Denver Crime Lab Suicide Study

Makayla Cowles

1. Introduction

The Denver Crime Lab in Colorado and coroner's offices in Las Vegas, Nevada and Milwaukee, Wisconsin provided demographic data for suicide victims for these three cities over different periods of time. "Along with the demographic data, the coroner's offices have dried blood cards for each of these victims which can be used to genotype each of the victims' DNA" (3). The future goal is to test if there exist a genetic predisposition to suicide by using the DNA of these victims to perform a Genome Wide Association Study (GWAS) of suicide victims.

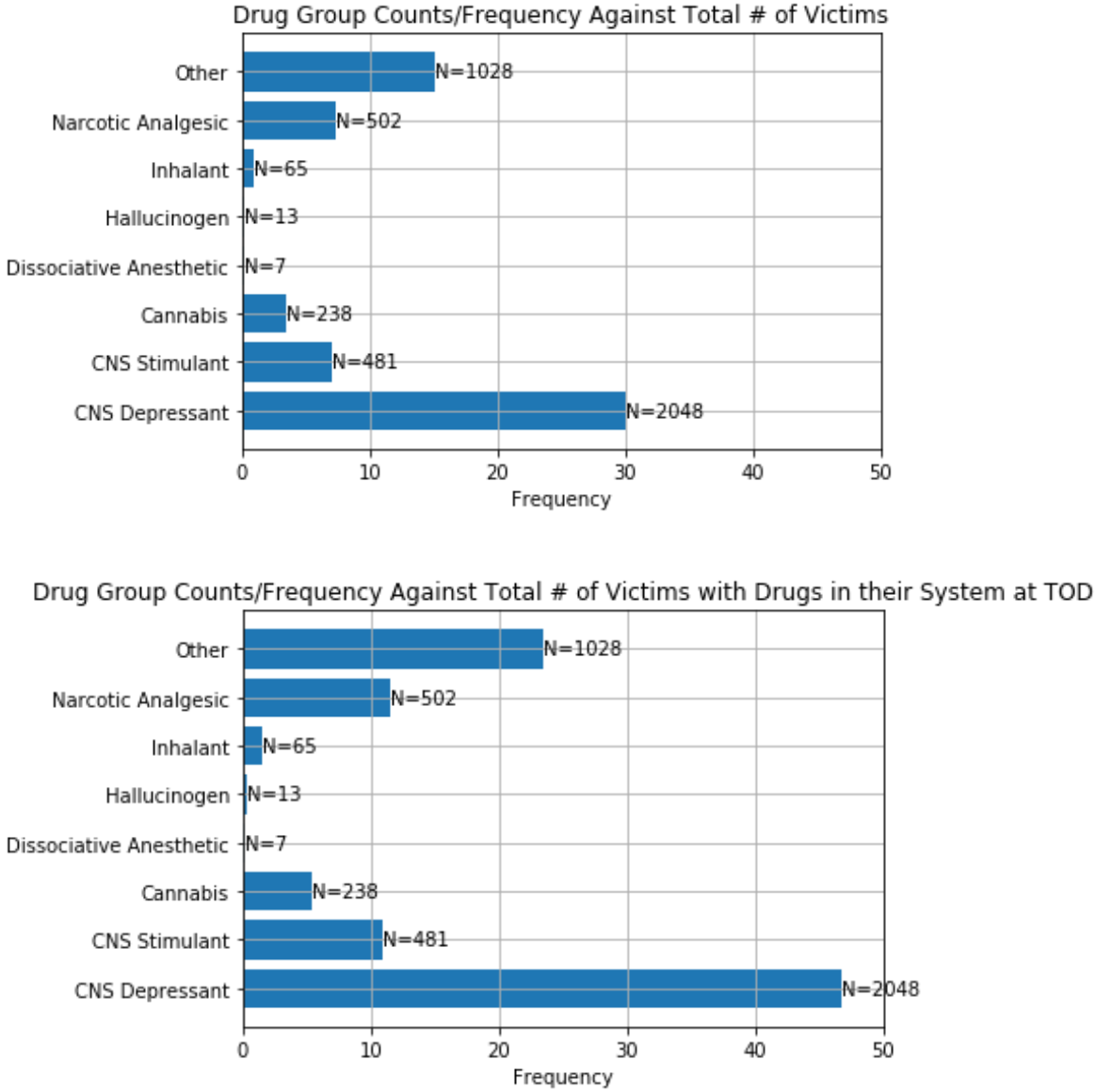
A previous analysis has been completed by Megan Duff to "identify relationships between demographic data and suicide" (3). Her results concluded that there exist relationships between sex and suicide and race and age in terms of suicide. To continue this work, toxic substance/drug data of the suicide victims has been added to the demographic data for additional analysis. The aim of this project is to perform exploratory analyses to identify relationships between the demographic data and the drug data. To perform exploratory analysis, the drug data was categorized into eight drug groups. The analyses conducted in this project tests for associations between sex, city, and race and drug groups. The secondary goal of this project is to provide a clean data set with a focus on the drug data to be used in future analyses. This is an observed study, and the population of interest is United States citizens.

2. Analysis and Results

There are a total 6,827 suicide victims observed in Denver, Las Vegas, and Milwaukee. Approximately, 41%, 2,808 victims, were identified with at least one toxic substance or drug in their system at time of death. Therefore, approx., 58.9% of victims were identified to not have a toxic substance or drug in their system at time of death. A total of 430 suicide victims were identified to have more than one toxic substance or drug in their system at time of death. There

exist 613 unique combinations of toxic substances found in victims at time of death that was reduced to the following eight drug groups: CNS depressants, CNS stimulants, cannabis, dissociative anesthetics, hallucinogens, inhalants, narcotic analgesics, and other. The total count for how often a drug group is observed in all suicide victims and just the victims identified with a toxic substance or drug in their system at time of death is shown in Figure 1.

