Introduction

Murder and murder trends are heavily scrutinized as various law enforcement agencies work to solve open murder cases and decrease murder rates. The dataset examined in this report comes from the Murder Accountability Project (MAP)[1] and includes case-level data from the Uniform Crime Report (UCR) in addition to data collected by MAP under the Freedom of Information Act. The original dataset included data on murders from 1976-2016 but has been cut down to allow a more targeted examination of the data. The dataset to be examined includes all records from 2000-2016.

The UCR is compiled annually by the Federal Bureau of Investigation (FBI) from reports sent in by law enforcement agencies across the United States. Because the FBI is not actively collecting all the data themselves, there is a higher possibility of error and missing information. Different states likely have different report styles and requirements; information required by one state, like ethnicity, may not be required in another. There is also a high chance of human error in the reporting process as many agencies are understaffed and overworked. One last possibility is intentional misreporting of murder. An agency might misreport murders to make it seem like they are doing a better job in solving murders or decreasing the murder rate than they are. In addition to these possible sources of error, only murders reported to or discovered by law enforcement agencies are recorded. It is almost certain that the dataset does not include information on every murder that has happened in the United States from 2000-2016. However, this is likely the most complete record available and referenced frequently when examining murder and crime in the United States.

The dataset includes a large amount of information for each murder recorded. A brief overview of each field is given below:

- ID – A unique identifier created based on the year, month, incident, and originating agency.
- CNTYFIPS – The state and county of the originating agency.
- ORI – A code representing the originating agency's state, county, and number within the state.
- STATE – The state of the originating agency.
- AGENCY – The name of the originating agency.
- AGENTYPE – The type of law enforcement agency (i.e. Sheriff, County Police, etc.).
- SOURCE – Identifies the source of the record, either FBI for the UCR or data collected by MAP under the Freedom of Information Act.
- SOLVED – Indicates if the murder was marked solved at the time of report creation.
- STATENAME – Abbreviation used by the FBI for the original reporting agency.
- YEAR – The year of the homicide or when the victim's body was recovered.
- MONTH – The month the homicide occurred or when the victim's body was recovered.
- INCIDENT – Describes the case number within the month that the homicide occurred, used in creating a unique ID.
- ACTIONTYPE – Describes the type of report (i.e. Normal Update or Adjustment to a previous report).
- HOMICIDE – Indicates whether the report refers to a "Murder or Non-negligent Manslaughter" or "Manslaughter by Negligence".
- SITUATION – Describes the number of victims and offenders in the incident.
- VICAGE – The age of the victim, may be unknown if the age of the victim could not be determined.
- VICSEX – The sex of the victim, may be unknown if incomplete remains were found.
- VICRACE – The race of the victim, may be unknown if the race of the victim could not be determined.
- VICETHNIC – Reports if the victim was Hispanic. However, many agencies do not report ethnicity, therefore a value of unknown in this field may mean ethnicity was not reported or could not be determined.
- OFFAGE – The age of the offender, may be unknown if the offender has not been identified.
- OFFSEX – The sex of the offender, may be unknown if the offender has not been identified.

[1] http://www.murderdata.org/p/data-docs.html

- OFFRACE – The race of the offender, may be unknown if the offender has not been identified.
- OFFETHNIC – Reports if the offender was Hispanic. However, many agencies do not report ethnicity therefore a value of unknown in this field may mean ethnicity was not reported or could not be determined.
- WEAPON – Describes the weapon used in the crime, if known.
- RELATIONSHIP – Describes the relationship between the victim and offender, if known.
- CIRCUMSTANCES – Describes the circumstances surrounding the incident, or a theory if exact circumstances are unknown.
- SUBCIRCUM – Describes conditions in which the victim is reported to have been a criminal offender (i.e. Felon killed in commission of a crime).
- FILEDATE – Describes the date the record was reported, not the date of the actual crime.
- FSTATE – Describes the state the homicide was reported in.
- MSA – The code of the Metropolitan Statistical Area from which a record was reported.

Research Questions

This report examines a dataset of murders in the United States from 2000-2016, and briefly murders from 1976-2016 in an attempt to identify trends, and what, if any, relationships exist between gender, age, and other aspects of a murder, how the data compares to common assumptions about murder.

Descriptive Statistics

The dataset includes 272,082 murder records. This section starts to break down the data by looking at descriptive statistics and graphs to allow for an overall understanding of the data.

Overview of Murder in the United States

Simple bar graphs and summary statistics allow for the overall understanding of the dataset and give an over view of murder in the United States. The following sections examine murder trends over the years, month, states, and weapon types included in this dataset. These graphs and statistics include all 272,082 records in the dataset.

Year

Counting murders by year allows for an understanding of murder frequency over time. Figure 1 shows the number of murders for each year from 2000-2016. Although there does seem to be a minor decline in the frequency of murders per year starting in 2008 and then an increase in frequency starting in 2015, the graph shows that the frequency of murders has not majorly fluctuated much over the 16 years being examined. The slight increase around 2016 seems to support the general public's opinion that the United States is more violent and dangerous now than it has been in the past. However, looking at the trends over 40 years, it is clear that the number of murders per year has been decreasing, this can be seen in Figure 2.

