of Caucasian (Non-Hispanic) and 21.5% of American Indian suicides were committed with a gun, while the rates of the remaining races are substantially lower and range from about 9-12%.

While it's important to recognize the pattern of gun-related suicides within races, it's worth noting that 83% of all suicides were of the Non-Hispanic Caucasian race when the Non-Hispanic Caucasian population proportion in the U.S. is 62.2%. This puts some perspective on the larger issue of suicide itself in our country and how some races more frequently commit suicide than others.

Sex and Marital Status
The next two most significant variables in determining if someone committed suicide with a gun are sex and marital status. We performed a relative risk test and found that in 2014, men have 7 times the risk of having a gun related suicide compared to females. We looked at how these variables interacted together and found that a higher proportion of both married men and married women committed suicide with a gun compared to single, divorced, and widowed people of both sexes. As shown in Figure 3, 39% of men who committed suicide with a gun were married, and 40% of women who committed suicide with a gun were also married.

Age and Education
The random forests model shows age and education to be significant predictors of if one's gun-related death was a suicide or not. The mean and median age for people who killed themselves with a gun are 50 and 51 years old, respectively; these statistics are the highest among all manners of death that are gun-related. Figure 4 displays a spread of age frequencies by manner of death in a histogram. The proportions of gun-related deaths that were from suicide vary between different levels of education. Decedents with some college or a college degree had the highest percentage of gun-related deaths that were from suicide: 60.1%. The rest of these proportions are presented graphically in Figure 5.

IV. Looking into significant variables for gun-related homicide

Race

For determining whether a person was involved in a gun-related homicide, the random forests procedure also showed race as being the most significant determinant. We conducted another chi-squared test to determine if the proportion of homicides that are gun-related is the same across all races. These proportions are displayed in Table 8.

H_0: Whether one's homicide is gun-related or not is independent of race.
H_a: Whether one's homicide is gun-related or not is dependent on race.

The test results can be viewed in Table 9. A p-value of <.0001 indicates that the test was significant (<.05), so similarly to the chi-square test done with suicides, we can conclude that whether one's homicide was gun-related or not is dependent on their race. These differences are presented visually in Figure 6, where the proportion of homicides that were gun-related are displayed by race. African Americans have the highest portion of homicides that are gun-related (78.82%), which is substantially higher than the overall rate, where 67.83% of homicides across all races were done with a firearm.

Sex and Marital Status

Similar to gun-administered suicide, sex and marital status are the next two most significant variables in predicting gun-related homicide. A relative risk test showed men have 5 times the risk of having a gun related homicide compared to females. We looked at the interaction between sex and marital status, which can be seen visually in Figure 7. In gun-administered suicides there was a higher proportion of people who used a gun to commit suicide if they were married; however, the opposite trend occurs with gun-related homicides. Of the men who died in gun-related homicides, 74% of them were single. Similarly, of the women who died in gun-related homicides, 49% of them were single.

Age and Education

With age as a significant variable of if one's gun-related death was a homicide or not, it is interesting to note that the mean age for people whose gun-related death was a homicide is 32.7 years old, while the median is 29.0. Figure 4 displays a spread of age frequencies by manner of death in a histogram. Education level has an influence on if one's gun-related death was a homicide or not. Decedents with their highest education as high school with no diploma had the highest percentage of gun-related deaths that were from homicide: 77.1%. The proportions for every education level are presented graphically in Figure 5, where the homicide rates can be compared to the accident and suicide rates.

Conclusion

The purpose of this analysis was to explore the variables and demographics that influence gun-mortality trends for the use of reducing gun fatalities. Most of our analysis focuses within gun-administered suicides and gun-related homicides separately because policy makers need to tailor prevention tactics differently for both of these manners of death.

For gun-administered suicides, non-Hispanic Caucasians and American Indians had a higher proportion of gun-administered suicides compared to those of other races. We discovered that males are disproportionately affected, with men having seven times the risk of having a gun related suicide compared to females. Also, married individuals had a higher proportion of gun-administered suicides. We also noticed a trend of older aged people committing gun-administered suicides, with a mean age of fifty years. Higher percentages of gun-administered suicides occurred from individuals with some college education or a college degree. Therefore, policy-makers should cater gun-administered suicide prevention tactics to non-Hispanic Caucasian and American Indian males who are married, have at least some college education, and around the age of fifty, as this demographic is at the highest risk for ending their life with a firearm. It is important to note that if a policy-maker is hoping to develop prevention strategies as effectively as possible, their efforts would be most effective if they focused on demographics in this order: race, sex, marital status, age, education level.

Gun-related homicides produced unique results of which demographics are most affected. The race with the highest proportion of people killed using a firearm was African Americans. Similar to gun-administered suicides, men have five times the risk of having a gun related homicide compared to females. In contrast to the trend seen in gun-administered suicides, individuals whose marital status is "single" had a larger proportion of gun-related homicides, compared to those who were "married", "divorced", or "widowed". Those who were younger in age also had a tendency to die from a gun-related homicide, with a mean age of around thirty-three. As education level increases, the proportion of gun-related homicides decreases. Therefore, prevention strategies would be most effective if directed towards African-American males who are single, young-adults, and have lower education levels. The influence of each of these variables is similar to the variables described in gun-administered suicides.