Figure 1. Counts and frequencies of each eight drug groups observed in all suicide victims (top) and only victims with a drug in their system at time of death (bottom)

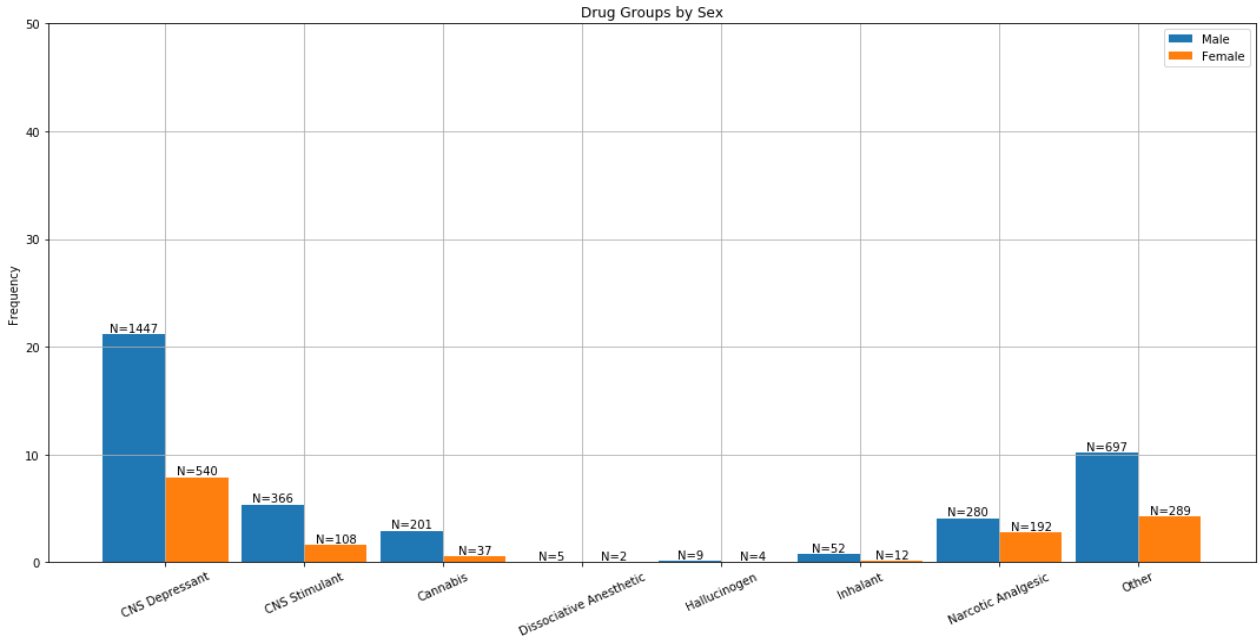


For exploratory analysis, counts were calculated for how often a drug group was observed within different demographics. Table 1 shows the observed counts of drug groups for sex in Denver and Las Vegas since the data for Milwaukee did not include sex. There are a total of 4,513 male suicide victims, and approximately 68% of male victims were identified to have at least one drug in their system at time of death. Additionally, there are 1,388 female suicide victims, and approximately 85% were identified to have at least one drug in their system at time of death. A permutation test of independence was performed to test for associations between drug group and sex for Denver and Las Vegas. The resulting p-value was 0.00001, showing that there exists enough evidence to reject the null hypothesis of no association.

Table 1. Total observed drug group counts for males and females in all Denver and Milwaukee

	CNS Depressant	*CNS Stimulant*	*Cannabis*	*Dissociative Anesthetic*	*Hallucinogen*	*Inhalant*	*Narcotic Analgesic*	*Other*	*Total*
Male	1,447	366	201	5	9	52	280	697	3,057 (67.7%)
Female	540	108	37	2	4	12	192	289	1,184 (85.3%)

Figure 2. Histogram of drug group frequency against all suicide victims and observed drug group counts for sex in Denver and Las Vegas

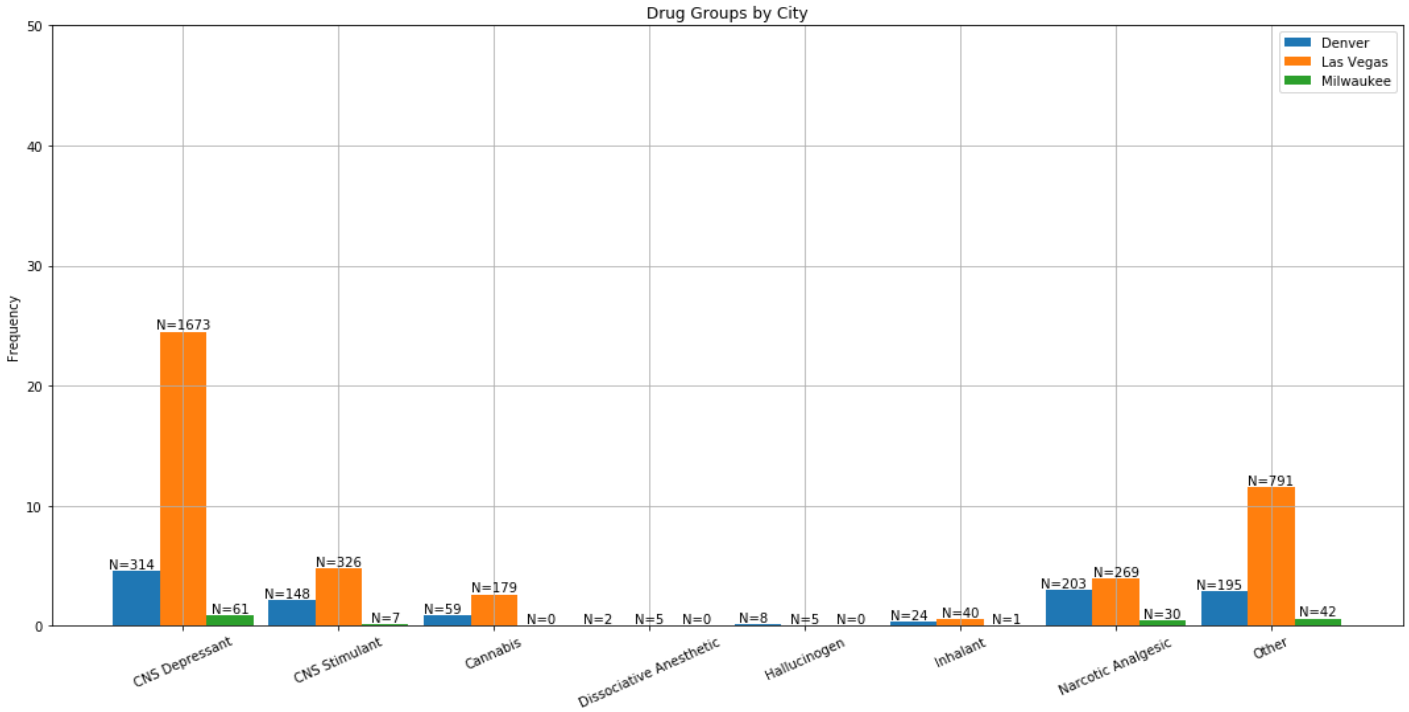


The total counts for how often a drug group is observed for Denver, Las Vegas, and Milwaukee is shown in Table 2. There are a total of 3,958 suicide victims in Las Vegas, 1,946 suicide victims in Denver, and 923 suicide victims in Milwaukee. Approximately, 83% of victims in Las Vegas, 49% of victims in Denver, and 15% of victims in Milwaukee were identified to have at least one drug in their system at time of death. A test for independence was performed between drug groups and city, with a resulting p-value of 0.00001.