Figure 1: Murder Count by Year

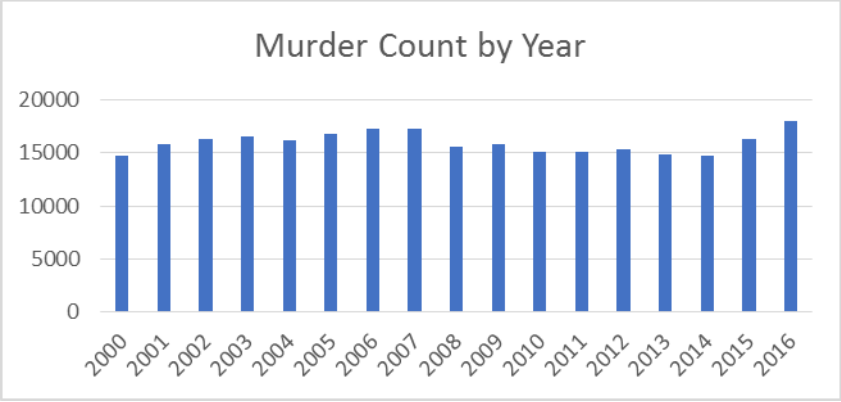


Table 1: Summary Statistics – Murder County by Year

Summary Statistics – Murders by Year	
Smallest Value	14,692 (2000)
First Quartile	15,126
Median	15,857 (2009)
Third Quartile	16,692
Largest Value	18,022 (2016)
Mean	16,005
Standard Deviation	959
Total	272,082

The summary statistics for murders from 2000-2016 can be seen in Table 1. Over the past sixteen years, 2000 had the least murders with 14,692 murders, 2016 had the most murders with 18,022, and 2009 has the median number of murders with 16,692. With a mean of 16,005 murders a year and a standard deviation of 959, it is clear that the variation over these sixteen years is not overly great.

Looking over the past 40 years shows a different picture. From Figure 2, it is clear that there are far less murders now then there have been, with the decline starting in 1994. Interestingly, the years with the greatest and least number of murders occurred within a decade of each other, 1993 and 1999 respectively. Additionally, 2016 is now the median, whereas before, when only looking at the past 16 years, it was the largest. With a mean of 18,349 murders per year and a standard deviation of 2,727, it is clear that the number of murders per year had varied more over this greater length of time.

Figure 2: Murder Count by Year 1976-2016

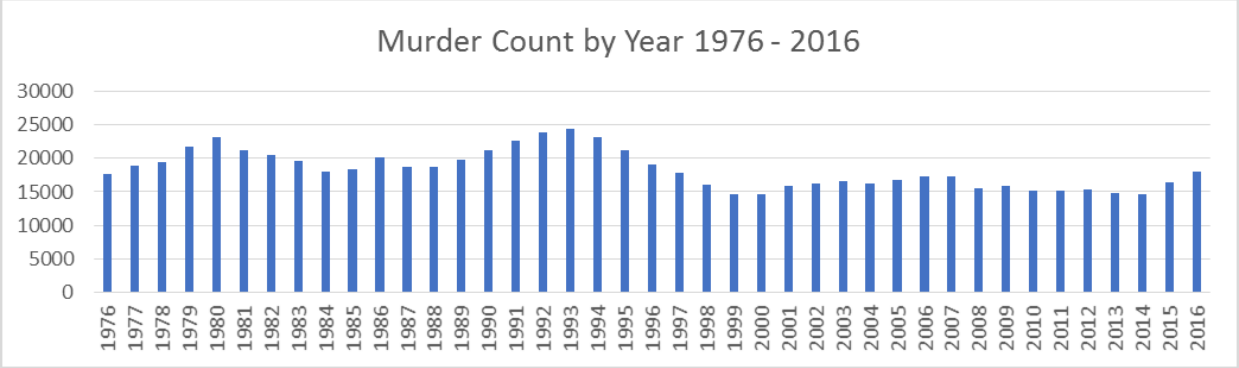


Table 2: Summary Statistics – 1976-2016

Summary Statistics – Murder Count by Month	
Smallest Value	18,248 (February)
First Quartile	21,818
Median	23,109 (September + October)
Third Quartile	23,537
Largest Value	25,568 (July)
Mean	22,674
Standard Deviation	1,752
Total	272,082

Month

Figure 3 shows murder frequency by month; there appears to be an increase towards the summer and a decrease towards the winter months. This trend matches well with the assumption that there is more crime, in general, during the summer months as more people are out of their houses and on vacations, leaving their homes vulnerable targets for property crimes.

Figure 3: Murder Count by Month

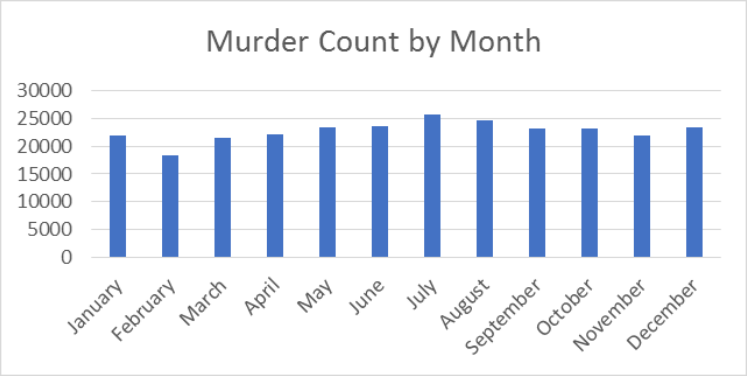


Table 3 shows the summary statistics for murders by month. February is the month with the least number of murders with 18,248 murders during that month over the past sixteen years. July has the most with 25,568, and September and October average out to the median of 23,109. With a mean of 22,674 murders each month and a standard deviation of 1,752, it is clear there is some variation in the number of murders committed each month.

Table 3: Summary Statistics – Murder Count by Month

Summary Statistics – Murder Count by State	
Smallest Value	173 (North Dakota)
First Quartile	746
Median	2,938 (Colorado)
Third Quartile	7,415
Largest Value	38,129 (California)
Mean	5,335
Standard Deviation	6,766
Total	272,082

State

Examining the murders by state shows a wide range and some interesting patterns. Notably, California and Texas have the two highest murder counts. Florida, New York, and Pennsylvania follow after. This directly correlates to the population of each state. California is the most populated state with almost 40 million residents, Texas, Florida, New York, and Pennsylvania are the next most populated states. It stands to reason that population and population density have a role to play in the frequency of murders. More people living in one area means there are more people with a disposition suited for killing and living in densely populated areas likely exasperates this issue. This is interesting as, when talking about murder and violence, many people will to bring New York City and Chicago into the conversation, which makes sense as both of the states those cities are in rank in the top 10 states with the largest population. However, rarely do people mention the murder rates in Florida or even Texas. Figure 4 shows the murder counts for each state and the District of Colombia from 2000-2016.

