Analyzing the United State's Nationwide Opioid Crisis and Socio-economic Factors using K-Means Clustering

Ryan McGinnis
Thomas Jefferson College of Biomedical Sciences
Thomas Jefferson University
Philadelphia, PA, USA
Email: ryan.mcginnis@students.jefferson.edu

Les Sztandera
Kanbar College of Design, Engineering, and Commerce
Thomas Jefferson University
Philadelphia, PA, USA
Email: les.sztandera@jefferson.edu

***Abstract*— In this paper, we lay out the severity of the opioid crisis by focusing on the current literature of socio-economic addiction factors. We also discuss how opioids have become so prominent in the US. The analysis was done by taking data from one major city per state, totaling 50 cities, and putting them through a K means clustering analysis. Our findings revealed that commute times and average annual income had proportional increase and decrease to their cities opioid death rates. This builds to the current literature of the different quality of life factors that contribute to one's likelihood to develop an addiction.**

Keywords – opioid; k-means; clustering; socio-economic; factors; statistical analysis.

I. INTRODUCTION

Over the past several decades, the number of deaths due to opioid use has been rising at an alarming rate. It has reached a critical level where deaths related to opioid use surpass the number deaths due to motor vehicle accidents, gun violence and HIV [3]. This means that opioids can be considered the third leading killer in the United States, beaten only by heart disease and cancer. The opioid epidemic has grown from multiple issues. There are several different reasons someone would use opioids. Opioids are used either by prescription for pain management, or as a result of addiction.

Opioids have become a widely used pain management medication for patients with acute or chronic pain. Opiates are prescribed to those who may be recovering from an injury or surgery and will use opioid to better manage their pain. They are also prescribed for those with chronic pain and patients that need pain management medication for long durations of their lives. It is not uncommon for some of these patients to take the wrong dosage of medication, or forget they already took their medication, and accidentally overdose [2] [4]. In addition to this, they are highly addictive medication and considered a controlled substance by the Food and Drug Adminstration (FDA). It is very common for patients who are prescribed opioids for pain management to become dependent on them [7]. As a result, patients will recover from their injuries addicted to opioids.

Outside of prescription medication, there are many forms of opioid abuse from illicit drugs. Prescription opioids, heron, fentanyl, and other synthetic opioids are often illegally sold on the streets to those suffering from addiction. Opioid addiction being a large contributor to the opioid death tool, it is much more dangerous using the street version of these drugs because they are often diluted with fillers or other drugs. Part of this is because of how accessible opioids and their synthetic counterparts have become. There has been network analysis done to track the distribution of opioids coming from South American countries into the United States [5]. It is no question that the East Coast has been hit the hardest in opioid epidemic simply because that is where the supply is being distributed.

The opioid epidemic also stays out of the public's eye because of how it is factored by the Center for Disease Control (CDC). The CDC considers any opioid related death an "accidental poisoning" and, therefore, falls in with many other categories of death that are considered accidents. This category of accidental deaths is the third leading killer in the United States. Breaking down the different accidents in this category, opioid related deaths would have the highest death toll. Because of this, opioid deaths have surpassed motor vehicle accidents, gun violence, and HIV [3]. However, all opioid accidents are avoidable, and addiction can be treated. With better management for pain treatment, opioid prescription frequency, and how we manage addiction in our cities, it is possible to minimize the opioid death toll.

There are many retrospective studies that have been done that analyze and discuss what demographic factors influence addiction. One study looked at 5,483 overdose patients and laid out all demographic factors for analysis [9]. The most significant finding was that more than half of the patients had an opioid prescription within 90 days of their death. This tells us that opioid prescription history is a massive contributor to one's likelihood to overdose. Having the initial exposure to opioids opens the gate to addiction that the patient otherwise would never have if they were prescribed a different medication. In addition to opioid prescription history, the results showed that the majority of the patients were white/non-Hispanic. Although it is difficult to confirm, many papers analyze ethnic backgrounds of their patients for trends. The challenge that this brings is the diversity of factors that could be

