

Fentanyl: A Study into Overdose Deaths

Martin Mezera, Agyei Sraha, Aryan Arora, Gavin Pilato, and Spencer Epley

DTSC 4302: Data Science for Social Good Part B

05/02/2024

Contents

Introduction	3
Literature Review & Background	4
Methodology	8
Analysis	13
Discussion & Conclusion.....	20
References	26

Introduction

Introduced in 1960 as a potent intravenous anesthetic, fentanyl quickly became the most efficient and fastest-acting opioid available [1]. Unfortunately, as of 2021, it has earned a reputation of being one of America's deadliest substances. The data published from UCLA Health reveals that fatalities resulting from fentanyl overdoses have escalated from a mere 235 deaths in 2010 to a total of 34,429 fatalities in 2021 alone [2].

Initially intended for managing severe cancer and surgical pain, fentanyl rapidly became the predominant synthetic opioid in clinical practice, boasting potency levels 50-100 times greater than morphine and 30-50 times stronger than heroin [1]. When rigorous policies were enforced on prescription opioids like Oxycontin, and law enforcement tightened regulations on heroin, fentanyl emerged as an alternative. In 2013, drug dealers began to mix fentanyl with other substances to increase profits without consumer awareness. Factors such as easy availability, low cost, high potency, and unintentional intake create a perfect storm for fentanyl overdose rates.

Fentanyl operates similarly to other opioids, targeting specific receptor systems in the brain that regulate emotions, pain perception, speech, and reward. Like heroin, common side effects include euphoria, confusion, respiratory depression, drowsiness, and nausea. More severe side effects include hallucinations, delirium, hypotension, coma, and even death [3]. Crucially, fentanyl's addictive nature stems from its interaction with reward receptors, worsening dopamine neurotransmission and addiction [4]. Once the brain adapts to fentanyl-induced dopamine surges, quitting fentanyl becomes challenging, with withdrawal often resulting in despair, depression, and suicidal ideation.

While existing literature extensively covers the fentanyl crisis, our focus is distinct. The goal of this study is to build an up-to-date profile of those at risk for fentanyl overdoses. We believe that by building this profile, we can aid those most likely to overdose. Additionally, we will provide local governments and community outreach programs with this data, so they can distribute resources based on these profiles. To accomplish this, our study analyzes the records in the Mortality Multiple Cause of Death database managed by the CDC, consisting of data from 2021. Our hypothesis posits that a machine learning approach will effectively determine the most crucial factors towards generating a model of populations vulnerable to fentanyl overdose. Furthermore, we will use this profile and features to create tailored prevention and harm reduction techniques.

Literature Review & Background

While reviewing the existing literature, efforts were made to comprehensively study the history of drug policy in America to better understand the fentanyl crisis today. On June 17th, 1971, then-President Richard Nixon dramatically shifted American Drug policy by declaring drugs to be "public enemy number one." This statement was influenced by concerns of heroin use among veterans returning from Vietnam and service members still deployed. To unify federal drug policies into a single entity, Nixon inaugurated the DEA (Drug Enforcement Agency) in 1973. However, it wasn't until the Reagan administration that the War on Drugs truly gained momentum and had substantial impact across society. Prior to Reagan's presidency, there was a focus on using medical interventions and rehabilitation programs to help drug users. However, Reagan's admin began to strictly enforce drug laws, resulting in non-violent drug related imprisonments skyrocketing from 50,000 in 1980 to 400,000 in 1997 [5].

While Reagan is no longer in office, the strict enforcement of non-violent drug offenses continues. With a rate of 629 incarcerations per 100,000 people as of 2021, the United States' incarceration rate ranked as the highest globally and is well above other world superpowers such as Russia and China, whose incarceration rates were 326 per 100,000 people and 119 per 100,000 people, respectively [6]. Although our study isn't concentrated on imprisonment, such conditions introduce concerns surrounding topics including repeat offenders and overdose upon release from prison. Studies have demonstrated varying rates of overdose deaths in recently released inmates, with most averaging a likelihood of 7 to 18 times the risk of overdose in their respective populations [7]. Additionally, a study from 2000 to 2015 in the American Journal of Public Health found that in the first 2 weeks after release, North Carolina inmates booked on any charge have an opioid overdose rate 40 times greater than that of the general NC population [8]. These findings suggest that imprisonment, while arguably beneficial for community safety, may worsen issues for drug users themselves. This study does not necessarily support decriminalization or anti-incarceration but emphasizes imperfections in the current system which do not deliver positive results for those exiting prison.

The prevalence of fentanyl in the U.S. forces examination of laws and policies that might have unintentionally contributed to this crisis. Drug legislation in America has traditionally centered around limiting substance use, however, drugs that are noted for their availability and potency, like fentanyl, have surfaced as alternatives and challenged these regulations. Work done by the University of British Columbia's Medicine Department claim that the surge of fentanyl overdoses can be linked to attempts to reduce opioid and heroin availability. They believe that these efforts have not addressed the needs of individuals struggling with dependence but rather, have led them to seek out more potent substances [9].

It is noteworthy how overwhelming this issue has become as the rate of overdose related deaths has doubled in the past 20 years. In 2001, the drug overdose death rate was 6.2 people per 100,000. That rate has increased to 14.7 in 2021 [10,11]. This journal also showed an increase in overdose deaths among all genders, ethnicities, regions, and age groups (for age groups above 25). While this source highlights the impact of the fentanyl crisis, it is imperative to identify factors, such as marital status, education, and employment status, that result in even greater risk. Despite these shortcomings, this source illustrates that every community is experiencing repercussions as the severity of the fentanyl crisis becomes worse.

In 2014, opioids contributed to 61% of all drug overdose deaths. As opioids became more prevalent, this rate rose to 75.4% in 2021. Interestingly, of opioid-related deaths in 2021, 88% are attributed to synthetic opioids, including fentanyl [12]. While the fentanyl crisis has been exacerbated partly due to policy decisions, it is beneficial to explore other factors that fuel the epidemic. David Kearns at American University has named the “three waves” of the opioid epidemic, which has now become the fentanyl epidemic. Beginning in the 90s, the first wave involved prescription pills like Oxycontin. The second wave started around 2010, where there was a shift to heroin because of policy reform that made prescription pills expensive and difficult to obtain. A legislative crackdown on heroin marked the end of the second wave. This resulted in the third wave, which is occurring now, with fentanyl as the primary opioid. Additionally, during this wave it has been found that fentanyl is much cheaper to produce than the opioids from earlier waves [13].