Table 2. Total observed drug group counts for suicide victims in Denver, Las Vegas, and Milwaukee

	CNS Depressant	*CNS Stimulant*	*Cannabis*	*Dissociative Anesthetic*	*Hallucinogen*	*Inhalant*	*Narcotic Analgesic*	*Other*	*Total*
Las Vegas	1,673	326	179	5	5	40	269	791	3,288 (83.1%)
Denver	314	148	59	2	8	24	203	195	953 (49.0%)
Milwaukee	61	7	0	0	0	1	30	42	141 (15.3%)

Figure 3. Histogram of drug group frequency against all suicide victims and observed drug group counts for Denver, Las Vegas, and Milwaukee

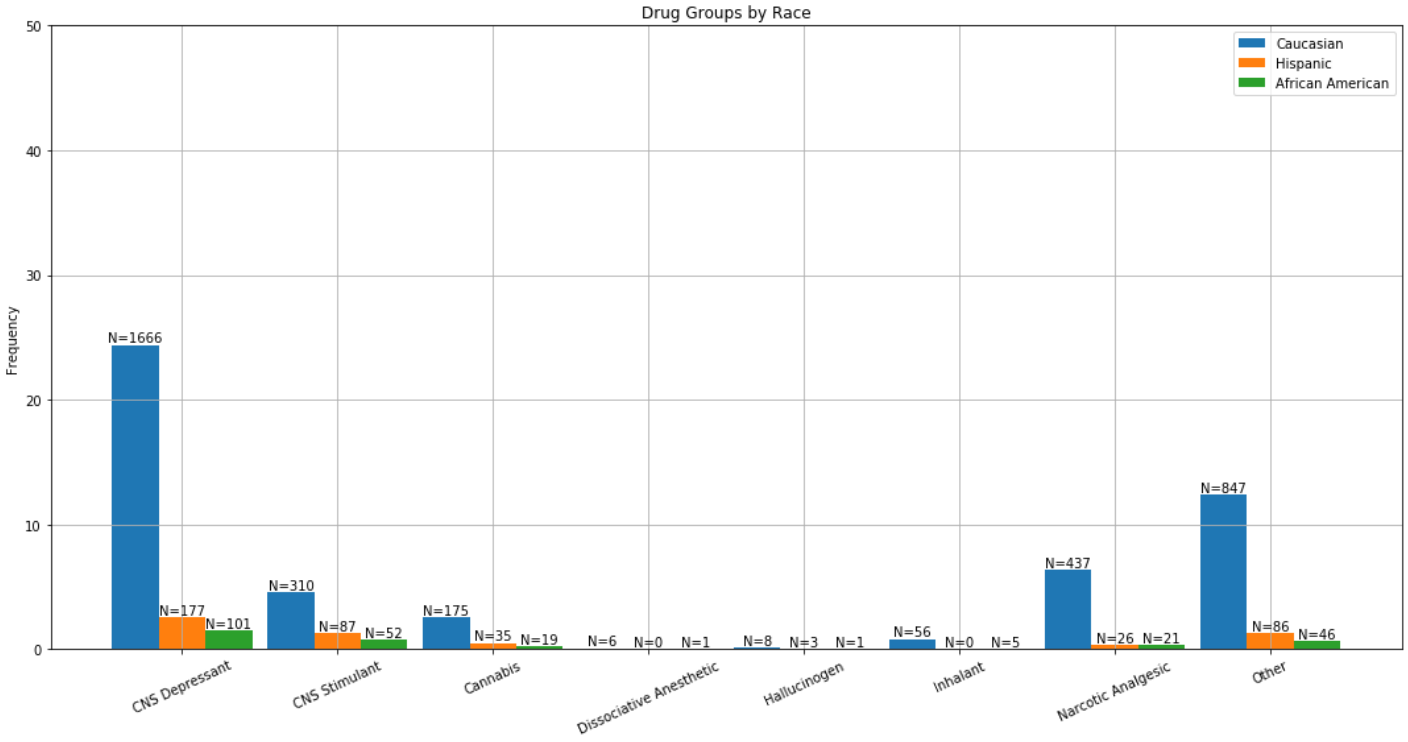


The total observed counts of drug groups identified in various races in all three cities are shown in Table 3. The races with the most drug data are Caucasian with approx. 66% of victims, Hispanic with approx. 58% of victims, and African American with approx. 50% of victims identified to have at least drug in their system at time of death. There are a total of 228 Asian/Pacific Islander suicide victims; however, none of these victims were identified with a drug in their system at time of death. Figure 4 shows the observed counts for only Caucasian, Hispanic, and African American since there is not enough drug data within the remaining races to include in the exploratory analysis. We performed a test for independence between drug group and the following races: Caucasians, Hispanics, and African Americans. The p-value calculated from this test was 0.00001.

Table 3. Total observed drug group counts for various races in Denver, Las Vegas, and Milwaukee

	CNS Depressant	*CNS Stimulant*	*Cannabis*	*Dissociative Anesthetic*	*Hallucinogen*	*Inhalant*	*Narcotic Analgesic*	*Other*	*Total*
Caucasian	1,666	310	175	6	8	56	437	847	3,505 (66.0%)
Eastern Indian	6	7	2	0	0	0	4	3	22 (81.5%)
Hispanic	177	87	35	0	3	0	26	86	414 (58.2%)
African American	101	52	19	1	1	5	21	46	246 (50.1%)
Asian/Pacific Islander	0	0	0	0	0	0	0	0	0 (0%)
Native American	3	1	0	0	0	0	0	0	4 (50.0%)
Multi-Racial	3	1	0	0	0	0	0	4	8 (44.4%)

Figure 4. Histogram of drug group frequency against all suicide victims and observed drug group counts for Caucasians, Hispanics, and African Americans in Denver, Las Vegas, and Milwaukee