 ISBN: 978-1-61208-891-4

influencing the data, outside of the scope of the experiment. For example, ethnic ratios are not likely to be proportionate in each city/state/country where the study is done and this will influence the discussion around the results. It is best to take ethnic background data on the opioid epidemic lightly until there is a larger collection of studies analyzing the data with different factors. Some interesting factors that were studied in the paper reference were education background and marital status. The highest overdose rates came from those with only a high school diploma. The rate of overdoses declined with each higher level of education, bachelor's degree, master's etc. Marital status opioid rates were relatively the same across married, divorced, or single, however there was a steep drop off in the widowed category. Those who were widowed were far less likely to overdose on opioids. Factors like these raise questions as to what exactly can lead one to an opioid dependency, and thus are important to analyze with as many studies as we can.

In "Analyzing the relationships between city opioid deaths and socioeconomic factors" [6], we looked at one major city from each state's opioid death toll, average commute time, budget for roads, and budget for arts and culture in 2018 for a total of 50 cities their respective variables. We also collected data from 2017 however not the full 50 cities due to the fact that budgets and their timelines are allocated differently for each city. The reason we collected these variables was so that we could analyze what different socio-economic factors were related to the opioid death toll, speculate as to how they factored into addiction, and reference other studies that were done to support our findings. The data was analyzed via ANOVA tests and Logistical Regression Tests. Logistical Regression was chosen for this experiment because it showed the trendline for the data. Our findings revealed that commute time yielded significant results in nearly all of the tests it was included in. This was complimented by budget for roads yielding significant results in some of the tests, while budget for arts and culture showed occasional significant results but with no pattern. This shows that commute times had a strong relationship with one's likelihood to overdose on opioids. Since the city's commute times are most likely determined by their respective budget for roads, it makes sense that that variable also share some significant results with commute time. We concluded that commute time and its opioid addiction relationship is connected by a quality-of-life factor.

The significance of these papers that analyze external factors form opioid deaths is that they are a novel approach to uncovering what drives addiction. There is already a massive quantity of papers that look at addiction related deaths retrospectively and compare what each patient had in common [8]. They often look at factors like ethnicity, financial status, education, or career, but they leave out all of the external factors. External factors may prove to be critical in identifying one's likelihood of addiction. External factors are what people are exposed to on a regular basis that can influence addiction, regardless of their demographic factors. By analyzing as many different factors as we can, we can create a foundation of external factors that are known to influence addiction. This will take many more studies of replication and novel approaches to build this foundation, but it will play a critical role in prediction and prevention in the world of addiction medicine.

In this paper, we aim to look at how external factors in one's surroundings may be a contributing factor to their likelihood of using opioids [1]. We will look at average commute times, budgets for roads, and budgets for arts and culture within each city and compare them their opioid death count of 2018. The significance of this, is that it gives cities predictor and preventative models to both prepare and avoid increasing opioid deaths. With K means cluster analysis we can group cities into different categories base on their average commute times and how they allocate their budget.

We will also compare the results with Cluster Mapping data from Harvard Business School and US Economic Development Administration, 2018 [8]. This allows us to look at the country by metropolitan county and see if there are any patterns that are comparable with our results. In Figure 1, we can see the average annual wage, clustered into their respective metropolitan regions. The different annual wage averages can give us some insight as to whether or not it has any effect on one's likelihood of opioid abuse. Looking to see if there are any financial thresholds where cities above or below any certain amount of funding could show some consistency with our clustering data. By comparing these patterns without clustering data, we could reason that financial status does play a role in a city's opioid crisis. If wealthier cities reflect to have higher opioid death rates, this can be used as predictor model.

II. MATERIALS AND METHODS

For this experiment, we looked at each state in the US and took one major city from them. For each of those cities, we took the following data from 2018: total opioid death toll, average commute time, budget for roads, and budget for arts and culture. We used IBM Statistical Package for the Social Sciences (SPSS) statistical software version 25 to analyze the data.

The first thing we did was standardize all the variables so that they could be compared to each other. This was done by converting them all to their Z scores. In SPSS, a descriptive test was run with all of the variables, and metrics set to their default settings.