Figure 4: Murder Count by State

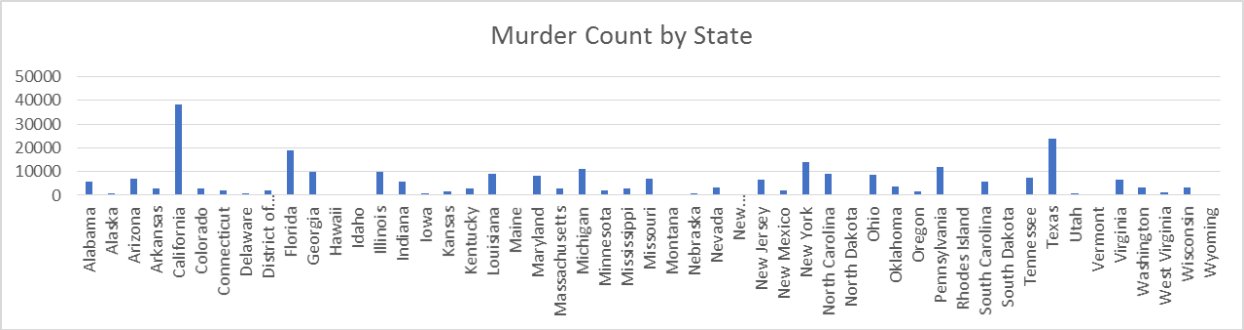


Table 4: Summary Statistics – Murder Count by State

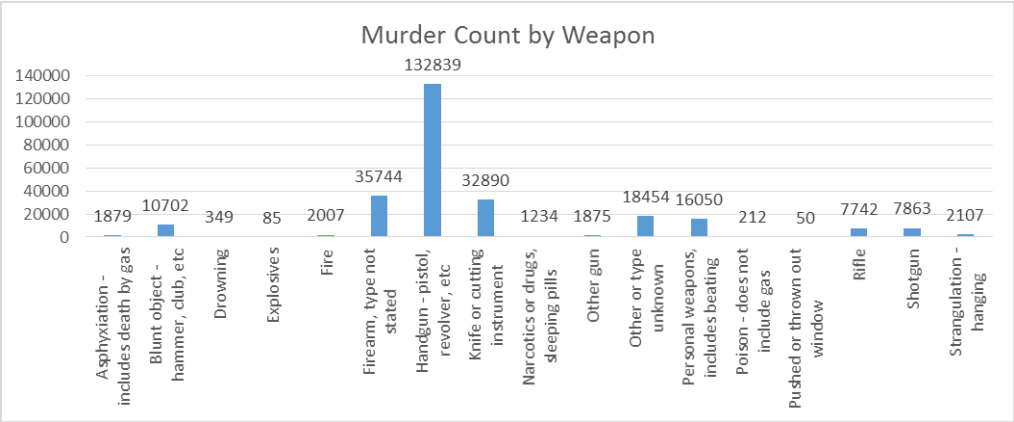
Summary Statistics – Murder Count by State	
Smallest Value	173 (North Dakota)
First Quartile	746
Median	2,938 (Colorado)
Third Quartile	7,415
Largest Value	38,129 (California)
Mean	5,335
Standard Deviation	6,766
Total	272,082

The state with the least murders is North Dakota with only 173 murders from the last sixteen years. That is in stark contrast to the 38,129 murders that have occurred in California. Among the 50 states, plus the District of Columbia, Colorado has the median number of murders at 2,938. From examining the graph, it is clear there is a large variation in the number of murders that occur in each state. This is also shown in the standard deviation of 6,766 and the mean of 5,335 murders in each state.

Weapon

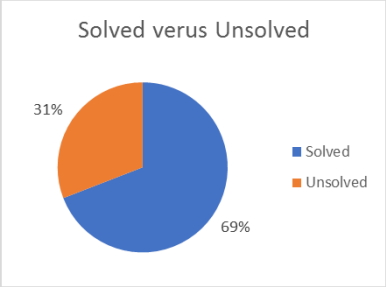
Figure 5 shows the breakdown of murders by weapon. Clearly, the majority of murders are committed using some type of firearm, such as a handgun or rifle. This is not particularly surprising as guns are fairly easy to purchase in the United States and do not require a lot of skill or ability to use. They would allow a smaller or physically weaker person to kill a larger or physically stronger person. Guns also allow one person to kill a larger number of people at a single time than a blunt object or knife would. That being said, knives, blunt objects, and personal weapons are the most common weapons following firearms.

Figure 5: Murder Count by Weapon



Before getting into the specifics of each crime, it is important to understand the limitations of this dataset. As seen in Figure 6, 31% of murders were not solved at the time the report was created. This means that, for those records, a decent amount of information was not available at the time of reporting, such as information about the offender. In addition, even though the murder has been reported as solved, there may still be data missing if it was not available, not included in the report, or unable to be determined.

Figure 6: Solved versus Unsolved



Victim and Offender Demographics

Looking at various aspects of the victim and offenders, such as their age, gender, and relationship, may allow for a better understanding of murder trends and the motive involved. This information may be used to help decrease the murder rates as we would know which groups and demographics are committing the most murder and can use that information to determine possible motives and trends, which can then be addressed by law enforcement and the entire community, if needed.

Gender

The first variable of major note for this dataset is the gender of the victim and the offender. From the graphs below, it is clear that males are far more likely than females to be both the victim and the murderer. It is also clear that females are more likely to be victims than they are to be the perpetrator. This is not particularly surprising as the general assumption is that males are involved in more crime than females are. The data here supports that assumption. However, it is a bit shocking that women make up less one quarter of the victims. Many assume that women would make up a larger portion of the murder victims, but, taking gang related violence into consideration, the number seems more appropriate. Figures 7 and 8 below show the breakdown of victims and offenders by gender.