Up-to-date information about groups of Americans who are affected the most by gun mortality is essential to reducing this public health crisis. There are various ways to change these differential outcomes among different demographics, such as gun regulations on reservations, mental-health checks for gun-owners, and reformed gun laws. However, we will not attempt to speculate as to why these specific patterns are occurring. We hope to inspire other researchers to carry out further investigation on this topic, as it is an extremely intricate issue of which we have only grazed the surface.

Suggestions for Future Studies

The lack of geographical information in the data restricts us from understanding contextual details about where the shooting occurred and from comparing gun deaths among regions in the United States. Data on the state, city, and setting where the death occurred (rural, urban, suburban, etc.) would permit us to make connections between gun laws, current events, policing, and health resources that are necessary in order to fully understand *why* these events occur and what specific governmental influences need to be adjusted in order to reduce firearm mortality. For example, *Slate Magazine* provides up-to-date gun injuries and fatalities on an interactive map of the U.S. that could be integrated with our analysis to provide insights in correlation between gun violence and geographic location[9].

While we understand that many details of the individual are to be kept private, information about the deceased such as mental health, gun registration/access, and family history could possibly be crucial factors that correlate with manner of death. Demographic variables such as sex, race, marital status, and education are helpful in analyzing gun-related fatalities, however, more individual characteristics need to be examined in order to account for diversity within these broad social categories.

Because our data refers to strictly gun-related mortality, for the purpose of deeper analysis on the relationship between certain demographics and firearms, it would be useful to also look at gun-related injuries. Combining injuries with fatalities would likely allow more insight into gun-related accidents, since our study only covers the portion of gun accidents that end in death. Accidental shootings are a huge issue in the topic of gun-policy and are extensive enough to warrant an entire study of their own.

The conclusions made in this study only hold true for recent gun-related deaths since the data only covers deaths from 2014; however, the Center for Disease Control and Prevention releases data on deaths in the United States for every year, so further analysis could compare gun-related trends over time.

References

[1] "AMA Calls Gun Violence 'A Public Health Crisis'"." *American Medical Association*. American Medical Association, 14 June 2016. Web. 07 Jan. 2017.

[2] Masters, Kate. "141 Medical Groups Urge Congress to Restore Funding for Gun Violence Research." *The Trace*. N.p., 06 Apr. 2016. Web. 06 Jan. 2017.

[3] Center of Disease Control and Prevention. (2014) Death in the United States [Data file and code book]. Retrieved from https://www.kaggle.com/cdc/mortality

[4] Agresti, Alan. *An Introduction to Categorical Data Analysis*. Second ed. New York: Wiley, 1996. Print.

[5] "The LOGISTIC Procedure." *Documentation.sas.com*. SAS® Institute Inc., n.d. Web. 05 Jan. 2017.

[6] "The HPFOREST Procedure." *Documentation.sas.com*. SAS Institute Inc., n.d. Web. 05 Jan. 2017.

[7] Cutler, Adele, Dr. "Introduction to Random Forests with Microarray Data." *YouTube*. YouTube, 28 Jan. 2013. Web. 03 Jan. 2017.

[8] Rose, Jen, and Lisa Dierker. "Building a Random Forest with SAS." *Coursera*. N.p., n.d. Web. 02 Jan. 2017.

[9] Kirk, Chris, and Alex Yablon. "How Many People Have Been Shot in Your Neighborhood This Year?" *Slate Magazine*. N.p., 31 Dec. 2015. Web. 03 Jan. 2017.

Appendix

Table 1: Variables and Indicator Variable Explanations

Variable	Meaning	Indicator Explanations
AccidentInd	Indicates whether the death was an accident or not	0 - Not an accident 1 - Accident
Age	Age of the deceased in years	
AmericanIndianInd	Indicates whether race of deceased was American Indian or not	0 - race of deceased was not American Indian 1 - race of deceased was American Indian
AsianInd	Indicates whether race of deceased was Asian or not	0 - race of deceased was not Asian 1 - race of deceased was Asian
BlackInd	Indicates whether race of deceased was African American or not	0 - race of deceased was not African American 1 - race of deceased was African American
DayOfWeekOfDeath	Day of week of which the death occurred	1 - Sunday 2 - Monday . . 7 - Saturday
Description	Description gives the specific cause of death	
Education	Highest education level of the deceased	1 - 8th grade or less 2 - 9th-12th grade, no diploma 3 - High school graduate or GED completed 4 - Some college credit or college degree
Education1Ind	Indicates whether or not the highest level of education received was 8th grade or less	0 - highest level of education received was not 8th grade or less 1 - highest level of education received was 8th grade or less
Education2Ind	Indicates whether or not the highest level of education received was 9th-12th grade, no diploma	0 - highest level of education received was not 9th-12th grade, no diploma 1 - highest level of education received was 9th-12th grade, no diploma