With the Z scores collected, the K means cluster analysis was ready to be run. The Z scores were added to the variable list with their labels classed by cities. The maximum number of iterations was set to 99. This is because with

 ISBN: 978-1-61208-891-4

TABLE 1. EACH CITY WITH THEIR RESPECTIVE CLUSTER LABEL FROM THE K-MEANS CLUSTER ANALYSIS.

City	Cluster
Philadelphia	1
Chicago	1
Phoenix	1
Baltimore	1
New York City	2
Detroit	3
Portland	3
Houston	3
Los Angeles	4
Boston	4
Providence	4
Newark	4
Charlotte	4
Indianapolis	4
Las Vegas	4
Atlanta	4
Seattle	4
Orlando	5
Charleston	5
Manchester	5
Louisville	5
Columbus	5
Albuquerque	5
Salt Lake City	5
Nashville	5
Portland	5
Wilmington	5
Hartford	5
Burlington	5
Milwaukee	5
St Louis	5
Columbia	5
Richmond	5
Anchorage	5
Oklahoma City	5
Denver	5
New Orleans	5
Birmingham	5
Cheyenne	5
Minneapolis	5
Des Moines	5
Little Rock	5
Jackson	5
Boise	5
Wichita	5
Fargo	5
Sioux Falls	5
Billings	5
Honolulu	5
Omaha	5

enough iterations that clusters will develop a pattern and average out. The cluster membership was set to save so in retrospect we could analyze which city fits in which cluster. The number of clusters was set to 5. The reason 5 clusters were chosen was because during the initial testing, too few didn't offer enough variation across clusters. Anymore than 5 clusters and we saw that groups started to replicate with no significant difference. 5 clusters gave us a good variation of the different groups that could emerge from the different ways commute time and budget allocation factors into opioid deaths. The ANOVA table was also saved to analyze variance and significance scores of the variables. All other metrics were set to their default settings.

III. RESULTS

In Table 1, we can see how the software categorized the different clusters. The software plotted all of the data points of the variables and grouped them into 1 of 5 clusters based on these generated values. With each iteration, there comes 5 new locations for where the clusters are centered. Each iteration offers a different way to group these cities into clusters. The more iterations that are done helped the software find an average cluster for each city until the city's cluster assignments become redundant. Once the software finds the average cluster that all the cities are assigned, the iterations end, and results are shown. From here we can analyze where each city falls in relation to the clusters.

In Table 1, we can see each city and in which cluster they were categorized. It is worth noting that the opioid crisis is so severe in New York City, that it is the only one in Cluster 2.

IV. DISCUSSION

There are a few key takeaways that this data reveals to us. Initially, it is clear that cities with higher populations tend to lean towards clusters 1, 3 or 4, while smaller cities tend to fall into cluster 5. A simple predictor of the opioid death toll can be population. The more people, the more likely there will be a higher opioid death toll.

In Figure 2, the commute time always sits next to the opioid death toll. This can infer that in many circumstances, the opioid death toll of a city could be predicted by looking at the average commute time for that city. Knowing this, cities can use this as a predictor model for gauging and preparing opioid casualties in their respective cities. The measures that could be taken can be to increase the amount of naloxone supply that emergency responders carry or increasing patrols in high-risk opioid overdose areas. Knowing that commute times are not something that can create addiction, we must ask ourselves how this factor is consistent with opioid deaths across cities. Like many other contributors of addiction, it is a reflection of the quality of life. For the portion that initially seek opioids in their illegal

 ISBN: 978-1-61208-891-4

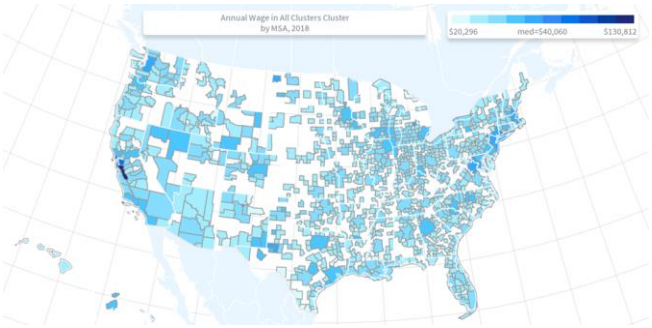

Figure 1. US Cluster Map by average salary per county.

forms or those that have addiction issues prior to their opioid prescription, quality of life is something that has major effects on one's likelihood of developing addiction.