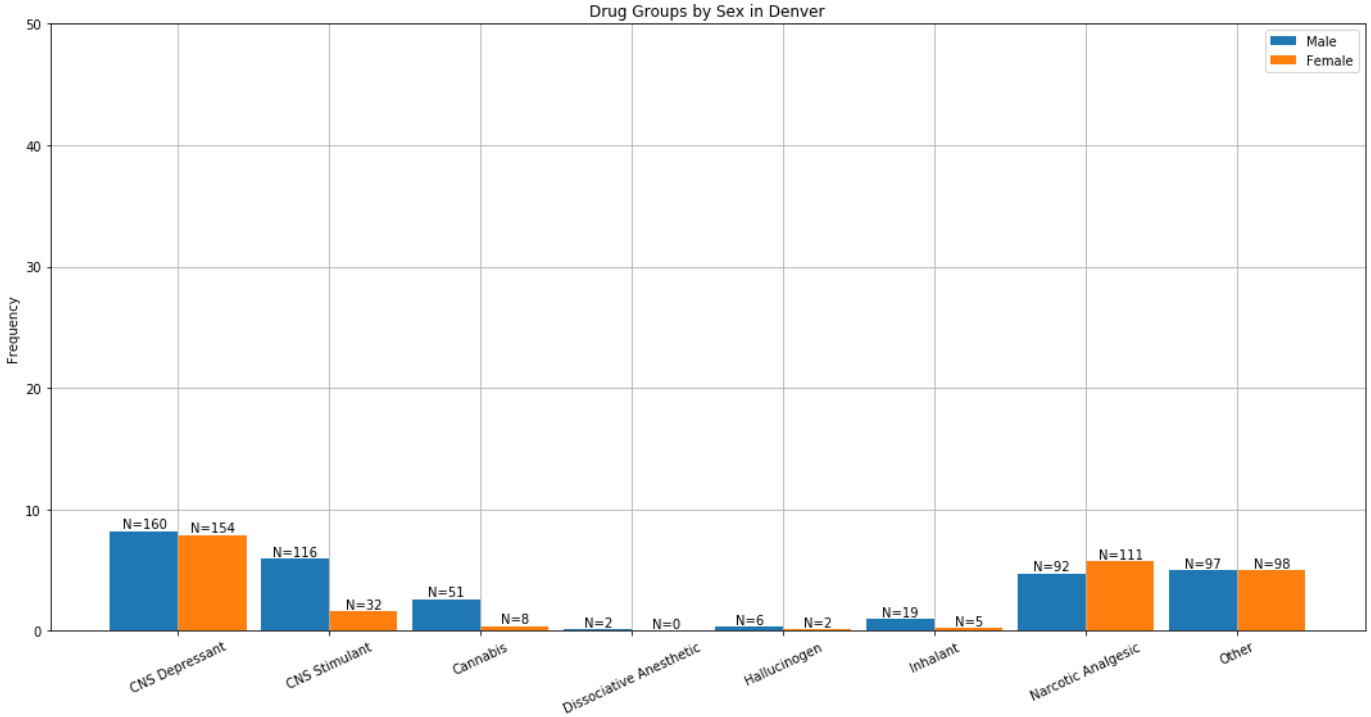
After calculating the total observed drug group counts for city, sex in Denver and Las Vegas, and race in all three cities, total observed drug groups were calculated for sex and race in each city separately. We first look only at data for Denver and calculate the total observed drug group counts for sex and race (Table 4). Approximately, 37% of males and 84% of females in Denver were identified with at least one drug in their system at time of death. Additionally, approx., 51% of Caucasians, 38% of Hispanics, and 46% of African Americans in Denver had at least one drug identified in their system at time of death. Tests for independence were performed to identify associations between drug group and both sex and race in Denver. For both tests, a p-value of 0.00001 was outputted.

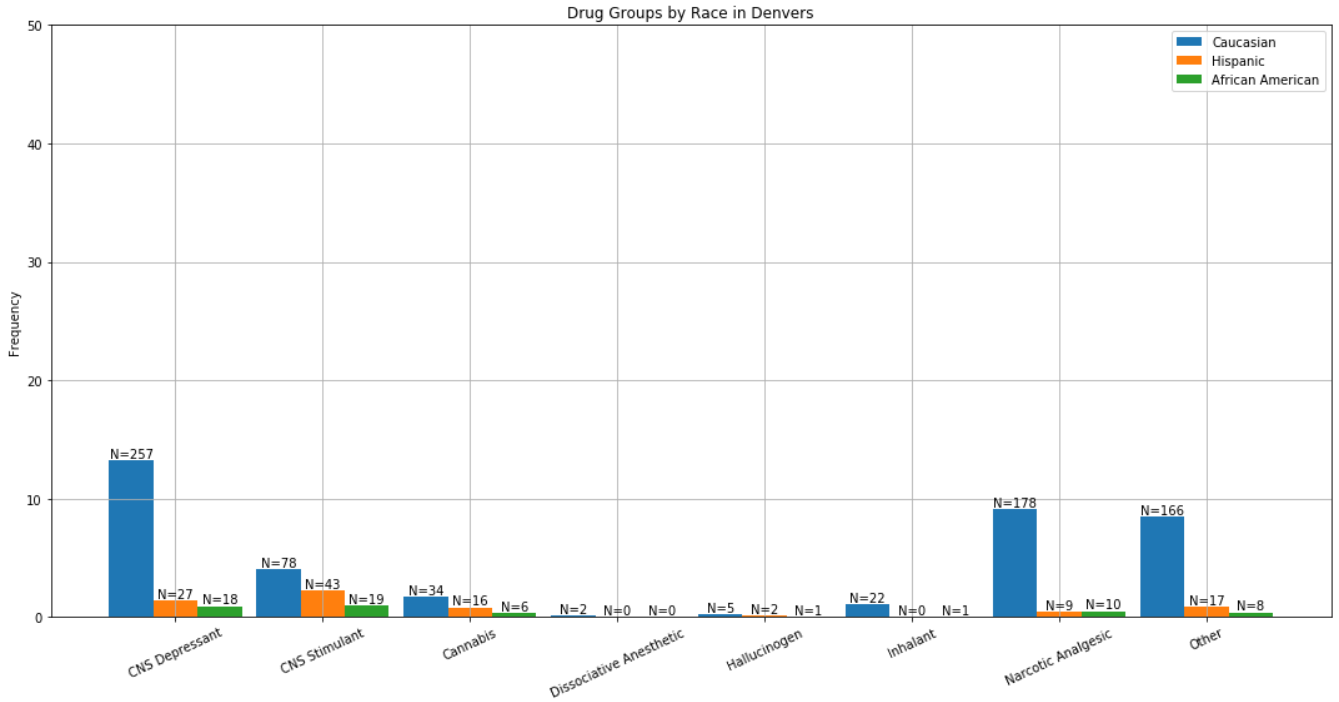
Table 4. Total observed drug group counts for sex (top) and Caucasians, Hispanics, and African Americans (bottom) for Denver