Education3Ind	Indicates whether or not the highest level of education received was high school (graduate) or GED completed	0 - highest level of education received was not high school (graduate) or GED completed 1 - highest level of education received was high school (graduate) or GED completed
Education4Ind	Indicates whether or not the highest level of education received was some college credit or college degree	0 - highest level of education received was not some college credit of college degree 1 - highest level of education received was some college credit of college degree
Gun	Indicates whether the death was gun-related or not	Z - Not gun-related Y - Gun-related
HispanicInd	Indicates whether race of deceased was Hispanic or not	0 - race of deceased was not Hispanic 1 - race of deceased was Hispanic
Home	Indicates whether the death was recorded in the home or not	0 - death not recorded in home 1 - death recorded in home
HomicideInd	Indicates whether the death was a homicide or not	0 - Not a homicide 1 - Homicide
HospitalDOA	Indicates whether the recorded location of the death was in a hospital, with the deceased dead on arrival	0 - death not recorded in hospital, individual dead on arrival 1 - death was recorded in hospital, individual dead on arrival
HospitalInd	Indicates whether the recorded location of the death was in a hospital (inpatient or outpatient) or emergency room or not	0 - location of death was not recorded as hospital inpatient, hospital outpatient, or emergency room 1 - location of death was recorded as one of the above
Icd10Code	Icd10Code gives an arbitrary code to the description	
Id	Identification number of the deceased	
InjuryAtWork InjuryAtWorkInd	Whether or not the injury that caused the death occurred at the deceased's workplace	Y - Yes N - No 0 - No 1 - Yes

MannerOfDeath	Manner in which the death occurred	1 - Accident 2 - Suicide 3 - Homicide 4 - Pending Investigation 5 - Could not determine 6 - Self-inflicted 7 - Natural *Note, our study focuses on 1, 2, and 3
MaritalStatus MaritalStatusInd	Marital status of the deceased	S - Single, never married M - Married W - Widowed D - Divorced 0 - Single 1 - Married 2 - Widowed 3 - Divorced
MaritalStatusDInd	Whether the deceased was divorced or not	0 - not divorced 1 - divorced
MaritalStatusMInd	Whether the deceased was married or not	0 - not married 1 - married
MaritalStatusSInd	Whether the deceased was single or not	0 - not single 1 - single
MaritalStatusWInd	Whether the deceased was widowed or not	0 - not widowed 1 - widowed
MonthOfDeath	Month of which the death occurred	1 - January 2 - February . . 12 - December
Nursing	Indicates whether death was recorded in hospice, nursing facility, or long-term care facility or not	0 - death not recorded in hospice, nursing facility, or long-term care facility 1 - death recorded in hospice, nursing facility, or long-term care facility
OtherAsianInd	Indicates whether race of deceased was Pacific Islander or other Asian or not	0 - race of deceased was not Pacific Islander or other Asian 1 - race of deceased was Pacific Islander or other Asian

OtherPlaceOfDeath	Indicates whether death was recorded in place categorized as "other" place or not	0 - death not recorded in place specified as "other" 1 - death recorded in place specified as "other"
OtherRaceInd	Indicates whether race of deceased was specified as "other" or not	0 - race of deceased was not "other" 1 - race of deceased was "other"
PlaceOfDeathAndDecedentsStatus	Location of which the person became recorded as deceased	1 - Hospital, inpatient, outpatient, or emergency room 3 - Hospital, but dead on arrival 4 - Home 5 - Hospice, nursing facility, or long-term care facility 7 - Other
PlaceOfInjury	Location of which the injury occurred that caused the person to become deceased	0 - Home 1 - Residential Institution 2 - School, other institution, or public administrative area 3 Sports and athletics area 4 - Street and highway 5 - Trade and service area 6 - Industrial and construction area 7 - Farm 8 - Other specified place
Race	Race of the deceased	0 - Other races 1 - Non-Hispanic white 2 - African American 3 - American Indian 4 - Asian 8 - Other Asian and Pacific Islander 9 - Hispanic
Sex SexInd	Sex of the deceased	M - Male F - Female 0 - Male 1 - Female
SuicideInd	Indicates whether the death was a suicide or not	0 - Not a suicide 1 - Suicide
WhiteInd	Indicates whether race of deceased was Caucasian (Non-Hispanic)	0 - race of deceased was not Caucasian (Non-Hispanic) 1 - race of deceased was Caucasian (Non-Hispanic)

Table 2: Results of Logistic Regression

Summary of Backward Elimination					
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq
1	Home	1	27	1.9292	0.1648
1	Age*Nursing	1	26	2.7207	0.0991
1	AmericanIndianInd	1	25	3.1378	0.0765
1	HospitalDOA	1	24	3.5935	0.0580
1	Age*MaritalStatusMIn	1	23	5.0998	0.0239
1	Age*MaritalStatusDIn	1	22	4.5106	0.0337
1	MaritalStatusDInd	1	21	6.0479	0.0139

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	7.7696	0.0907	7343.2302	<.0001
Age	1	0.000948	0.00109	0.7600	0.3833
MaritalStatusSInd	1	-0.6797	0.0652	108.7382	<.0001
MaritalStatusMInd	1	-0.2270	0.0269	71.4454	<.0001
SexInd	1	0.6659	0.0622	114.5744	<.0001
HospitalInd	1	0.4662	0.0255	333.6031	<.0001
Nursing	1	2.6552	0.1674	251.5750	<.0001
AccidentInd	1	-2.0717	0.0719	829.9787	<.0001
SuicideInd	1	-6.3139	0.0502	15788.0357	<.0001
HomicideInd	1	-8.5971	0.0530	26350.6845	<.0001
WhiteInd	1	-0.2187	0.0334	42.9122	<.0001
BlackInd	1	-0.5647	0.0373	228.9593	<.0001
AsianInd	1	0.1448	0.0261	30.8766	<.0001
OtherAsianInd	1	0.3465	0.1099	9.9432	0.0016
Education1Ind	1	1.5573	0.0836	347.0264	<.0001
Education2Ind	1	-0.2860	0.0735	15.1448	<.0001
Education3Ind	1	-0.1746	0.0611	8.1544	0.0043
Age*MaritalStatusSIn	1	0.0171	0.00144	142.0453	<.0001
Age*SexInd	1	0.00573	0.00129	19.8493	<.0001
Age*Education1Ind	1	-0.0216	0.00159	185.1246	<.0001
Age*Education2Ind	1	0.00748	0.00166	20.3883	<.0001
Age*Education3Ind	1	0.00363	0.00121	8.9500	0.0028