Figure 7: Victim Gender

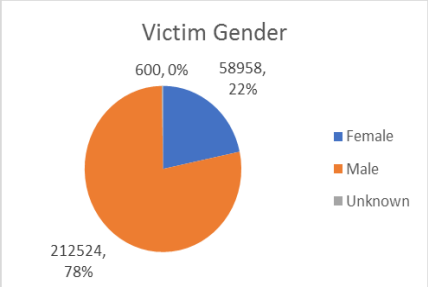
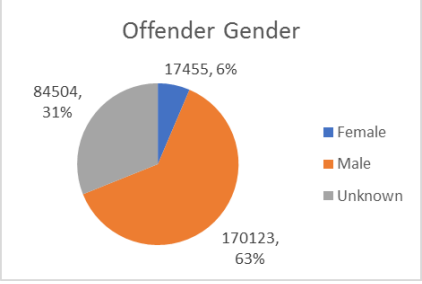


Figure 8: Offender Gender



Age

The next variable of interest is the age of the victim and offender. For this dataset, an age of “0” indicates a person younger than one year old, an age of “99” indicates a person aged 99 or older, and an age of “999” indicates the age is unknow. The trends for victims and offenders are similar in this case as well. The trends for offenders may be flawed as a great number, almost 100,000, of them do not have an age that was known at the time of reporting. However, from these graphs, it is clear that as a person ages, they become more likely be involved in a murder until they are 30, at which point it becomes far less likely. This would match with the societal perception that younger people, particularly those around 18-24, are more likely to be involved in crime than older people are.

Figure 9: Ages of Victims

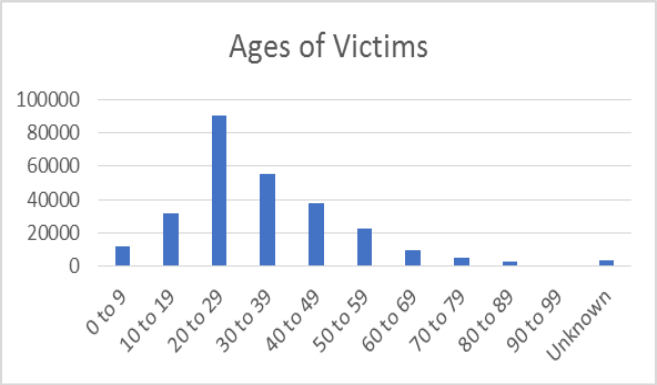


Table 5: Summary of Statistics – Victim Age

Summary Statistics – Victim Age	
Smallest Value	0
First Quartile	22
Median	29
Third Quartile	42
Largest Value	99
Mean	32.88
Standard Deviation	16.19
Margin of Error	0.0613
Total	268,260

Figure 10: Ages of Offenders

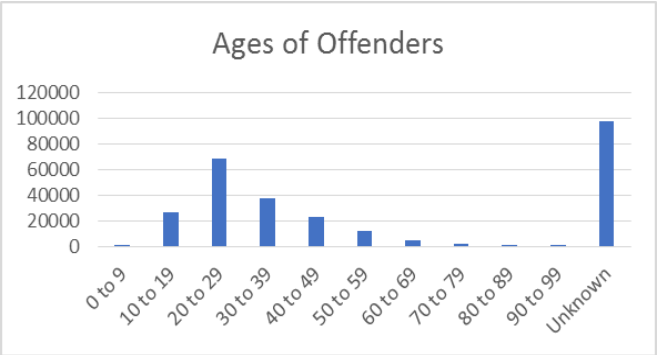


Table 6: Summary Statistics – Offender's Age

Summary Statistics – Offender Age	
Smallest Value	1
First Quartile	21
Median	28
Third Quartile	39
Largest Value	99
Mean	31.58
Standard Deviation	13.19
Margin of Error	0.0613
Total	174,242

The statistics in this case exclude records where the age was unknown, as it would skew the results. If the offender's age was unknown, the record was excluded from the offender statistics but not the victim statistics. This was done to ensure the statistics calculated were as accurate and representative of all known murder cases as possible. Figures 9 and 10 make it obvious that the trends for victim and offender ages are similar. The statistics back this up. The youngest victim is less than one year old, likely a newborn. The youngest offender is around one. The median age for victims and offenders is just one year off from each other, 29 for victims and 28 for offenders. The oldest victims and offenders are both 99 or older. From the means and standard deviations, it is clear that there is a decent amount of variation in the ages of victims and offenders, a fact which is also supported by the figures mentioned previously.

Lastly, Table 7 shows the relationship between the sex of the offender and the sex of the victim. Male-Male murders consist of almost half of all murders reported. Females seem more likely to kill males than other females. And, aside from the victim sex unknown category, Female-Female violence makes up the smallest percentage. This just further shows that males are more likely to be involved in violence and murder than females are.

Table 7: Relationship Between Victim and Offender Sex

Offender Sex	Victim Sex		
	Female	**Male**	**Unknown**
Female	1.74%	4.65%	0.02%
Male	15.91%	46.43%	0.08%
Unknown	4.01%	26.93%	0.12%

Race

The next variable of particular note is the race of the victim and offender. Figures 11 and 12 show the breakdown of victim and offenders by their race. Again, it is possible that an offender's race is unknown if they were never caught. In addition, a victim's race might by unknown if incomplete remains were found. White and black persons make up the majority of offenders and victims which is not surprising considering they make up a majority of the population as well. White persons make up a slightly larger percentage of the offenders but a slightly smaller percentage of victims than black persons. This contrasts with the trends seen between victims and offenders previously, where the largest portions of offenders also made up the largest portion of victims. This also contrasts with society's general assumption that black persons commit more crime and murder than white persons do.