Furthermore, Doctors Richard Frank and Harold Pollack have highlighted the economic incentives behind fentanyl. They found that manufacturing heroin costs \$65,000 per kg, while illicit fentanyl costs a mere \$3,500 per kg. This price difference incentivizes drug dealers to

manufacture illicit fentanyl without investing in proper equipment or knowledge [14]. Due to illicit fentanyl being significantly cheaper than heroin, dealers often mix it with other substances to increase potency. While pure fentanyl is relatively safe, illicit fentanyl poses significant danger due to inadequate production and lack of quality control. Thus, even small doses of fentanyl mixed into other drugs can be fatal to users. To compound this issue, users often are not aware that they are consuming fentanyl as drug dealers rarely advertise this.

Harm and usage reduction are popular methods to combat the opioid crisis. Use reduction strives to lower rates of illegal drug use, while harm reduction attempts to decrease negative repercussions associated with drug usage. Two common harm reduction strategies are Naloxone and supervised consumption facilities. Naloxone, commonly known as Narcan, is an antidote for overdose that can reverse the effects of opioids. Supervised consumption facilities are a way users can consume drugs more safely. These facilities provide users with pure formulations and supervise users during consumption [14]. This can help combat the fentanyl crisis, as users can be guaranteed a drug without mixed-in fentanyl. Supervised consumption facilities have been made widely available to the public at no cost.

During our research, it was discovered that a major limitation of existing literature on fentanyl is a lack of up-to-date information. Particularly, studies that built and used profiles were outdated, using data from 2016, and only used about half as many features as we found usable within our data set. We intend to address this limitation by building a modern profile with data from 2021, as well as examining how this crisis has evolved over recent years. This updated profile will help identify the combination of factors that currently contribute to fentanyl overdoses. Understanding these factors will equip us to propose methods to reduce community

harm and usage. Moreover, setting up a profile can help us find the most vulnerable communities and generate tailored prevention and harm-reduction strategies.

Methodology

In our study, data was used from the 2021 Mortality Multiple Cause file published by the Centers for Disease Control (CDC), which holds 3.5 million reported deaths in the United States. The data contains variables such as marital status, education level, age, occupation, and manner of death. This data will allow us to build a profile and determine the most significant variables contributing to fentanyl overdoses. While many profiles have been built using the Mortality Multiple Cause files, none have been formed in recent years. To address this, updated profiles will be built to better analyze the variables contributing to this crisis. While researching overdose deaths, it was noticed that most literature focused on drug overdose at large, rather than fentanyl specifically. This study will bridge that gap by focusing solely on the fentanyl crisis.

Since our dataset contains many entries unrelated to our research, all inapplicable records were carefully filtered out, resulting in a dataset of 72,236 fentanyl deaths. Additionally, fentanyl overdose deaths were split into three categories that encompass the manner of death: accidental, suicide, and undetermined intent. Given that our data frame consists of 93% accidental deaths, our main goal is to determine which factors most contribute to accidental fentanyl overdoses. To limit ambiguity and focus on whether the user died from accidental drug poisoning or not, suicide and undetermined intent were grouped into “Others”. Using this data will identify which communities are most likely to overdose on fentanyl accidentally, inform them of the dangers of illicit fentanyl properly, and help them seek proper treatments.

The data frame consisted of 30 different ethnicities, including White, Black; individuals that were of multiple races, i.e., Black and Asian, Black, Asian and White, etc.; individuals that

were of Asian descent, i.e., Japanese, Chinese, etc.; individuals that were Islanders, or others, i.e., Guamanian, Hawaiian, etc. All individuals that were of multiple races, Asian descent, and Islanders were binned in specific groups: Mixed, Asian, and Other/Islanders, respectively, to reduce variability and facilitate a more concise analysis.

The contents of the age variable were also binned to reduce complexity and improve representation. The original data frame consisted of ages ranging from 1 to 98; these have been divided into six groups: 18 and under, 19 to 30, 31 to 40, 41 to 50, 51 to 65, and 65 and up. These bins allow us to explore the many distinct stages of human life, including youth, early adulthood, adulthood, and the elderly.

Various machine learning methods were implemented alongside data analysis tools to examine the factors contributing to fentanyl overdose and make predictions accordingly. Both an unsupervised and supervised algorithm were used, clustering and extreme gradient boosting, respectively. Clustering algorithms are optimized to place data into clusters to identify similar groups and create profiles. Given that our dataset consists primarily of categorical features, the lack of a metric for similarity among categorical data presents a challenge for clustering [15]. To address this issue, we decided to use a technique known as k-mode clustering. Unlike the popular k-means clustering that uses Euclidean distance to group numerical data, k-mode clustering works by identifying the modes, or most frequent values, within the clusters to choose its centroid [16]. Further, it was important to perform hyperparameter tuning to optimize our clusters by using the lowest cost parameters. The ‘cost’ attribute in k-mode clustering refers to the sum of the dissimilarities of all points to their respective cluster modes and is used to examine the quality of the clusters [17]. In turn, when the cost is minimized, the data points within the cluster are closest to the modes.

Parameter tuning was performed for the method of initialization (init): the number of times the algorithm will run with different centroid seeds (n_init) and the maximum number of iterations (max_iter) [17]. To determine K, or the number of clusters (n_clusters), the elbow method was used. The optimal K was chosen as the value where the reduction in cost began to level out [17]. The following parameters were optimal for our clusters: ‘Huang’ for init, which assigns the most frequent categories equally to the K initial cluster centers [18], ‘12’ for n_init, ‘100’ for max_iter, and ‘4’ for n_clusters.