Table 3: Random Forest Variable Importance Output – Accident

Loss Reduction Variable Importance					
Variable	Number of Rules	Gini	OOB Gini	Margin	OOB Margin
Education1Ind	359	0.000225	0.00008	0.000451	0.00032
HospitalInd	556	0.000251	0.00006	0.000502	0.00031
AsianInd	292	0.000010	-0.00002	0.000020	-0.00002
Nursing	199	0.000033	-0.00002	0.000066	0.00002
WhiteInd	596	0.000209	-0.00003	0.000418	0.00020
OtherAsianInd	305	0.000033	-0.00005	0.000067	-0.00004
MaritalStatusSInd	515	0.000115	-0.00006	0.000229	0.00005
BlackInd	543	0.000140	-0.00007	0.000279	0.00003
Education2Ind	679	0.000105	-0.00010	0.000210	-0.00001
MaritalStatusMInd	738	0.000150	-0.00012	0.000300	0.00004
Education3Ind	762	0.000155	-0.00016	0.000309	-0.00002
SexInd	1060	0.000235	-0.00019	0.000470	0.00005
Age	788	0.000951	-0.00046	0.001901	0.00056

Table 4: Random Forest Variable Importance Output – Suicide

Loss Reduction Variable Importance					
Variable	Number of Rules	Gini	OOB Gini	Margin	OOB Margin
WhiteInd	490	0.092626	0.09280	0.185252	0.185561
BlackInd	296	0.071542	0.07133	0.143085	0.142677
HospitalInd	735	0.029692	0.02826	0.059384	0.057837
Age	1358	0.013320	0.00710	0.026639	0.020570
SexInd	1695	0.007675	0.00616	0.015350	0.013973
MaritalStatusMInd	1046	0.007215	0.00551	0.014430	0.012846
Education2Ind	1115	0.004577	0.00288	0.009155	0.007608
MaritalStatusSInd	867	0.003762	0.00273	0.007525	0.006438
Education1Ind	1140	0.002774	0.00094	0.005548	0.003766
Education3Ind	1583	0.002439	0.00030	0.004877	0.002822
Nursing	345	0.000305	-0.00000	0.000610	0.000209
AsianInd	516	0.000934	-0.00003	0.001867	0.000970
OtherAsianInd	443	0.000601	-0.00009	0.001202	0.000513

Table 5: Random Forest Variable Importance Output – Homicide

Loss Reduction Variable Importance					
Variable	Number of Rules	Gini	OOB Gini	Margin	OOB Margin
WhiteInd	518	0.095026	0.094099	0.190051	0.189158
BlackInd	276	0.075814	0.076033	0.151628	0.151902
HospitalInd	728	0.026341	0.025131	0.052681	0.051387
SexInd	1765	0.008372	0.006598	0.016744	0.015090
Age	1432	0.012210	0.005845	0.024420	0.018355
MaritalStatusMInd	1065	0.006961	0.005319	0.013923	0.012400
Education2Ind	1092	0.004874	0.003423	0.009748	0.008300
MaritalStatusSInd	877	0.004107	0.002810	0.008215	0.006942
Education1Ind	1137	0.002372	0.000531	0.004744	0.003009
AsianInd	508	0.000914	0.000085	0.001829	0.001029
Education3Ind	1652	0.002467	0.000050	0.004935	0.002564
OtherAsianInd	440	0.000774	0.000041	0.001547	0.000831
Nursing	363	0.000286	0.000003	0.000571	0.000190

Table 6: Gun-related suicides by race

	Caucasian (Non-Hispanic)	African-American	American Indian	Asian	Pacific Islander/ Other Asian	Hispanic	Total
Number of suicides	35,822	2,379	468	709	469	3,292	**43,139**
Number of gun-related suicides	7,521	290	101	63	45	425	**8,445**
Percent of suicides that are gun-related	21.00	12.19	21.58	8.89	9.59	12.91	**19.58**

Table 7: Results of Chi-Squared Test for Race (Suicide)

Statistics for Table of Gun by Race

Statistic	DF	Value	Prob
Chi-Square	5	303.5250	<.0001
Likelihood Ratio Chi-Square	5	338.0454	<.0001
Mantel-Haenszel Chi-Square	1	179.2149	<.0001
Phi Coefficient		0.0839	
Contingency Coefficient		0.0836	
Cramer's V		0.0839	

Sample Size = 43139

Table 8: Gun-related homicides by race

	Caucasian (Non-Hispanic)	African-American	American Indian	Asian	Pacific Islander/ Other Asian	Hispanic	Total
Number of homicides	5,413	8,032	269	160	164	2,802	**16,840**
Number of gun-related homicides	2,949	6,331	124	79	97	1,843	**11,423**
Percent of homicides that are gun-related	54.48	78.82	46.10	49.38	59.15	65.77	**67.83**

Table 9: Results of Chi-Squared Test for Race (Homicide)

Statistics for Table of Gun by Race

Statistic	DF	Value	Prob
Chi-Square	5	981.2004	<.0001
Likelihood Ratio Chi-Square	5	985.3036	<.0001
Mantel-Haenszel Chi-Square	1	0.4584	0.4984
Phi Coefficient		0.2414	
Contingency Coefficient		0.2346	
Cramer's V		0.2414	

Sample Size = 16840

Figure 1: An example of a single tree in a random forest exploring accident gun-related deaths.