	CNS Depressant	*CNS Stimulant*	*Cannabis*	*Dissociative Anesthetic*	*Hallucinogen*	*Inhalant*	*Narcotic Analgesic*	*Other*	*Total*
Male	160	116	51	2	6	19	92	97	543 (37.3%)
Female	154	32	8	0	2	5	111	98	410 (83.5%)

	CNS Depressant	*CNS Stimulant*	*Cannabis*	*Dissociative Anesthetic*	*Hallucinogen*	*Inhalant*	*Narcotic Analgesic*	*Other*	*Total*
Caucasian	257	78	34	2	5	22	178	166	742 (51.0%)
Hispanic	27	43	16	0	2	0	9	17	114 (37.6%)
African American	18	19	6	0	1	1	10	8	63 (46.0%)

Figure 5. Histogram of drug group frequency against all suicide victims in Denver and observed drug group counts for male and females (top) and Caucasians, Hispanics, and African Americans (bottom) in Denver.





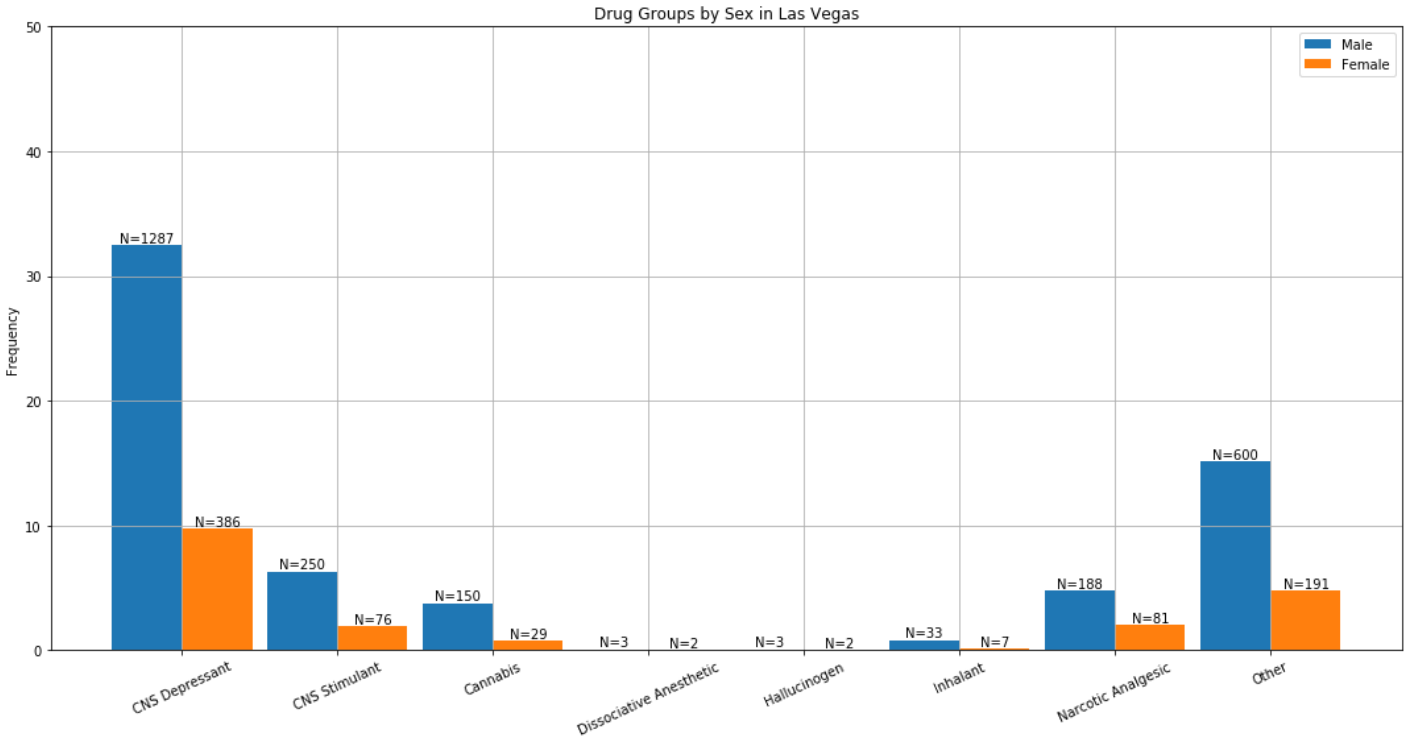
Next, we look at only data from Las Vegas, and calculate observed drug group counts for sex and race (Table 5). In Las Vegas, approximately, 82% of male suicide victims and 86% of female suicide victims were identified with at least one drug in their system at time of death. In addition, approximately 83% of Caucasians and Hispanics and 84% of African Americans were identified with at least one drug in their system at time of death. The last two tests of independence were performed to find associations between drug group and sex and race in Las Vegas. The p-value for independence between drug group and sex in Las Vegas was 0.00004; the p-value for independence between drug group and race in Las Vegas was 0.02265.