Alternatively, we can see that the city's budget for arts and culture do not show any consistent patterns together. In all clusters, we can see that opioid deaths and budget for arts and culture are inconsistent with each other. As opioid deaths rise and fall, budget for arts does not rise or fall or vice versa with it. This means that budget for arts and culture acted as our control group. By showing us that some groups have no relationship, we can focus on the groups that show relationships like opioid deaths and commute times. Looking at Figure 1, we can see the annual wage averages by metropolitan area of 2018. By comparing the cluster groups to their city's annual wage, we can see that almost all of the high annual wage cities have a proportional relationship with high opioid death tolls. New York City, being in cluster 2 has one of the highest annual wages of $73,000. Close to New York there is Philadelphia in cluster 1 with an annual wage of $59,000 and Baltimore in cluster 1 with an annual wage of $57,000. The surrounding metropolitan areas have a lower annual wage and by more than $10,000 and lower opioid death counts. However, if we look at Los Angeles in cluster 4, it has an annual wage of $60,000. The opioid crisis, being as complex as it is, leads us to believe that this cannot be predicted by commute time, roads, or wage alone, but by a combination of actors. Knowing that the opioid crisis is worse on the East coast, we can see that reflected here by how death rates are higher in wealthier cities along the East Coast with harsher commute times, but not along the West coast.

Illicit opioids and their synthetic counterparts are some of the most lethal illegal drugs in the US right now. With massive amounts of opioids being brought into the US internationally and devastating the east coast, it has become a two pronged attacked on those susceptible to addiction. Not only do those who work their tolerance up eventually become addicted to opioids, but it is very common who have finished their prescription with a dependence on the drug, continue to seek, use and thus turning to the more accessible and cheaper illegal sources. For these illegal opioids, it is not simply a matter of population density, otherwise we would see different results in our study. For example, Los Angeles being in cluster 4 but having one of the highest populations in the country. Illegal opioids are much more widely distributed along the East Coast. Synthetic opioids can be cheap or diluted to be sold at a margin of their price, however purchasing prescription opioids on the street are far more expensive. Knowing this, we must also consider the economic state that some of these critically hit areas are in.

Looking at Figure 2, we can see that the majority of our critically impacted cities, being in clusters 1 or 2, are all metropolitan areas that average over $55,000 annual wage. Knowing this, it stands to reason that these areas have the financial means to afford these opioids either via prescription or illegally. Looking at other cities on the chart, it is interesting to see how as soon the average wage falls roughly below the $55,000 annually, these cities end up in clusters 3, 4, and 5. This tells us that economic status can play a significant role in one's likelihood of addiction to opioids. This is most likely due to the fact that opioids are one of the most expensive drugs to abuse both legally and illegally. Naturally those that cannot afford prescription opioids, or their illegal street counterparts, will not have the same level of exposure as those that can. Those that can afford opioids both in their prescription form or their illegal street form have a much higher risk of becoming addicted. If one is in a pain management situation and can afford an opioid prescription, they are at a much higher risk than it may seem. If they can afford the opioid prescription, they have the risk of developing a dependency. Not only are they at the initial risk from their prescription but following their prescription their odds of developing an addiction only increase if they live in one of these wealthier cities, on the East coast, and make an average income greater than $55,000.

Knowing that those making above $55,000 annually may be a factor contributing to addiction, we can compare that to the findings of the study discussed earlier in this paper. There is a steady decline in overdose likelihood with degree of higher education. Typically, higher salaries are earned by those with a strong educational background. Naturally there are exceptions, but there must be some middle ground of those with only a high school diploma, although still in a high-income job. It is possible that the patients that made up the high school diploma category are from an older generation when going to college was not a standard. This generation would fit the age group where it is very common to take the wrong dose of opioid medication and passing away in their sleep. This means they would also fit the category of having an opioid prescription within 90 days of their death. Evidently, knowing what factors contribute to addiction helps us build the circumstances as to one's likelihood of overdosing.