The variables used in this example tree are not necessarily the variables produced in analysis. Also, many more nodes would have occurred in a real tree. This tree used two variables for the purpose of easy visualization.

Figure 2: Proportion of Gun-Administered Suicides within Race.
A segmented bar graph displaying the proportion of people who committed suicide with a gun out of all suicides within race. The five races represented are Caucasian (Non-Hispanic), African American, American Indian, Asian, Pacific Islander/Other Asian, and Hispanic.

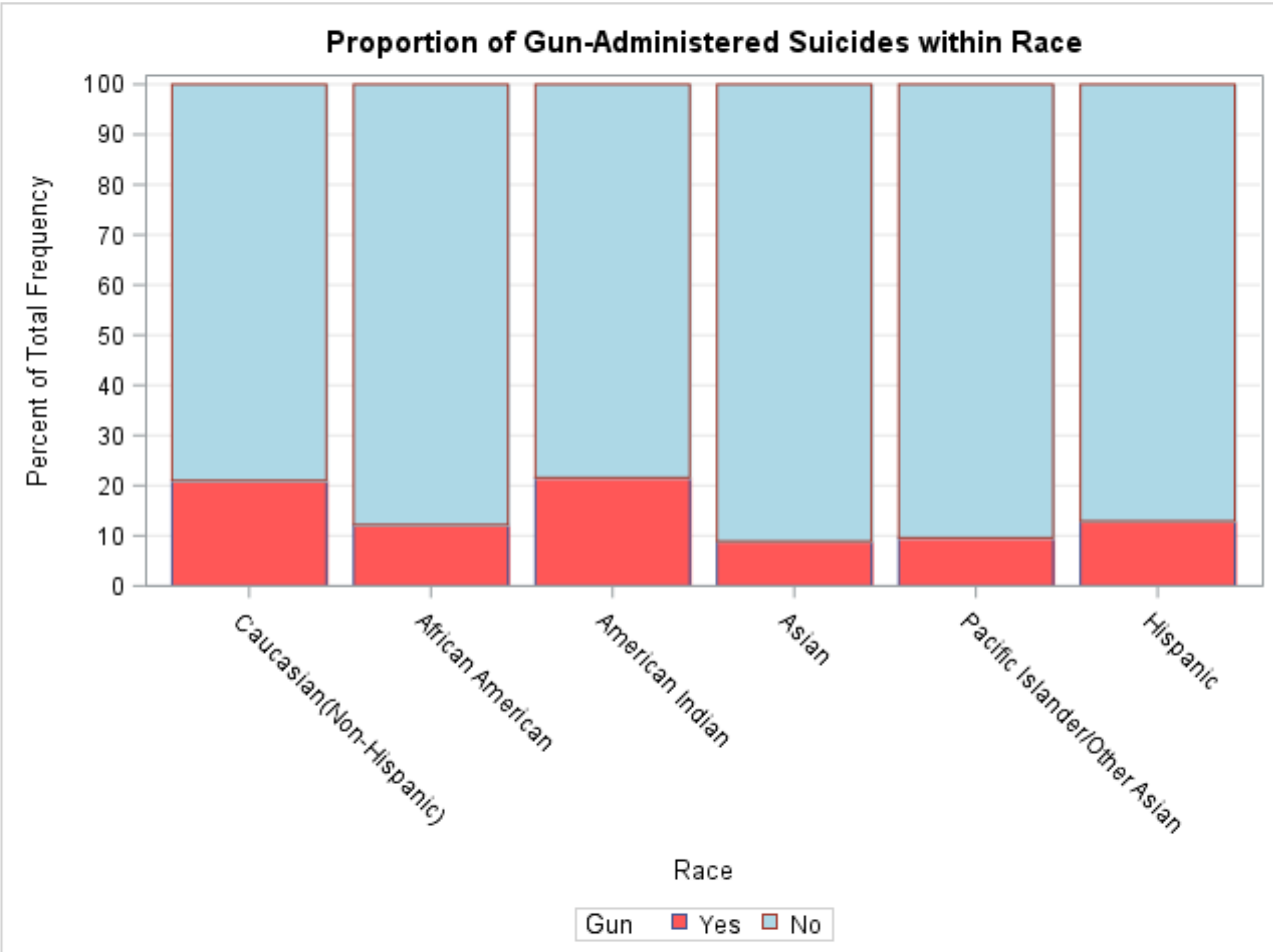

Figure 3: Proportion of Gun-Administered Suicides by Marital Status within Gender.
A butterfly bar chart displaying the proportion of men (left) and women (right) who committed suicide with a gun by marital status. The four statuses are widowed (W), single (S), married (M), and divorced (D).