Figure 11: Victim Race

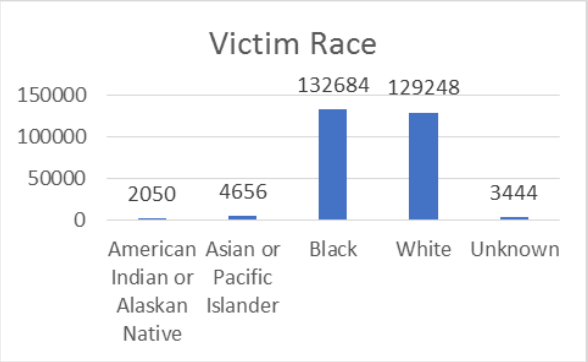
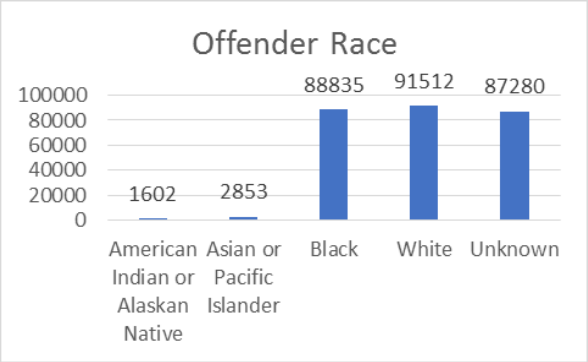


Figure 12: Offender Race



Relationship

The last variable of interest as a part of the summary is the relationship between the victim and the offender. For this dataset, the relationship is listed in terms of the offender. For example, a record with a relationship listed as "Wife" means that the victim was the offender's wife. A count of each occurrence can be seen in Figure 13, however, due to the large number of categories, the graph is a bit difficult read and slightly misleading. The "relationship" category with the most records is "Relationship not determined," which likely has to do with the number of unsolved cases included in this dataset. That category is followed by Acquaintance and Stranger. Due to the number of categories, it is difficult to tell just how likely a person is to be murdered by someone they know. Figure 14 is presented to condense this information. The "Known to Victim" category contains the following subcategories: Acquaintance, Boyfriend/Girlfriend, Brother/Sister, Common-law Husband/Wife, Daughter/Son, Employee, Employer, Ex-husband/wife, Father/Mother, Friend, Homosexual Relationship, Husband/Wife, In-Law, Neighbor, Other – Known to Victim, Other Family, Stepdaughter/son, and Stepmother/father. The "Unknown to Victim" includes Strangers. From this, it is clear that a person is more likely to be murdered by someone they know than someone they do not. Figure 15 further breaks down the "Know to Victim" category. From this graph, it is clear that a person is also more likely to be murdered by a person they know that is not their family than they are to be murdered by one of their family members. This family member category contains the following subcategories: Brother/Sister, Common-law Husband/Wife, Daughter/Son, Ex-husband/wife, Father/Mother, Husband/Wife, In-Law, Other Family, Stepdaughter/son, and Stepmother/father. The Non-Family categories contains the following subcategories: Acquaintance, Boyfriend/Girlfriend, Employee, Employer, Friend, Homosexual Relationship, Neighbor, and Other – Known to Victim.

The relationship aspect is particularly interesting as much of the media focuses on murders wherein the victim and offender did not know each other; murders that seem more like random acts of violence than what may be considered "private matters" of a family. This, however, causes people to misunderstand the nature of murder and over focus on the idea that a stranger could murder them, causing them to forget that their family is just as capable.

Figure 13: Relationship between Victim and Offender

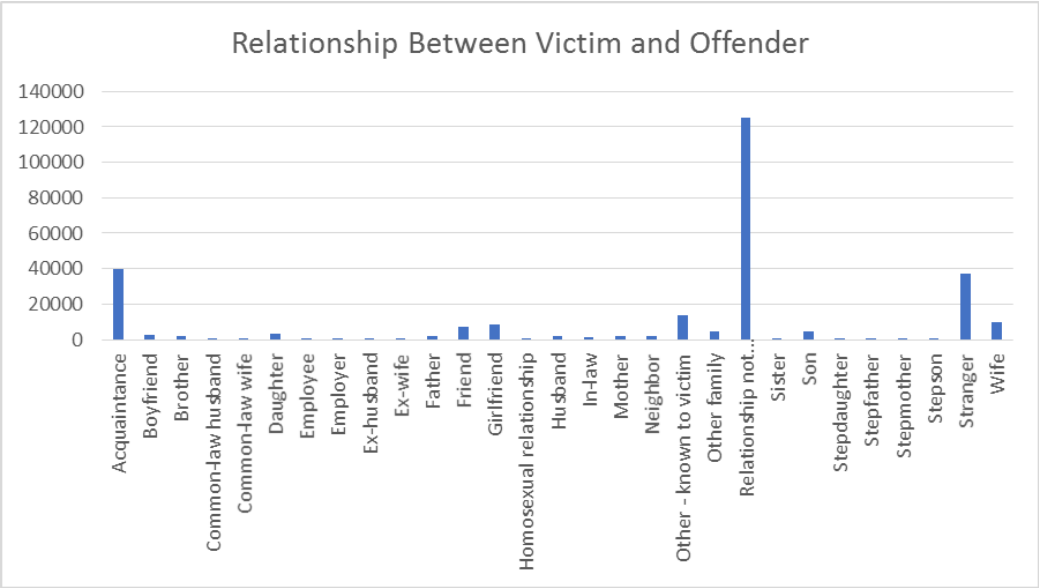


Figure 14: Relationship by Category

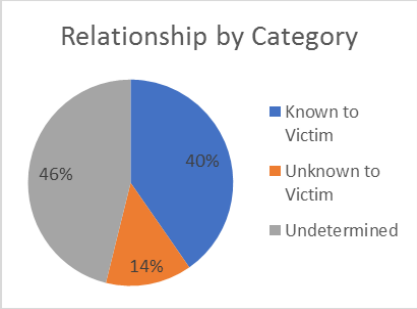
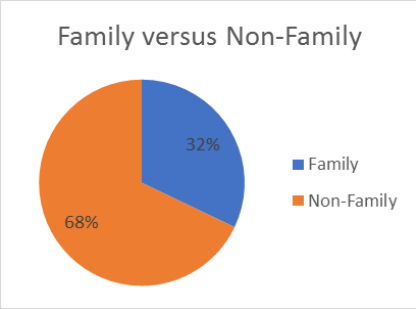