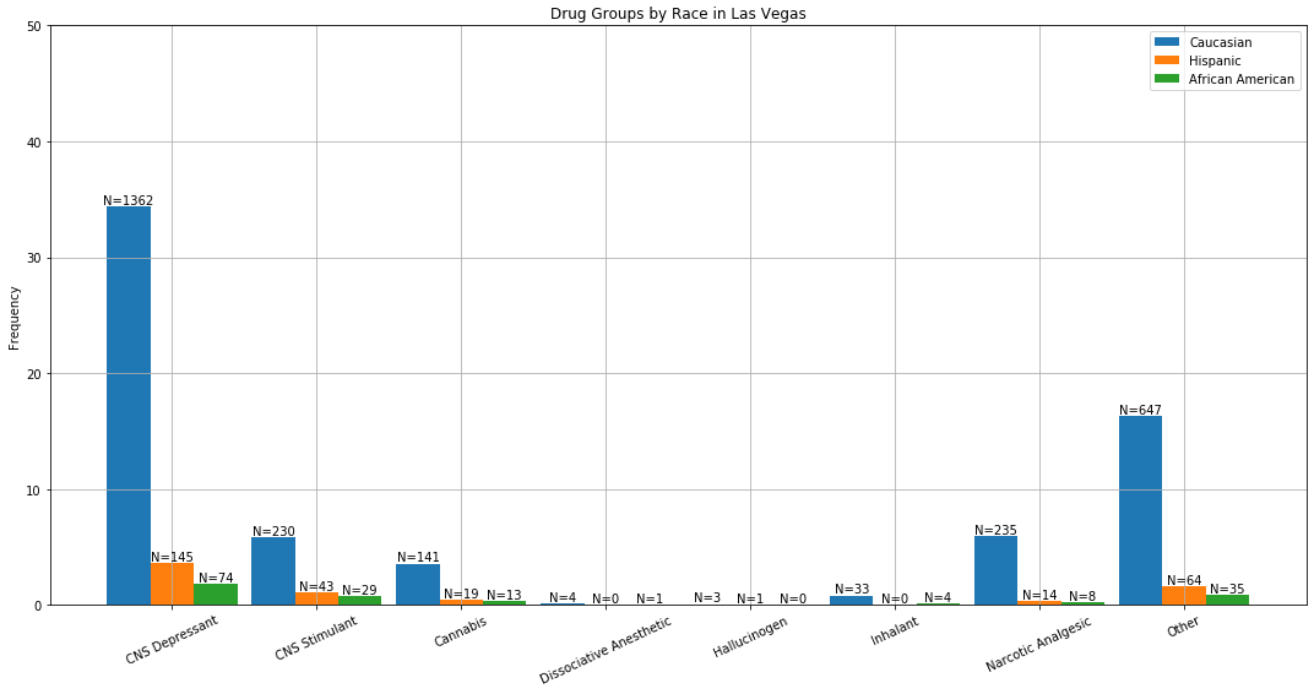
Table 5. Total observed drug group counts for sex (top) and Caucasians, Hispanics, and African Americans (bottom) for Las Vegas

	CNS Depressant	*CNS Stimulant*	*Cannabis*	*Dissociative Anesthetic*	*Hallucinogen*	*Inhalant*	*Narcotic Analgesic*	*Other*	*Total*
Male	1,287	250	150	3	3	33	188	600	2,514 (82.2%)
Female	386	76	29	2	2	7	81	191	774 (86.3%)

	CNS Depressant	*CNS Stimulant*	*Cannabis*	*Dissociative Anesthetic*	*Hallucinogen*	*Inhalant*	*Narcotic Analgesic*	*Other*	*Total*
Caucasian	1,362	230	141	4	3	33	235	647	2,655 (83.3%)
Hispanic	145	43	19	0	1	0	14	64	286 (83.3%)
African American	74	29	13	1	0	4	8	35	164 (84.1%)

Figure 6. Histogram of drug group frequency against all suicide victims in Las Vegas and observed drug group counts for male and females (top) and Caucasians, Hispanics, and African Americans (bottom) in Las Vegas.





Lastly, we calculate observed drug group counts for race in just Milwaukee (Table 6). Drug groups were identified in approx. 16% of Caucasians, 22% of Hispanics, and 12% of African Americans at time of death in Milwaukee.

Table 6. Total observed drug group counts for Caucasians, Hispanics, and African Americans for Milwaukee

	CNS Depressant	*CNS Stimulant*	*Cannabis*	*Dissociative Anesthetic*	*Hallucinogen*	*Inhalant*	*Narcotic Analgesic*	*Other*	*Total*
Caucasian	47	2	0	0	0	1	24	34	108 (16.2%)
Hispanic	5	1	0	0	0	0	3	5	14 (21.5%)
African American	9	4	0	0	0	0	3	3	19 (11.9%)

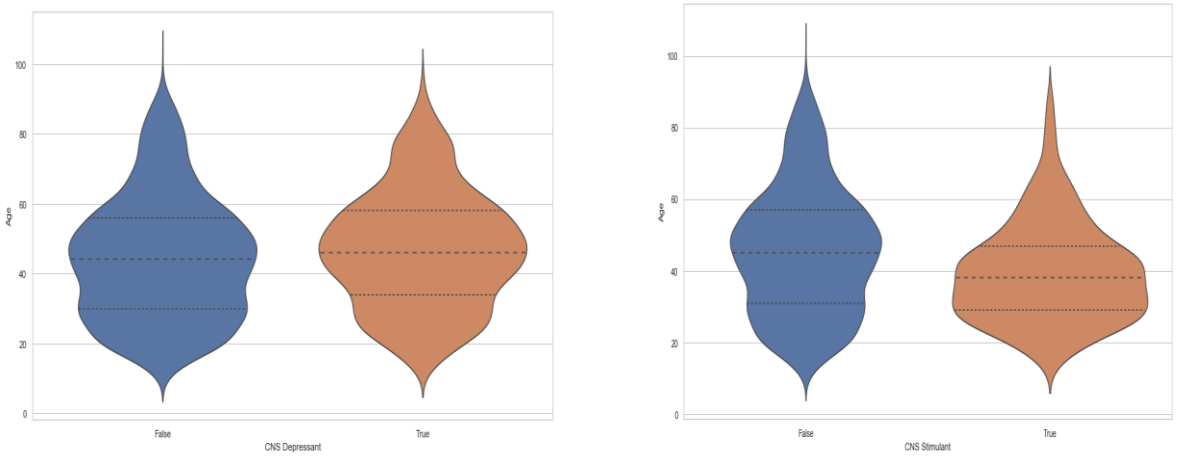
Data analysis was also conducted on age regarding the drug data. In Table 7, observed drug group counts were calculated for each pre-define age range for all suicide victims. The lowest proportion of victims identified with at least one drug in their system at time of death to all victims occurs in the 1-19 years age range. Approximately, 43% of suicide victims aged 1-19

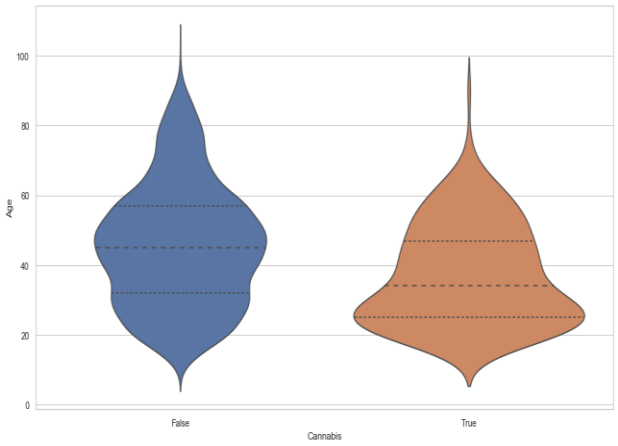
were identified with a drug in their system at time of death. Whereas the largest proportion occurs with suicide victims aged 60-69 with 74% of these victims identified with at least one drug in their system at time of death. Figure 7 shows the distributions of ages for each drug group; in addition to the median and first and third quartiles. The mean age and standard error are shown in Table 8 for each drug group.