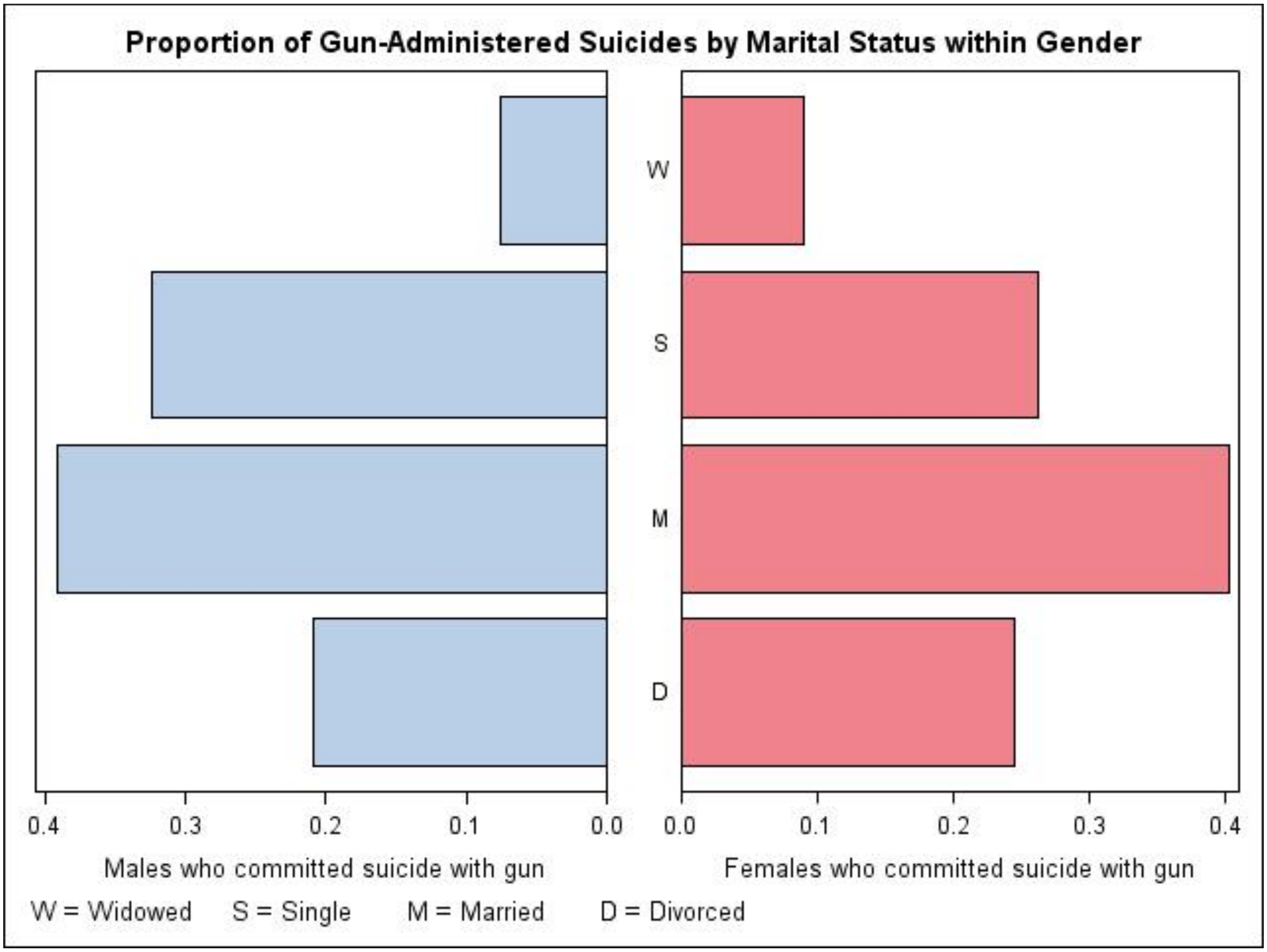

Figure 4: Age of Gun-Related Deaths by Manner of Death

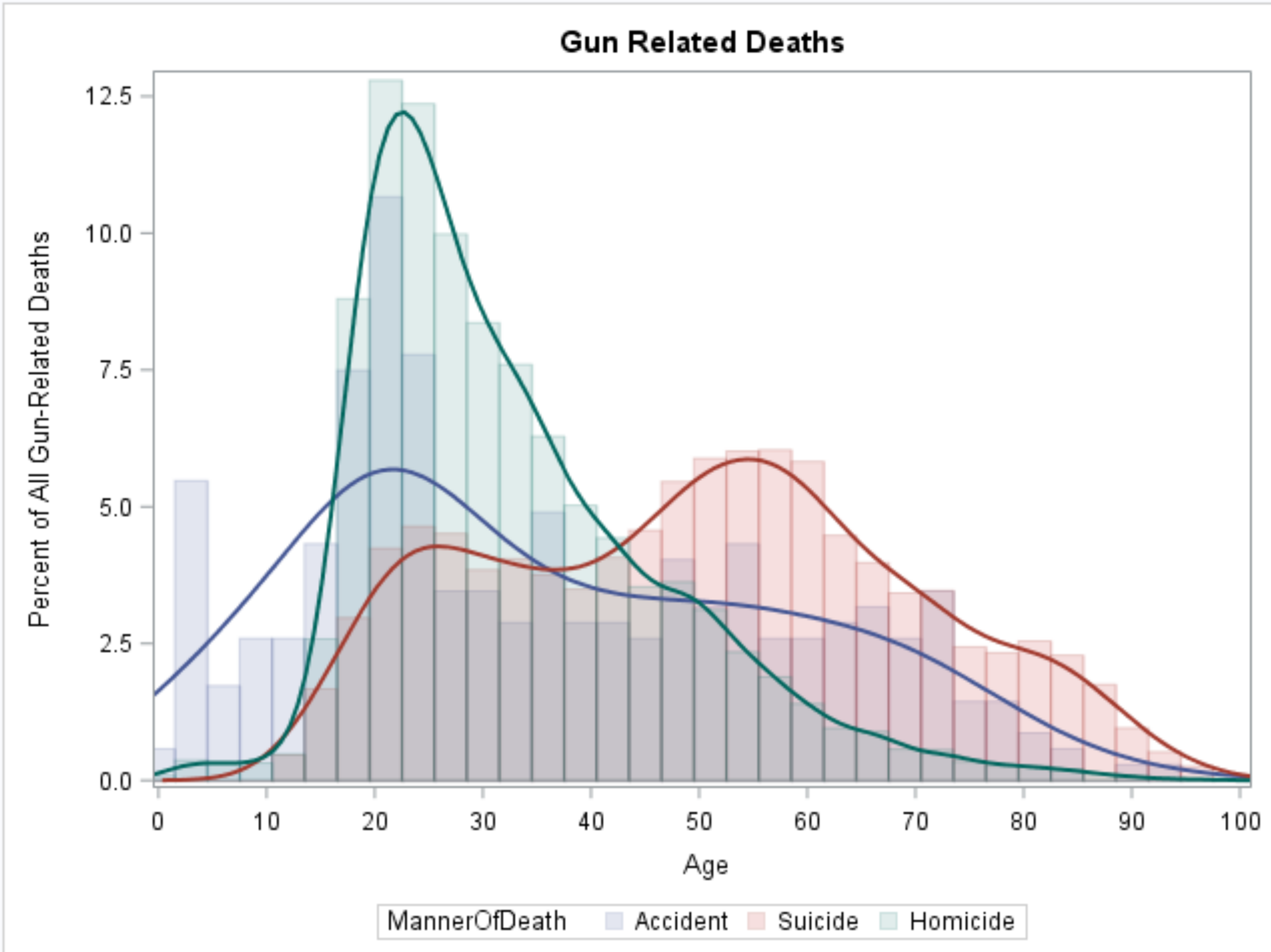

Figure 5: Type of Gun-Related Death by Education Level

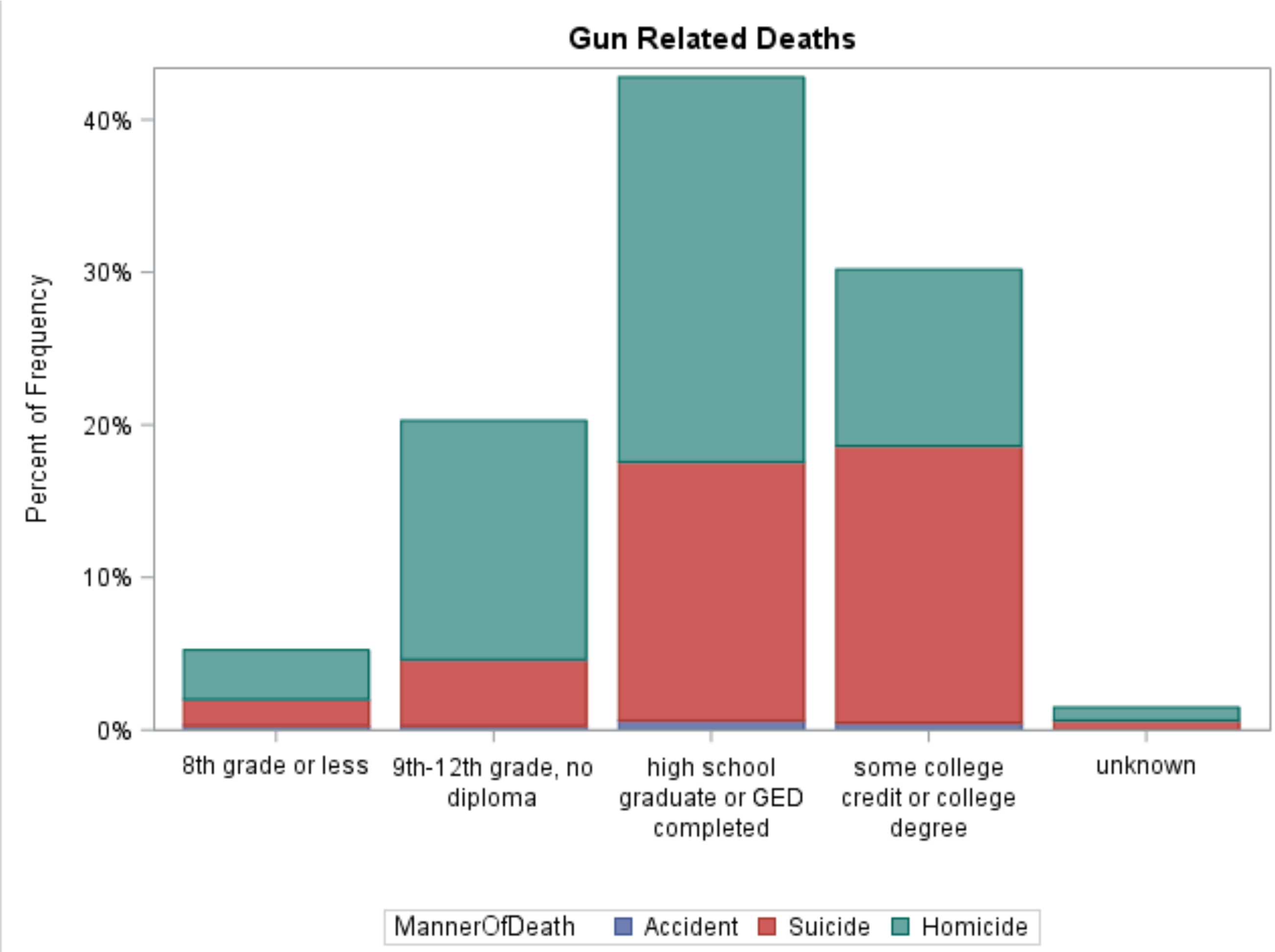

Figure 6: Proportion of Gun-Related Homicides within Race.