Figure 15: Family versus Non-Family



Inferential Statistics

Inferential statistics allow for a random sample of the data to be taken and examined. From that examination, conclusions about the dataset can be drawn. In the following sections, conclusions about the relationship between age, gender, relationship, and weapon choice will be drawn using inferential statistics techniques. Each of the samples used for the following tests have been taken from a subset of the entire dataset, where some records have been removed if variables of interest were missing or unknown. The two tests are the test of independence and ANOVA. A test of independence is a test of association between two categorical variables. The null hypothesis is that the two categorical variables being tested are independent. The alternative hypothesis is that the two categorical variables are not independent. For this test, the observed and expected frequencies are used to compute the test statistic. If the test statistics is greater than or equal to the critical value, the null hypothesis can be rejected. This critical value is based on the degrees of freedom, which, in turn, is based on the number of rows and columns in the table and the chi-square distribution table. ANOVA examines the equality of three or more population means. This test has three assumptions: (1) the dependent variable is normally distributed, (2) the variance of the response variable is the same for all populations, and (3) all observations are independent. The null hypothesis is that all population means are equal, the alternative being that they are not. To reject the null hypothesis, the *F* value must be greater than or equal to the critical value. The critical value is based on an *F* distribution table, the number of groups, and the number of elements in those groups.

Relation Between Age and Relationship

In the previous section, Figure 14 showed that victims are killed by someone they know more often than by someone they do not know. To examine the possible relationship between age and relationship of the victim to the offender, a random sample of 1000 records was created. This sample came from a subset of the dataset in which all records where the age or sex of the offender or victim or the relationship was unknown have been excluded. The relationships have also been grouped together to allow for easier understanding of the results and address the gendered relationships such as wife and husband. The "Family" category contains the following subcategories: Brother/Sister, Common-law Husband/Wife, Daughter/Son, Ex-husband/wife, Father/Mother, Husband/Wife, In-Law, Other Family, Stepdaughter/son, and Stepmother/father. The "Non-Family Known" category contains the following subcategories: Acquaintance, Boyfriend/Girlfriend, Employee, Employer, Friend, Homosexual Relationship, Neighbor, and Other – Known to Victim. The "Unknown to Victim" category only includes Strangers. From this, the following observed frequencies and expected frequencies were calculated, shown in Tables 8 and 9 respectively. Two hypotheses were developed to test this relationship. The first, the null hypothesis, states: "The relationship between the victim and offender is independent of age." The second, the alternative hypothesis, states: "The relationship between the victim and the offender is not independent of age."

Using these tables and hypotheses, a test of independence was run. The critical value to reject the null hypothesis is 26.30 and, with a c^2 value of 72.30, the null hypothesis can be rejected at the .05 level of significance. Therefore, the assumption that the relationship between the victim and offender is independent of age is rejected. This is not particularly surprising as many relationships themselves are dependent on age. At age 13, a person is not likely to have a significant other or employer, but they will have family or friends. As a person ages and change, so do the relationships in their life. For a murderer, that would mean their potential victim base also changes. This also answers one of the research questions asked in this report—there is a relationship between age and relationship between the victim and offender.

Table 8: Observed Frequencies – Age and Relationship

Observed Frequencies – Age and Relationship									
	10-19	**20-29**	**30-39**	**40-49**	**50-59**	**60-69**	**70-79**	**80-89**	**90-99**
Family	16	78	56	40	28	7	4	6	2
Non-Family Known	81	207	108	87	42	5	5	2	0
Unknown to Victim	49	108	40	16	7	4	2		

Table 9: Expected Frequencies – Age and Relationship

Expected Frequencies – Age and Relationship									
	10-19	**20-29**	**30-39**	**40-49**	**50-59**	**60-69**	**70-79**	**80-89**	**90-99**
Family	34.60	93.14	48.35	33.89	18.25	3.79	2.61	1.90	0.47
Non-Family Known	78.40	211.04	109.55	76.79	41.35	8.59	5.91	4.30	1.07
Unknown to Victim	33.00	88.82	46.10	32.32	17.40	3.62	2.49	1.81	0.45

Relation Between Gender and Relationship

Next, the possible relationship between gender and relationship is examined. It is relatively well known that female serial killers are more likely to kill or target people they know than their male counterparts. This dataset is not limited to serial killers, but it would be interesting to see if that assumption can be applied to all female murderers. The same sample used in the previous section, "Relation Between Gender and Relation," has been used to examine the possible relationship between gender and relationship of the victim to the offender. From this, the following observed frequencies and expected frequencies were calculated, shown in Tables 10 and 11 respectively. Two hypotheses were developed to

test this relationship. The first, the null hypothesis, states: “The relationship between the victim and offender is independent of gender.” The second, the alternative hypothesis, states: “The relationship between the victim and the offender is not independent of gender.”

Using these tables and hypotheses, a test of independence was run. The critical value to reject the null hypothesis in this case is 5.99 and, with a c^2 value of 37.51, the null hypothesis can be rejected at the .05 level of significance. Therefore, the assumption that the relationship between the victim and offender is independent of gender is rejected. This result is not surprising. Given that females are less likely to commit murder than males, it makes sense that the murder they commit is against those they know rather than strangers. In addition, males are more likely to be involved in physical fights that may result in death and gang activity, meaning they are more likely than females to murder someone they do not know. This test also serves to answer another research question about the possible relationship between gender and relationship between victim and offender. This test shows that there is a relationship between the gender of the offender and the relationship to the victim.