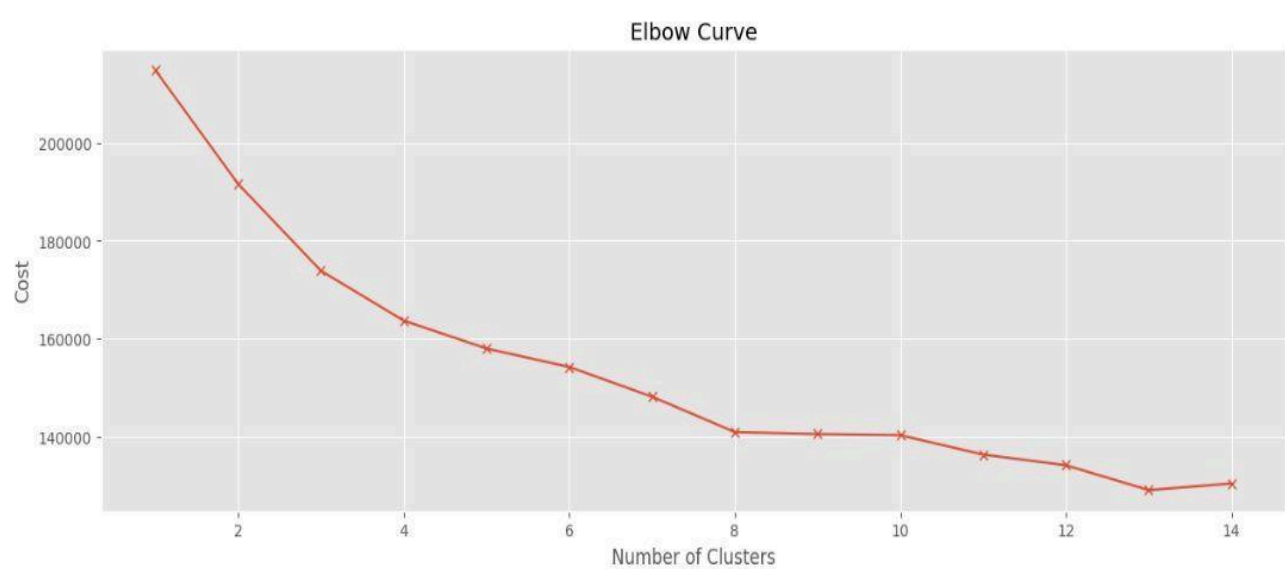


Figure 1- Elbow Curve Showing Optimal Number of Clusters

Once the optimal clusters were found, clustering was performed three separate ways: (1) using every feature excluding the ‘Death’ column, (2) using every feature including the ‘Death’ column, and (3) using every feature including the ‘Death’ column where ‘Other’ deaths were oversampled. Since 93% of the deaths were accidental, clusters that used the ‘Death’ column were skewed towards accidental, resulting in clusters with a cost of 163,772. When performing clustering where we oversampled ‘Other’ deaths to achieve a 50/50 split, the cost significantly increased to 356,375 due to an increase in the size of the data frame. The results were heavily

biased and did not represent actual patterns within the dataset, while potentially increasing the risk of overfitting. When performing clustering without the 'Death' column, the cost dropped to 161,943 and the clusters began to represent the data frame more accurately. We then calculated the ratio of the cost to the number of rows to compare and choose the method with the lowest cost, regardless of size. The ratios were found to be 2.24, 2.27, and 2.62 for methods 1, 2, and 3, respectively, and therefore K-modes clustering without the 'Death' column was chosen as the optimal method.

Once the profiles were built, the data frames were split, and models were built on each. Clusters 1, 2, 3 and 4 had a size of 32,040, 12,255, 12,899, and 15,029, respectively. Since all features in the data are categorical and most machine learning algorithms are designed to understand numerical data, we performed one-hot encoding to prepare our data for modeling. One-hot encoding converts a categorical variable into a binary column, where a '1' represents the presence of the variable, and a '0' represents absence.

However, before modeling began, a common trend was observed among all four clusters: a discrepancy between the number of 'Accidental Drug Poisoning' deaths and 'Other' deaths. Even though our primary focus was predicting 'Accidental Drug Poisoning', the models did not have enough data on 'Other' deaths to train effectively. To address this and create a balanced model, 'Other' deaths were oversampled by 20% to achieve an 80/20 split for 'Accidental Drug Poisoning' and 'Other', respectively. Various supervised learning algorithms were performed, including decision tree, random forest, naive bayes, and xgboost to predict the manner of overdose based on the four different profiles. Xgboost, even with default parameters, consistently yielded the highest accuracy metrics for all four clusters. Thus, xgboost was chosen as our sole algorithm.

Once our model was established, gridsearchcv was used to perform parameter tuning for the following parameters: the number of trees (n_estimators), max depth of the tree (max_depth), step size (learning_rate), subsample ratios of the training instances (subsample), the different columns (colsample_bytree) and minimum sum of instance weight (min_child_weight) [19]. Given that each cluster was a distinct size, parameters varied and changed, accordingly. The following parameters were the most optimal for XGBoost modeling:

	Cluster #1	Cluster #2	Cluster #3	Cluster #4
<u>Colsample_bytree:</u>	0.8	0.7	0.9	0.9
<u>Learning_rate:</u>	0.1	0.1	0.1	0.1
<u>Max_depth:</u>	9	9	9	7
<u>Min_child_weight</u>	1	1	1	1
<u>N_estimators:</u>	500	500	500	400
<u>Subsample:</u>	0.8	0.9	0.8	0.8

Figure 2- Key cluster parameters

Additionally, xgboost was used for feature importance to analyze how individual features affect the likelihood of overdose. Understanding how specific features contribute to overdose for a given profile can help distribute vital information to build proper use and harm-reduction strategies, such as to non-profit organizations or local governments.

Analysis

After the data was filtered to only include fentanyl overdose deaths, exploratory analysis began to gain better insight on the data. Type of death was the first attribute analyzed. In the data, there were two types of death: ‘Accidental Drug Poisoning’ and ‘Other’. The ‘other’ deaths represented deaths that were either unknown, undermined, or suicide. This analysis ended up reinforcing that accidental drug poisoning was the leading cause of death by far, which aligned with findings in relevant literature.