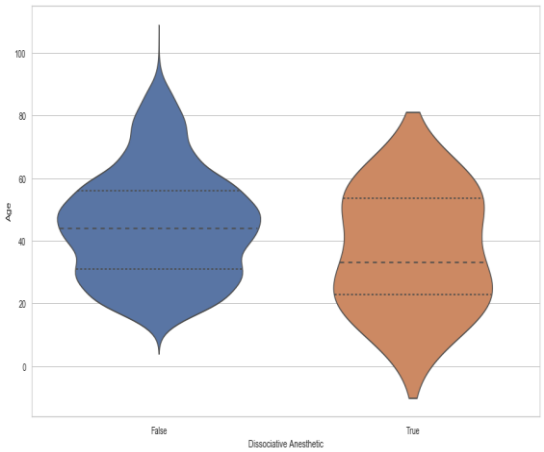
Table 7. Total observed drug group counts for different age ranges in Denver, Las Vegas, and Milwaukee

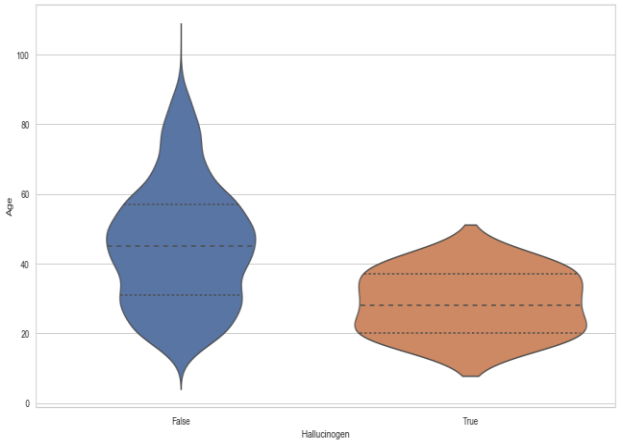
	CNS Depressant	*CNS Stimulant*	*Cannabis*	*Dissociative Anesthetic*	*Hallucinogen*	*Inhalant*	*Narcotic Analgesic*	*Other*	*Total*
1-19	70	17	18	1	2	2	8	38	156 (43.1%)
20-29	287	114	82	2	5	6	41	113	650 (58.2%)
30-39	343	132	45	1	4	12	88	156	781 (63.5%)
40-49	466	116	40	0	2	14	136	228	1,002 (68.3%)
50-59	435	58	33	3	0	15	108	215	867 (67.1%)
60-69	228	28	18	0	0	11	57	149	491 (74.2%)
70-79	149	9	1	0	0	4	41	65	269 (66.6%)
80-89	64	7	0	0	0	1	20	53	145 (56.4%)
90+	6	0	1	0	0	0	3	8	18 (60.0%)

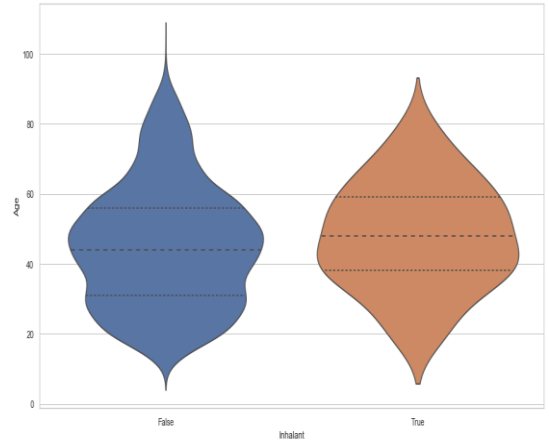
Figure 7. Violin plot showing the distribution, medians, and quartiles of ages for each drug group

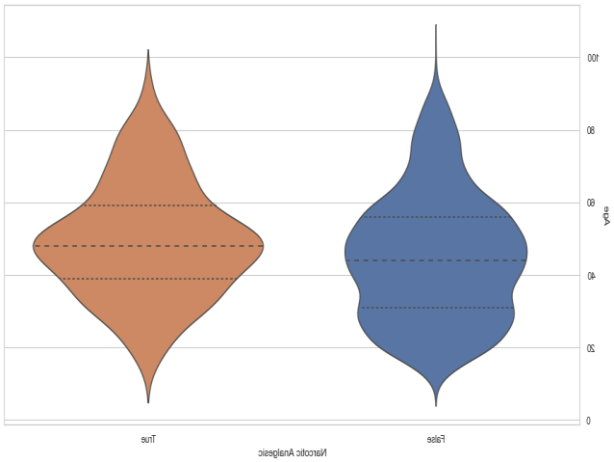



Age
0
20
40
60
80
100
False
True
Cannabis


Age
0
20
40
60
80
100
False
True
Dissociative Anesthetic


Age
0
20
40
60
80
100
False
True
Hallucinogen


Age
0
20
40
60
80
100
False
True
Inhalant


Age
0
20
40
60
80
100
False
True
Narcotic Analgesic

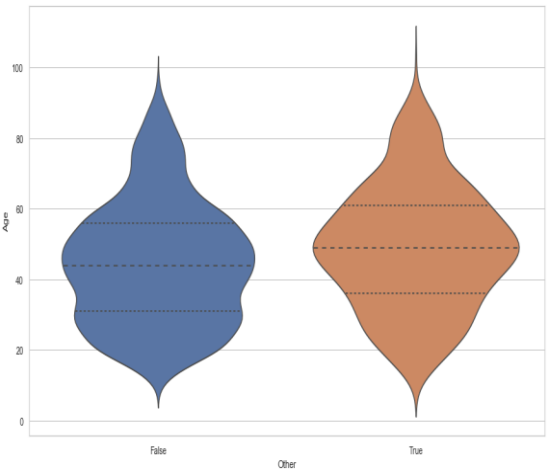

Age
0
20
40
60
80
100
False
True
Other

Table 8. Mean age and standard error for each drug group

	Mean Age	*Standard Error*
CNS Depressants	46.96	0.37
CNS Stimulants	39.56	0.63
Cannabis	36.83	0.93
Dissociative Anesthetics	36.71	6.74
Hallucinogens	28.85	2.36
Inhalants	47.89	1.89
Narcotic Analgesics	49.47	0.71
Other	49.03	0.55

3. Discussion

For the exploratory analysis, multiple associations were identified between demographic data and drug group data. The identified associations were between drug groups and each of the following: city (Las Vegas, Denver, Milwaukee), sex (male, female) in Denver and Las Vegas, race (Caucasian, Hispanic, African American) in Denver Las Vegas, and Milwaukee, sex (male, female) in Denver, race (Caucasian, Hispanic, African American) in Denver, sex (male, female) in Las Vegas, and race (Caucasian, Hispanic, African American) in Denver. The results from our permutation test for independence show very small p-values. Using a significance level of 0.05, enough evidence exists to reject the null hypothesis of no associations for each 7 permutation tests we run. Identifying each of these associations is promising regarding the future goal of performing a genome wide association study of suicide victims. However, given the very small p-values outputted from the results, it would increase confidence if the associations were performed again.