A segmented bar graph displaying the proportion of people killed in gun-related homicides out of all homicides within race. The five races represented are Caucasian (Non-Hispanic), African American, American Indian, Asian, Pacific Islander/Other Asian, and Hispanic.

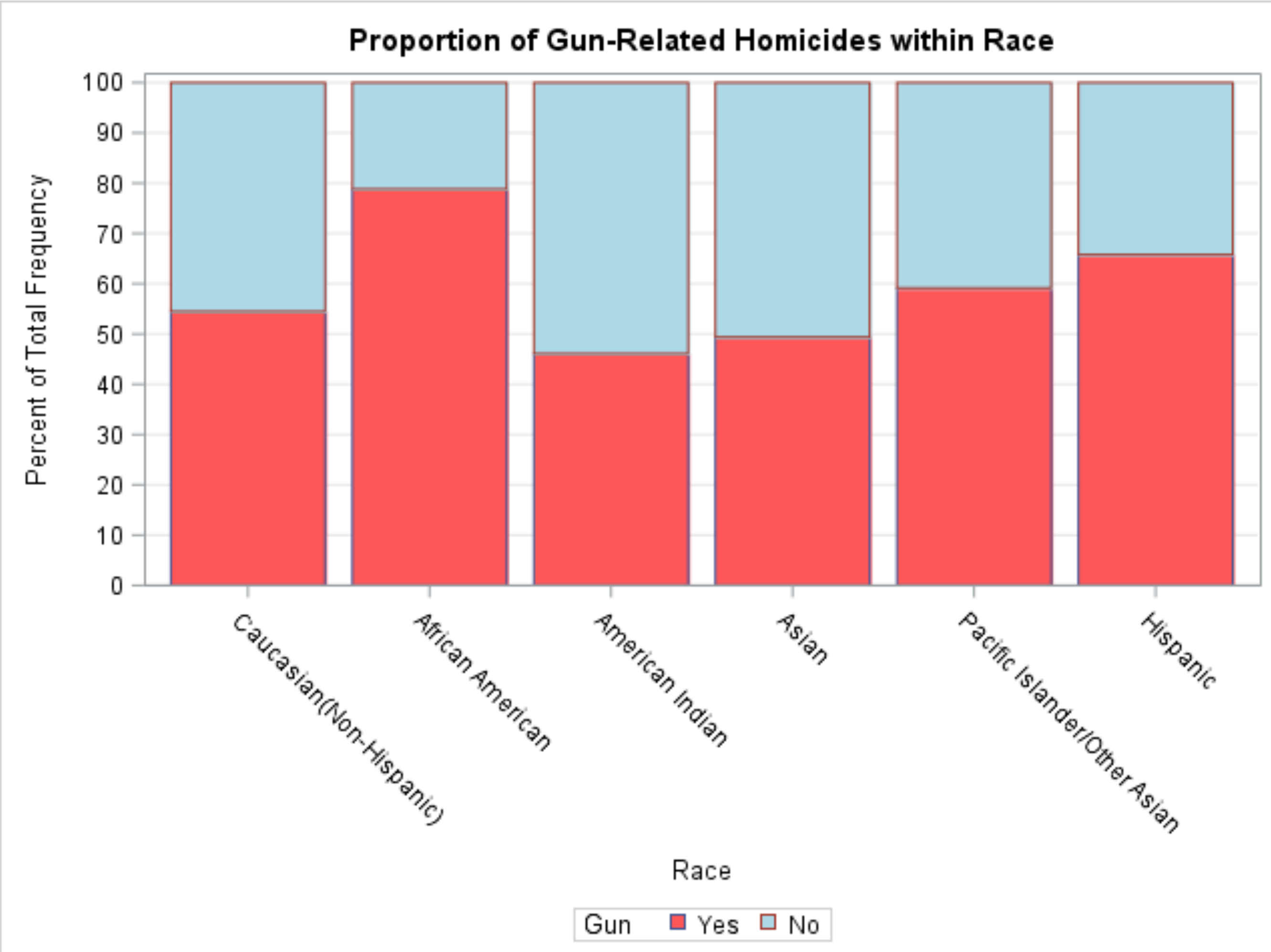

Figure 7: Proportion of Gun-Related Homicides by Marital Status within Gender.
A butterfly car chart displaying the proportion of men (left) and women (right) who were killed in a gun-related homicide by marital status. The four statuses are widowed (W), single (S), married (M), and divorced (D).