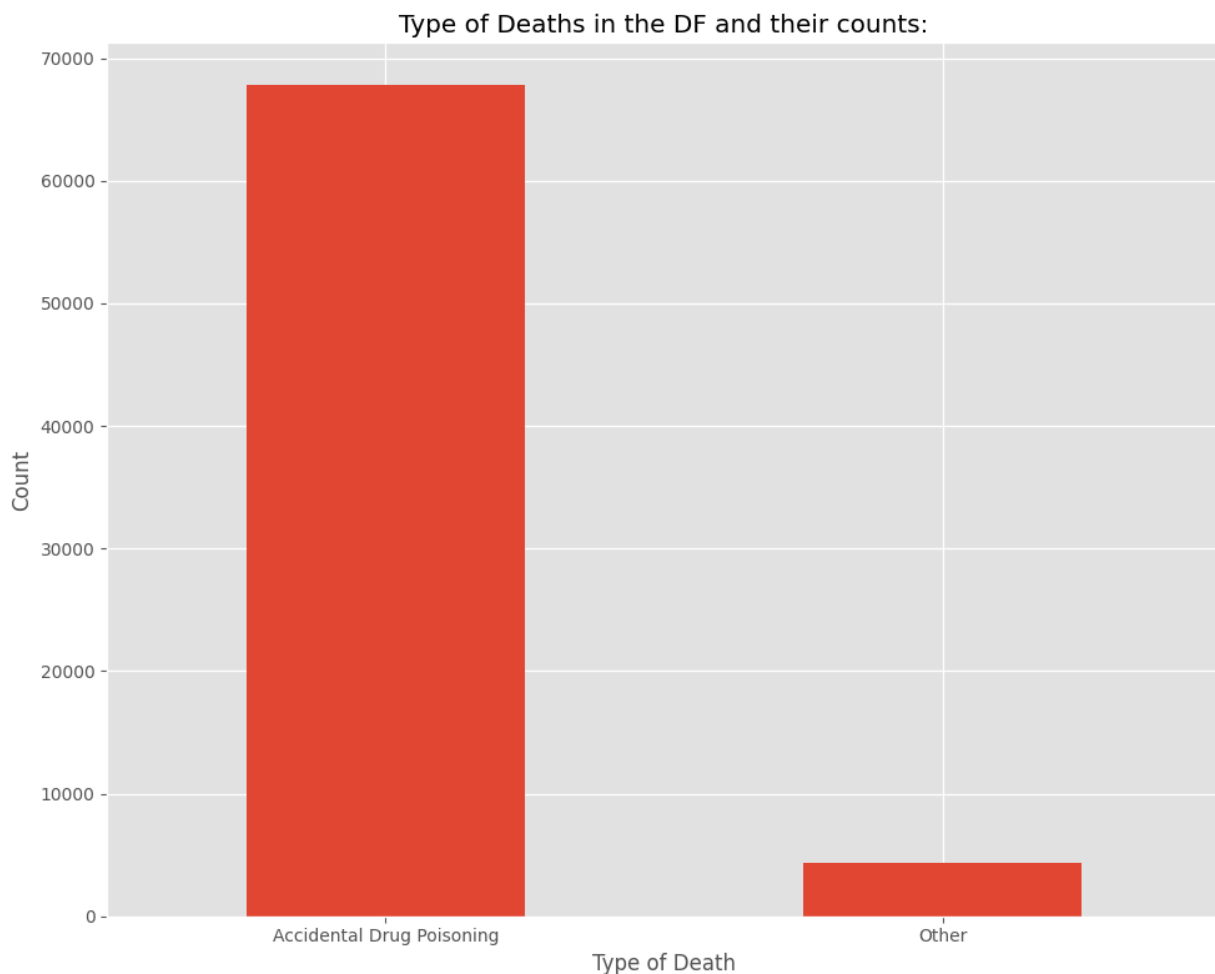


Figure 3- Graph shows count of death types

In our literature review, the leading theory as to why the fentanyl epidemic has reached unprecedented proportions is because of the contamination of other drugs with illicit fentanyl. This means the user has no idea they are ingesting fentanyl, and the number of accidental drug poisoning deaths reinforces this hypothesis. Also, it was immediately noticed how many overdose deaths involved individuals with a high school diploma or less education. According to the US Census Bureau, approximately 37.9% of Americans either haven't completed high school or at most have a high school diploma, compared to around 68.71% of fentanyl overdose victims [12,20]. Furthermore, it was found that while around 72% of all fentanyl overdose victims were

men, the US general population is around 49% men.