Table 10: Observed Frequencies – Gender and Relationship

Observed Frequencies – Gender and Relationship		
	Female	**Male**
Family	43	194
Non-Family Known	46	491
Unknown to Victim	4	222

Table 11: Expected Frequencies – Gender and Relationship

Expected Frequency – Gender and Relationship		
	Female	**Male**
Family	22.04	214.96
Non-Family Known	49.94	487.06
Unknown to Victim	21.02	204.98

Relation Between Gender and Weapon Choice

The next relationship to be examined is the possible relationship between gender and weapon choice. As seen in a previous section, firearms are by far the most popular murder weapon, likely due to the ease at which they can be used to kill. To examine this possible relationship, a new random sample of 1000 was created. This sample came from a subset of the dataset wherein any record in which the victim or offender age or sex was unknown has been excluded. From this, the following observed and expected frequencies were calculated, shown in Tables 12 and 13 respectively. Two hypotheses were developed to test this relationship. The first, the null hypothesis, states: “Weapon choice is independent of gender.” The second, the alternative hypothesis, states: “Weapon choice is not independent of gender.”

Using these tables and hypotheses, a test of independence was run. The critical value to reject the null hypothesis in this case is 23.69 and, with a c^2 value of 63.27, the null hypothesis can be rejected at the .05 level of significance. Therefore, the assumption that weapon choice is independent of gender is rejected. The conclusion that weapon choice is dependent on gender is not surprising. Murdering a person with a blunt object or even a knife would likely require the ability to overpower the victim. This would be difficult, though not impossible, to do if the offender is physically weaker than the victim. In this case, a firearm of some sort would be a more effective choice. It would also require a brutality that most females do not exhibit. That is not to say they are not capable, just that it is not a trait typically cultivated by societal norms. This test answers another research question and shows that there is a relationship between weapon choice and the gender of the offender.

Table 12: Observed Frequency – Gender and Weapon Choice

Observed Frequency – Gender and Weapon Choice		
	Female	**Male**
Asphyxiation - includes death by gas	3	4
Blunt object - hammer, club, etc	5	43
Drowning	1	1
Fire	2	6
Firearm, type not stated	2	100
Handgun - pistol, revolver, etc.	30	425
Knife or cutting instrument	29	134
Narcotics or drugs, sleeping pills	3	1
Other gun	1	4
Other or type unknown	8	45
Personal weapons, includes beating	11	69
Poison - does not include gas		2
Rifle	1	28
Shotgun	2	33
Strangulation - hanging	1	6

Table 13: Observed Frequency – Gender and Weapon Choice

Expected Frequency – Gender and Weapon Choice		
	Female	**Male**
Asphyxiation - includes death by gas	0.693	6.307
Blunt object - hammer, club, etc	4.752	43.248
Drowning	0.198	1.802
Fire	0.792	7.208
Firearm, type not stated	10.098	91.902
Handgun - pistol, revolver, etc.	45.045	409.955
Knife or cutting instrument	16.137	146.863
Narcotics or drugs, sleeping pills	0.396	3.604
Other gun	0.495	4.505
Other or type unknown	5.247	47.753
Personal weapons, includes beating	7.92	72.08
Poison - does not include gas	0.198	1.802
Rifle	2.871	26.129
Shotgun	3.465	31.535
Strangulation - hanging	0.693	6.307

Relation Between Age and Weapon Choice

The last relationship to be examined is the possible relationship between age and weapon choice. The sample used to test this possible relationship is the same random sample from the previous section. From this sample, the following observed and expected frequencies were calculated, shown in Tables 14 and 15 respectively. Two hypotheses were developed to test this relationship. The first, the null hypothesis, states: “Weapon choice is independent of age.” The second, the alternative hypothesis, states: “Weapon choice is not independent of age.”

Using these tables and hypotheses, a test of independence was run. The critical value to reject the null hypothesis in this case is 137.70 and, with a c^2 value of 113.47, the null hypothesis cannot be rejected at the .05 level of significance. Therefore, the assumption that weapon choice is independent of age cannot be rejected. While this is not the expected result, it is also not surprising. As mentioned previously, firearms are a particularly useful weapon to murder a person with. They require little skill to handle, are relatively easy to acquire in the United States, and are just generally effective weapons. This test answers

another research question and shows that there is no definitive relationship between weapon choice and age of the offender.

Table 14: Observed Frequency – Age and Weapon Choice

Observed Frequency – Age and Weapon Choice									
	0-9	**10-19**	**20-29**	**30-39**	**40-49**	**50-59**	**60-69**	**70-79**	**80-89**
Asphyxiation		1	1	3	2				
Blunt object		5	20	10	10	3			
Drowning			1		1				
Fire		1	1	2	4				
Firearm		24	40	20	7	8	2		1
Handgun		73	200	93	50	23	11	4	1
Knife		22	65	34	31	9	2		
Narcotics			2	1	1				
Other gun		2	2			1			
Other or Type Unknown		6	20	8	11	8			
Personal weapons	1	9	37	21	6	5	1		
Poison					2				
Rifle		6	9	8	6				
Shotgun		4	16	6	4	3	2		
Strangulation			4	1		1	1		

Table 15: Expected Frequency – Age and Weapon Choice

Expected Frequency – Age and Weapon Choice									
	0-9	**10-19**	**20-29**	**30-39**	**40-49**	**50-59**	**60-69**	**70-79**	**80-89**
Asphyxiation	0.01	1.07	2.93	1.45	0.95	0.43	0.13	0.03	0.01
Blunt object	0.05	7.34	20.06	9.94	6.5	2.93	0.91	0.19	0.10
Drowning	0.00	0.31	0.84	0.41	0.27	0.12	0.04	0.01	0.00
Fire	0.01	1.22	3.34	1.66	1.08	0.49	0.15	0.03	0.02
Firearm	0.10	15.61	42.64	21.11	13.77	6.22	1.94	0.41	0.20
Handgun	0.46	69.62	190.2	94.19	61.43	27.76	8.65	1.82	0.91
Knife	0.16	24.94	68.13	33.74	22.01	9.94	3.1	0.65	0.33
Narcotics	0.00	0.61	1.67	0.83	0.54	0.24	0.08	0.02	0.01
Other gun	0.01	0.77	2.09	1.04	0.68	0.31	0.1	0.02	0.01
Other or Type Unknown	0.05	8.11	22.15	10.97	7.16	3.23	1.01	0.21	0.11
Personal weapons	0.08	12.20	33.44	16.56	10.8	4.88	1.52	0.32	0.16
Poison	0.00	0.31	0.84	0.41	0.27	0.12	0.04	0.01	0.00
Rifle	0.03	4.44	12.12	6.00	3.92	1.77	0.55	0.12	0.04
Shotgun	0.034	5.36	14.63	7.25	4.73	2.14	0.67	0.14	0.07
Strangulation	0.01	1.07	2.93	1.45	0.95	0.43	0.13	0.03	0.014