To continue this work, it is important to investigate other methods for organizing the toxic substance and drug data to avoid non-mutually exclusive data. One possible method is to duplicate each individual who is shown to have more than one drug in their system at time of death. For example, if an individual has cocaine, alcohol, and morphine found in their system at time of death, then this individual's data would be duplicated. All the data is the same for each duplication except for the drug data; in the toxic substance and drug column, the first duplication would say cocaine, the second duplication would say alcohol, and the third duplication would say morphine. Applying this method of organizing the data would significantly increase the data size, but the data would be mutually exclusive. Statistical methods can directly be applied to mutually exclusive data, avoiding the use of permutations. With mutually exclusive data, the analyses completed in this project could be ran again directly using the chi-square test of independence. Additional analyses could be more easily completed as well, such as an ANOVA test for difference in means among ages for each drug group.

The toxic substance and drug data is a valuable addition to the analysis of data regarding suicide victims. To enhance the results of the analysis performed in this project and future analyses, a more refined way of grouping the drugs should be explored. Drug recognition experts classify drugs into one of seven categories (1). These seven categories were used to group drugs in this analysis. However, it might be more effective to look at grouping the drugs into even more categories. At the very least, it is necessary for an expert to double check the categorizations used in this analysis.

In moving forward with this study, more data is essential. Given the data received, some demographics could not be included in the drug association analysis due to there not being enough observations. Regarding race, there were four racial groups that were not included in this

drug analysis due to small sample sizes related to drug data: Eastern Indian, Native American, Asian/Pacific Islander, Multi-Racial. Additionally, Milwaukee did not include sex in their data, so Milwaukee had to be left out of the exploratory analyses that tested for independence between drug groups and sex. There is a better chance of having more drug data for demographics left out of this analysis if data can be added from more cities, or if more observed years were included in the data from Denver, Las, Vegas, and Milwaukee,

4. Model and Methods

Harmonize/Clean Data

Jupyter Notebooks was used to harmonize and clean the data using pandas and numpy python packages. Columns of similar data between the three data sets were combined, e.g., Location of Death and Place of Death, and Cause of Death and Final Cause of Death. In addition to the toxic substance and drug data, these columns were cleaned: race, age, and marital status. The drug data for Denver was provided as five separate columns (Drug 1, Drug 2, Drug 3, Drug 4, Drug 5) of drug codes. A dictionary was created to convert drug codes to drug names and all five columns containing drug data for Denver were combined into one column (Drug). The newly created drug column was combined with a toxic substance column present in the data for Las Vegas. To identify the drug data for Milwaukee, drug related injuries were searched within the injury type column, then the column containing drug names (Cause A) was combined with the toxic substance column that contains both data for Las Vegas and Denver. To clean the toxic substance column, extra words that were not drug names were removed, extra spaces and characters were removed, and drug names were separated by semicolons. A new column was created for the clean drug data, and in a different column drug names were replaced by their corresponding drug group. Eight Boolean columns were created for each drug group, and an

additional Boolean column was created to indicate whether the observed individual did or did not have a drug in their system at time of death. The location of death and cause of death were not cleaned, but still included in the final data set.

Drug Groups

A total of eight drug groups were used to categorize each drug identified in the data. International Association of Chiefs of Police website defines seven drug categories: central nervous system (CNS) depressants, CNS stimulants, hallucinogens, dissociative anesthetics, narcotic analgesics, inhalants, and cannabis (1). The categorization of each drug was based off the descriptions and examples provided by the International Association of Chiefs of Police website. An "other" category was created for drugs that did not clearly fit into the pre-define seven drug categories. The following were examples of drugs that were categorized into each eight drug groups: CNS depressants (alcohol, anti-depressants, anti-anxiety, sleep inducers), CNS stimulants (cocaine, nicotine, methamphetamine), hallucinogens (LSD, ecstasy), dissociative anesthetics (PCP, ketamine), narcotic analgesics (opioids, codeine, morphine, heroin), inhalants (gasoline, hair spray, paint thinner, anesthetic gases), cannabis (THC, cannabinoids, synthetics), and other (anticonvulsants, antibiotics, anticoagulants, allergy medications). Also included in "other" were various drug codes not included in the drug code dictionary used to convert drug codes to drug names.

Permutation Test for Independence of Two Variables

The drug data was organized into 8 Boolean variables for each pre-determined drug group. Due to this organization of the drug panel data, and the event that an individual can be identified with more than one drug from multiple drug groups, we do not have mutually exclusive events. Non-

mutually exclusive events means that the events can occur at the same time. To account for non-mutually exclusive events, we run a permutation test.

Our exploratory analysis includes testing for associations between two variables (demographic variables and drug group variable). The variables we consider testing associations with drug group are sex, race, and city. All variables are categorical; hence we use a chi-square statistic for our permutation tests.

To run a permutation test for independence using a chi-square statistic, we first create a chi-square contingency table with the observe counts for each variable. We calculate the expected counts for each cell, and only keep the variable with an expected cell count larger than 5 (2). The chi-squared test statistic is calculated on the remaining variables and used to calculate the p-value from the permutations. Random permutations are applied to the column containing demographic data, and a chi-square statistic is calculated for the permutated data. This process is repeated 99,999 for precision. Finally, we calculate the p-value as the fraction of times the random chi-square statistics exceed the original.

Note, observations that had missing data for one or more variables were excluded from the analysis.

5. References

1. "7 Drug Categories." *International Association of Chiefs of Police*, https://www.theiacp.org/7-drug-categories.
2. Chihara, L.M. and Hesterberg, T.C. (2018). Introduction to Hypothesis Testing. In Mathematical Statistics with Resampling and R (eds L.M. Chihara and T.C. Hesterberg). https://doi.org/10.1002/9781119505969.ch3
3. Duff, Megan. (2020). "Denver Crime Lab Client Report".

6. Appendix

https://github.com/mcowles33/Denver-Crime-Lab-Suicide-Study.git