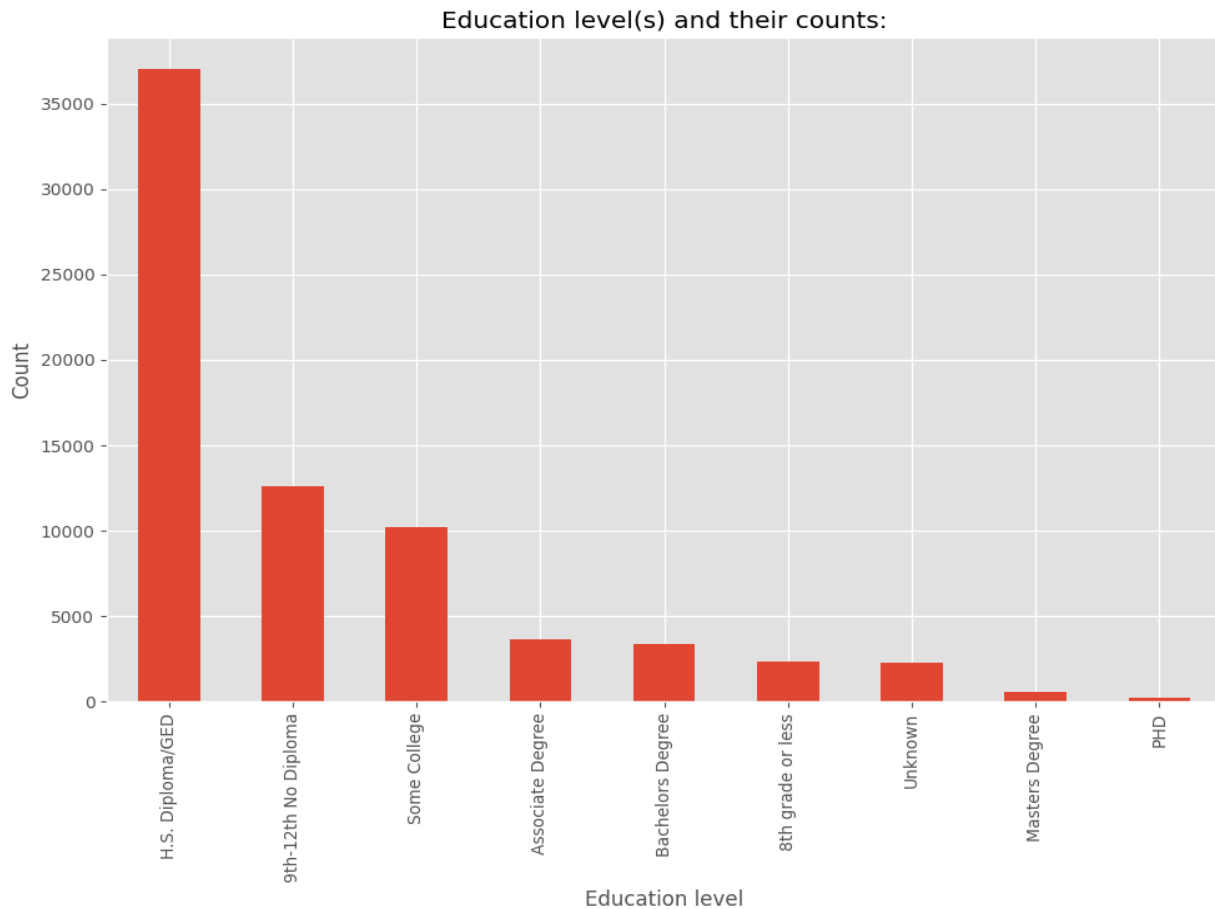


Figure 4- Education level of fentanyl overdose victims

Understanding these discrepancies can aid in more effective analysis of the profiles created. Since we found that the population of fentanyl overdose victims holds a demographic that is vastly different than the general US population, tailored harm and use reduction strategies can begin to be formulated based on this difference.

At this point, construction began on our profiles using k-mode clustering. The death column was excluded, and profiles were made of the remaining features. The death column was dropped because it was so heavily skewed towards 'Accidental Drug Poisoning' and we wanted

to avoid any potential bias. As stated earlier, when using the elbow method, the ideal number of clusters was found to be 4, which resulted in these profiles:

	Profile 1	Profile 2	Profile 3	Profile 4
<u>Education:</u>	High school diploma or GED	9-12 th grade, no diploma	High school diploma or GED	High school diploma or GED
<u>Gender:</u>	Male	Male	Male	Male
<u>Age:</u>	31 to 40	41 to 50	51 to 65	19 to 30
<u>Race:</u>	White	White	White	White
<u>Marital Status:</u>	Never married, single	Married	Divorced	Never married, single
<u>Occupation:</u>	Food Service	Other, Misc.	Construction & Extraction operations	Other, Misc.

Figure 5 – Table featuring our profiles

Once the profiles were created, the distribution of other variables within them had to be analyzed. Firstly, racial frequency was studied within our profiles, which can provide us with valuable insight. Furthermore, it aids in understanding race in the context of our clusters. For example, even though cluster 1 is dominated by the white ethnicity, there is still a large black minority within it. This is imperative in understanding which communities are most affected by accidental fentanyl overdose:

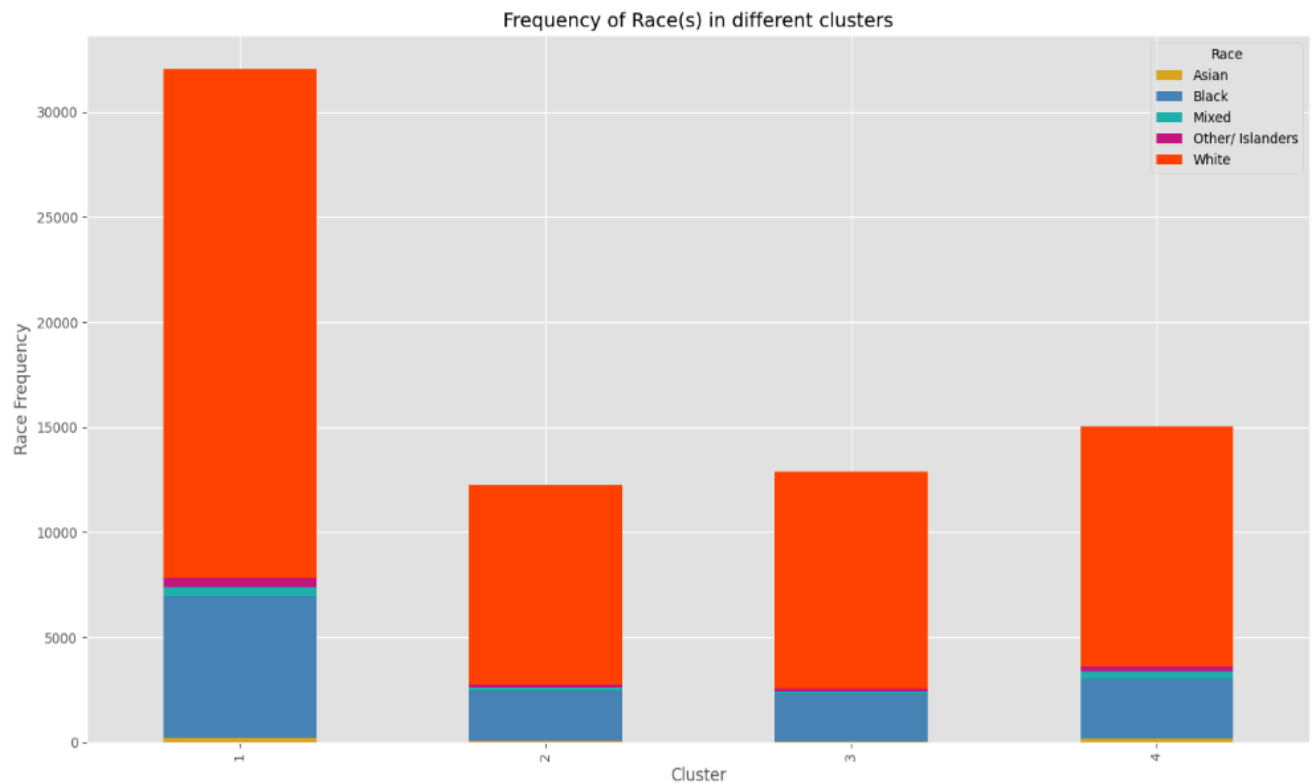


Figure 6 – Race frequency within our profiles

After that, the age distribution in our profiles was studied as well. One of the first observations made was how our age variable was used to establish the clusters. All our clusters and profiles had distinct age groups, granting us another avenue to study which communities or groups are most affected by this tragedy. Like ethnicity and education, these continue to give us necessary insights on who is affected, and how to tailor reduction strategies based on the profiles.