ANOVA

The last test used to examine and analyze this data set is an analysis of variance, also called ANOVA. This test is used to determine the equality of three or more population means. Two analyses were run, the first examined the victims while the second examined the offenders. In each case, four population means are being examined. These are the four races specifically identified in this dataset: American Indian or Alaskan Native, Asian or Pacific Islander, Black, and White. For the purposes of this analysis, no record with race listed as “Unknown” was used. This test attempts to discern if the mean age is equal among all these groups. An unequal mean may indicate that different races are involved in crime, specifically murder, at different ages, which may indicate varying motives for murder. It would also indicate that a murder-rate reducing tactic aimed at one group may not work well on another. In a sense, it may mean that race plays a role in murder trends and motives.

For this test, a random sample of 250 was taken from each group. As mentioned previously, no records were the race was listed “Unknown” was used. In addition, no records where the age was known was

used. Two hypotheses were developed and applied to both tests. The first, the null hypothesis, states: “All population means are equal.” The second, the alternative hypothesis, states: “Not all population means are equal.”

The box plot in Figure 16 and Tables 16 and 17 below summarize the results for victim age.

Figure 16 – Box Plot of Victim Age by Race

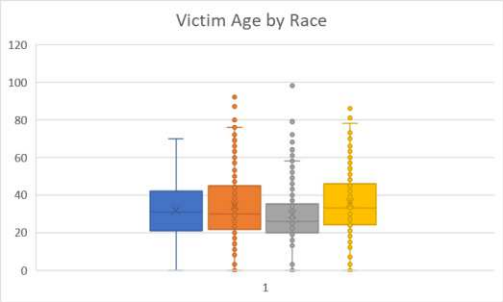


Table 16: Summary – Sample of Victim Age

Summary – Sample of Victim Age			
	Count	Average	Variance
American Indian or Alaskan Native	250	31.72	244.63
Asian or Pacific Islander	250	34.29	312.80
Black	250	29.63	203.62
White	250	35.75	279.45

Table 17 ANOVA – Victim Age

ANOVA – Victim Age						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	5531.73	3	1843.91	7.09	0.0001	2.61
Within Groups	259085.2	996	260.13			
Total	264616.9	999				

The box plot indicates that there is a difference between the means and variation of the four populations. This is supported by the ANOVA, as the F value, 7.09, is greater than the critical value of 2.61. This means the null hypothesis can be rejected and it can be concluded that not all of the population means are equal.

Figure 17 and Tables 18 and 19 summarize the result for the offenders.

Figure 17 – Box Plot of Offender Age by Race

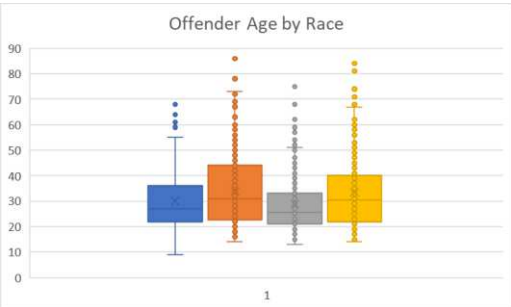


Table 18 Summary – Sample of Offender Age

Summary – Sample of Offender Age			
	Count	Average	Variance
American Indian or Alaskan Native	250	30.06	120.42
Asian or Pacific Islander	250	33.89	181.55
Black	250	28.79	120.14
White	250	33.39	195.65

Table 19 ANOVA – Offender Age

ANOVA – Offender Age						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4681.38	3	1560.46	10.104	1.47E-06	2.61
Within Groups	153821.53	996	154.44			
Total	158502.91	999				

In this case, the differences are more obvious in the box plot and are again reflected in the ANOVA. The F value, 10.10, is greater than the critical value of 2.61. This means the null hypothesis can again be rejected, and it can be concluded that not all of the population means are equal.

Conclusion

The general public and the media have a very specific view on crime and murder. Many people believe that they are more likely to be killed by a stranger than someone they know. Unfortunately, most people who are murdered are killed by someone they know and may love. However, that kind of murder is harder to sensualize than murder that happens within gangs, as part of a bar brawl or between strangers. This means that so much of what people believe about crime is not true. As shown by the descriptive statistics, males are most likely to be both the offender and victim of murder. White persons are just barely most likely to be the offender, with black persons following close behind, while black persons are just barely most likely to be the victim, with white persons following close behind. Both victims and offenders are young, typically in their 20's. Some of what the general population believes they know about crime is correct, while other bits are just made up, such as the idea that pre-1990 had any less murder than now. Knowing and understanding the trends behind murder make it more possible to devise tactics that will actually reduce the murder rate both at a community and law enforcement level. The inferential statistics showed that there are relationships been gender and weapon choice, gender and relationship, and age and relationship, but not between age and weapon choice. The ANOVA showed that there is a difference between the population means of different races for both victims and offender. This may indicate there is a racial element to murder motivation. This report makes no grand claims about murder motivations, but it does show that much of what is assumed about murder by the general population and the media may not be statistically true and that there are many fascinating trends that could be used and examined in efforts to decrease murder rates and solve open cases.