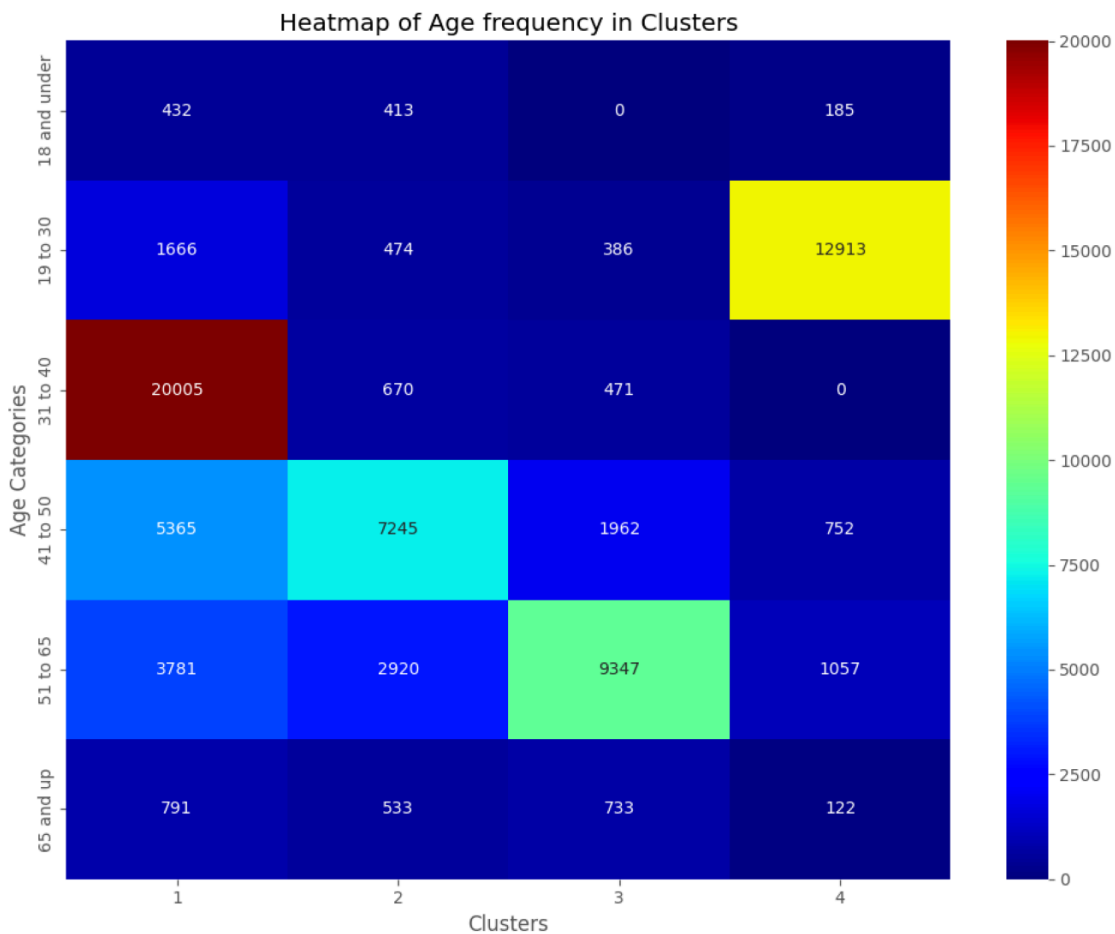


Figure 7- Age breakdown in our clusters, values represent count in cluster

Our four distinct profiles help paint a clearer picture as to who is most vulnerable to fentanyl overdose. This would be a low educated, white male, from ages 19 to 65. These men can either be single, married, or divorced and are most likely to be working in the food service industry or construction, among other occupations. Initial reactions were surprised to see that all the profiles were white, however after revisiting the data, it was found that out of all fentanyl overdose deaths in 2021, 76% of those were white.

After establishing our profiles, it was time to employ the xgboost model to our profiles. This approach was used because we wanted to predict the likelihood of overdosing based off

these profiles, and wanted to conduct feature importance to paint an even clearer picture of how the different variables contribute to fentanyl overdose.

	<u>Accuracy</u>	<u>Precision</u> <u>(0)</u>	<u>Precision</u> <u>(1)</u>	<u>Recall</u> <u>(0)</u>	<u>Recall</u> <u>(1)</u>	<u>F1-Score</u> <u>(0)</u>	<u>F1-Score</u> <u>(1)</u>
<u>Cluster 1:</u>	92.6%	0.18	0.94	0.07	0.98	0.10	0.96
<u>Cluster 2:</u>	90.5%	0.06	0.94	0.04	0.96	0.05	0.95
<u>Cluster 3:</u>	91.0%	0.05	0.94	0.03	0.97	0.03	0.95
<u>Cluster 4:</u>	93.2%	0.05	0.95	0.01	0.98	0.02	0.96

Figure 8- Model statistics for each cluster/profile, 0 represents other deaths and 1 represents accidental

Overall, our models were concise and almost entirely predicted ‘Accidental Drug Poisoning’ for every cluster. Each model was trained very effectively and yielded a high accuracy and good metrics for predicting the majority class in our target variable. Moreover, it also meant that our feature importance yielded accurate results as to which features within the profiles contribute most to overdose deaths. Our feature importance results can be seen here:

	<u>Cluster 1</u>	<u>Cluster 2</u>	<u>Cluster 3</u>	<u>Cluster 4</u>
<u>Feature 1:</u>	Age 18 & under	Occupation Farming, Fishing, Forestry	Occupation Unknown	Occupation Unknown
<u>Feature 2:</u>	Occupation unknown	Asian	Occupation Legal	Divorced
<u>Feature 3:</u>	Occupation Healthcare	Occupation Arts, Design, Entertainment	Occupation Life, Physical, Social Science	Widowed
<u>Feature 4:</u>	Occupation Legal	Occupation Unknown	Occupation Protective Service	Occupation Computer and Mathematical
<u>Feature 5:</u>	Occupation Farming, Fishing, Forestry	Doctorate or Professional Degree	Occupation Military	Occupations Agriculture and Engineering

Figure 9- Feature importance table, in descending order

After feature importance was completed, it was time to take a step back and study our results. Each cluster had similar results, but while we analyzed feature importance, we noticed that age, occupation, and race seemed to have the most impact on predicting accidental overdose deaths. This was interesting, as most of these features were not represented in the clusters. This also represented a unique challenge; as there isn't a way to tailor reduction strategies based on age and race. As seen in the feature importance, occupation played a major part in predicting accidental overdose deaths. Likewise, since education was our most consistent feature across profiles, we decided to use these two for tailoring reduction strategies. Also, there exists a strong relationship between these two features. A lower level of education directly correlates with worse career prospects; and vice versa.

Discussion & Conclusion

After completing modeling and clustering, we concluded that the clusters are effective in producing a model of populations vulnerable to accidental fentanyl overdose. Moreover, we were able to tailor reduction strategies based on features. Our models returned high accuracy levels, and extremely high precision values. Since precision measures how accurately a model predicts true positives, it is vital for our case to have high precision. Since precision was high for all models, we can confidently conclude that these models are effective at predicting accidental overdose deaths.

Since education and occupation are so highly correlated, we focused on education primarily while tailoring reduction strategies. We theorized that if you can improve the level of one's education, occupation prospects will follow. Our research yielded four main strategies with these two features in mind.

Education

The first strategy is to simply get students more involved in their education, and their schools. Student/Adolescent connectedness is when students feel that adults and peers in school care about their learning and for them as individuals. This can include a sense of being cared for, supported, and belonging at school. A study by the American Academy of Pediatrics found that a high sense of connectedness has positive implications on several factors, including prescription drug misuse and other illicit drug use. Comparatively, those with lower connectedness experience higher rates of substance abuse. In fact, higher connectedness is associated with a 48%-66% lower chance of health risk behaviors in adulthood [21]. Since we noticed that low education is so prevalent in our profiles, we believe that school connectedness efforts could go a long way in addressing fentanyl overdoses.

Standards-Based Health Education

While school connectedness is important, it is also imperative to educate students on drugs. Ideally, this would come after connectedness had improved. One of the main techniques for addressing drug use and establishing long lasting skills is standards-based health education. This education not only teaches students about drugs and their dangers, but also educates on safe substance use, and teaches life skills such as decision making, goal setting, and self-management [22]. Instilling positive habits in young people has been shown to reduce drug usage later into adulthood; even simply teaching students about safe substance use can positively affect disease transmission rates in the immediate area.

Low-Cost Treatment

As of April 2024, the National Center for Drug Abuse Statistics reports that the average cost of outpatient treatment ranges from \$2,000 to \$10,000. Inpatient treatment also costs on

average between \$6,000 and \$30,000 [23]. This is extremely expensive, as the median income of Americans is only \$74,580. Along with other expenses, most Americans are not able to afford outpatient treatment, much less inpatient treatment. Furthermore, many of the profiles we generated consist of occupations that are lower income/ completely unemployed. If someone in these profiles were to wish to seek treatment, it is likely they would be turned away. In fact, it was found by the National Council for Mental Wellbeing that 4 in every 10 American adults who needs treatment is unable to receive so, citing affordability as one of the main reasons [24].

Fentanyl Test Strips

Moving on to harm reduction, providing fentanyl test strips can be implemented to address the overall problem of accidental overdose. Our research proved that accidental overdose deaths are the main driver of the fentanyl crisis. One of the primary solutions to addressing this is using fentanyl test strips and making them widely accessible. Fentanyl test strips are a type of harm reduction strategy that can detect fentanyl in a wide variety of drugs such as cocaine, meth, heroin etc., and can even detect fentanyl in different forms such as pills, powder, and injectables [25]. In April of 2021, the CDC along with the Substance Abuse and Mental Health Services Administration (SAMHSA) announced that federal funds can be used to buy fentanyl test strips. However, this doesn't mean that all programs or outreach initiatives have begun to do so. We believe that implementing test strip programs can be effective at providing substance users an opportunity to safely screen their drugs before usage.

When researching strategies, we tried to implement a balanced approach between long-term and short-term solutions, and we believe that both in tandem can drastically reduce accidental overdose deaths. Not only did our models and profiles provide us with a better

understanding of what contributes to accidental overdose, they allowed us to build fully tailored reduction strategies.

When beginning research, we intended for non-profits and local governments to receive this information. We believe that because these organizations are closer to the crisis at hand, they will be able to allocate resources to vulnerable communities more effectively. Vulnerable communities, such as the individuals shown in our profiles, would be the ones receiving these resources and solutions.

However, we recognize that this study has limitations. The most glaring issue is the misrepresentation of minority groups within our clusters. Originally, we intended to perform a separate analysis on minority communities, but, due to time constraints, we were unable to complete this aspect of our research. While our initial results represented the data, more research is needed on minority communities, such as Black and Asian Americans. These people are still present within our clusters, but they deserve their own in-depth analysis.

Furthermore, we noticed that while accidental overdose deaths are by far the most common, they are not the only kind. Our models are poor at predicting other deaths and this illustrates some of the bias within our model. Further analysis on other death types such as suicide could uncover different profiles of vulnerable communities.

Lastly, it is important to note that more research needs to be done on tailored reduction strategies. While we determined four strategies that we consider potentially successful, we still believe there are more options and strategies out there that may perform even better. Additionally, while we proposed strategies, we were unable to test their effectiveness. Further

research could build upon this and determine quantitative measures of our proposed strategies' effectiveness.

Over the course of this paper, we studied the long standing and controversial history of America and drugs. Policy stemming from the War on Drugs was researched and applied to current times, examining the effects on society today. Discussion of what fentanyl is, how it works, and what makes it so addictive supported the framework of our literature and established the scope of the epidemic. We discovered the main contributor to the severity of the crisis is accidental consumption, and we promptly focused our research on potential solutions.

We built profiles based on the data and then built models based off these profiles, to predict accidental death and understand which features contribute to accidental overdose deaths. Once we found which features were most important, we did research into reduction strategies with these features in mind. It was concluded that a combination of long-term and short-term techniques would provide the most impact. Using these strategies, we simultaneously targeted younger, less educated people, and those most likely to accidentally consume fentanyl.

As stated above, more research could have been done into different reduction strategies and more analysis is needed on minority communities. However, our models and profiles provided fruitful analysis. All models performed well at predicting accidental overdose deaths and determining important features, and all strategies encompass the most important aspects of accidental overdose deaths. We believe this project to be a great starting point for further research and discussion into putting an end to the fentanyl crisis.

References

- [1] Han, Y., Yan, W., Zheng, Y., Khan, M. Z., Yuan, K., & Lü, L. (2019). The rising crisis of illicit fentanyl use, overdose, and potential therapeutic strategies. *Translational Psychiatry*, 9(1). <https://doi.org/10.1038/s41398-019-0625-0>
- [2] Overdose deaths from fentanyl laced stimulants have risen 50-fold since 2010. (2023, September 13). UCLA Health. <https://www.uclahealth.org/news/overdose-deaths-fentanyl-laced-stimulants-have-risen-50-fold>
- [3] Ramos-Matos, C. F. (2023, May 29). Fentanyl. StatPearls - NCBI Bookshelf. <https://www.ncbi.nlm.nih.gov/books/NBK459275/>
- [4] Han, Y., Cao, L., Yuan, K., Shi, J., Yan, W., & Lü, L. (2022). Unique pharmacology, brain dysfunction, and therapeutic advancements for fentanyl misuse and abuse. *Neuroscience Bulletin*, 38(11), 1365–1382. <https://doi.org/10.1007/s12264-022-00872-3>
- [5] The Editors of Encyclopaedia Britannica. (2024, January 25). War on Drugs | History & Mass incarceration. Encyclopedia Britannica. <https://www.britannica.com/topic/war-on-drugs>
- [6] Fair, H., & Walmsley, R. (n.d.). Research & Publications. Research & Publications | World Prison Brief. https://www.prisonstudies.org/research-publications?shs_term_node_tid_depth=27

- [7] S. E. Wakeman, S. E. Bowman, M. McKenzie, A. Jeronimo, and J. D. Rich (2009, April) “Preventing Death Among the Recently Incarcerated: An Argument for Naloxone Prescription Before Release,” Retrieved April 9, 2024 from <https://doi.org/10.1080/10550880902772423>
- [8] Ranapurwala, S. I., Shanahan, M., Alexandridis, A. A., Proescholdbell, S., Naumann, R. B., Edwards, D., & Marshall, S. W. (2018). Opioid overdose mortality among former North Carolina inmates: 2000–2015. *American Journal of Public Health*, 108(9), 1207–1213. <https://doi.org/10.2105/ajph.2018.304514>
- [9] Socias, M. E., & Wood, E. (2017). Epidemic of deaths from fentanyl overdose. *The BMJ*, j4355. <https://doi.org/10.1136/bmj.j4355>
- [10] Rudd, R. A., Aleshire, N., Zibbell, J. E., & Gladden, R. M. (2016). Increases in drug and opioid overdose deaths — United States, 2000–2014. *Morbidity and Mortality Weekly Report*, 64(50–51), 1378–1382. <https://doi.org/10.15585/mmwr.mm6450a3>
- [11] Richter, F. (2022, May 12). Fentanyl fuels surge in U.S. drug overdose deaths. Statista Daily Data. <https://www.statista.com/chart/18744/the-number-of-drug-overdose-deaths-in-the-us/>
- [12] Drug overdose deaths | Drug overdose | CDC Injury Center. (n.d.). <https://www.cdc.gov/drugoverdose/deaths/index.html>
- [13] Kearns, D. & American University of Washington D.C. (2023, May 17). To the Point: The Fentanyl Crisis, Why Now, Why So Deadly? American University. <https://www.american.edu/cas/news/to-the-point-the-fentanyl-crisis.cfm>

- [14] Frank, R. G., & Pollack, H. A. (2017). Addressing the fentanyl threat to public health. *The New England Journal of Medicine*, 376(7), 605–607.
<https://doi.org/10.1056/nejmp1615145>
- [15] Andritsos, P., & Tsaparas, P. (1970, January 1). Categorical data clustering. SpringerLink.
<https://link.springer.com/referenceworkentry/10.1007/978-0-387-30164->
- [16] Bonthu, H. (2023, September 14). KModes clustering algorithm for categorical data. Analytics Vidhya.
<https://www.analyticsvidhya.com/blog/2021/06/kmodes-clustering-algorithm-for-categorical-data/#:~:text=KModes%20is%20a%20clustering%20algorithm,cluster%20to%20determine%20its%20centroid.>
- [17] Nicodv. (n.d.). kmodes/kmodes/kmodes.py at master · nicodv/kmodes. GitHub.
<https://github.com/nicodv/kmodes/blob/master/kmodes/kmodes.py>
- [18] Nguyen, H. H. (2017). Clustering categorical data using community detection techniques. *Computational intelligence and neuroscience*.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5753022/>
- [19] XGBoost parameters👉. XGBoost Parameters - xgboost 2.0.3 documentation. (n.d.).
<https://xgboost.readthedocs.io/en/stable/parameter.html>
- [20] US Census Bureau. (2022, February 24). Census Bureau releases new educational attainment data. Census.gov.
<https://www.census.gov/newsroom/press-releases/2022/educational-attainment.html>

- [21] Álvarez, B. (n.d.). All hands on deck: School-based programs to stem substance abuse. NEA.
<https://www.nea.org/nea-today/all-news-articles/all-hands-deck-school-based-programs-stem-substance-abuse>
- [22] Steiner, R. J., Sheremenko, G., Lesesne, C., Dittus, P. J., Sieving, R. E., & Ethier, K. A. (2019, July 1). Adolescent connectedness and adult health outcomes. American Academy of Pediatrics.
<https://publications.aap.org/pediatrics/article/144/1/e20183766/37106/Adolescent-Connectedness-and-Adult-Health-Outcomes?autologincheck=redirected>
- [23] Average cost of drug rehab [2023]: By type, state & more. NCDAS. (2024, March 25). <https://drugabusestatistics.org/cost-of-rehab/>
- [24] National Council for Mental Wellbeing, & Majlessi, S. (2022, May 31). More than 4 in 10 U.S. Adults Who Needed Substance Use and Mental Health Care Did Not Get Treatment - National Council for Mental Wellbeing. National Council for Mental Wellbeing.
<https://www.thenationalcouncil.org/news/more-than-4-in-10-us-adults-who-needed-substance-use-and-mental-health-care-did-not-get-treatment/>
- [25] Centers for Disease Control and Prevention. (2022, September 30). Fentanyl test strips: A harm reduction strategy. Centers for Disease Control and Prevention.
<https://www.cdc.gov/stopoverdose/fentanyl/fentanyl-test-strips.html>