V. CONCLUSION

When considering strategies to minimize the death toll of opioids, one must first consider the complexity what

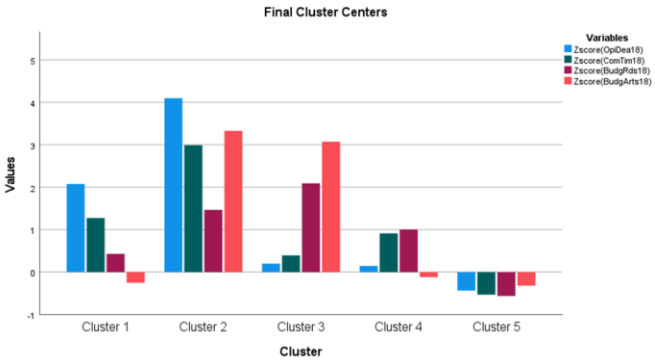

Figure 2. Results from k-means cluster analysis.

influences it. We know that they are a pillar of pain management in modern day society, highly addictive, and have little to no substitutes for chronic pain. They are so addictive that prescribed patients are likely to develop some level of dependence of these drugs, either causing addiction to the prescription opioids themselves or starting a lifestyle of addiction to similar drugs. In addition to having a prescription to an opioid, there are demographic factors that contribute to one's likelihood of addiction. This gives insight as to who is likely to become dependent on their opioid prescriptions and developing addiction, however using this information to prescribe opioids and their strength is at the discretion of the provider and is not uniform. Because of all this, it is to be expected that opioid deaths are to be higher in more densely populated areas. The more people, the more doctors, the more circumstances there will be to prescribe opioids. However, this is not the only pathway that opioids contribute to their death toll.

Opioids have proven to become one of the most dangerous drugs in our society. Being one of the only medications on the market suitable for pain management and its high likelihood of developing a dependency, it creates a high volume of addiction in our society. In addition, opioids are being illegally distributed across our country granting even easier access to them. Legally or illegally, opioids are one of the strongest and easiest drugs to get a hold of in this country. With opioid deaths only continuing to rise, the best thing we can do is uncover what are the factors that drive this epidemic. From this paper, we discovered that a city's commute plays a major role in their opioid death toll. As commute times increased or decreased, the death followed. We also found that annual income is also a factor. Having the financial means to afford these drugs, either legally or illegally, only builds to the accessibility of these drugs. The opioid epidemic is driven by many compounding factors. Knowing that commute times and annual income rates are some of these factors can give cities more insight to prepare and manage their populations.

Further studies can be done that analyze other quality of life factors to see if any of them have a proportionate role in opioid death rates. The model of the study could also be changed. Instead of looking at cities, additional studies could analyze the data by zip code or county. One of the difficulties that this study faced is that, although we analyzed city's budget for roads and arts and culture, there was no way knowing exactly where these funds were allocated to. In 2018, it is possible that a large portion of the budget went to one project as opposed to spread evenly across the city. Because of this, it would be interesting to look at these cities on a smaller scale. If the data is available, analyze where the budgets were allocated (filling potholes, new bridges, multi neighborhood) and compare to how the opioid death toll was affected. One could also use the Harvard Business School Clustering Map, to analyze other factors that may be proportional to a location's opioid death rate. There is more work to be done in order to have a better understanding as to all the different factors that influence addiction and opioid death rates.

REFERENCES

[1] A. Alalawi and L. Sztandera, "Leveraging Statistical Methods and AI Tools for Analysis of Demographic Factors of Opioid Overdose Deaths." International Journal on Advances in Life Sciences: 12(1&2): 24 – 33.

[2] J. Baird et al., "A retrospective review of unintentional opioid overdose risk and mitigating factors among acutely injured trauma patients." Drug Alcohol Depend. 2017; 178: 130 - 135. DOI: 10.1016/j.drugalcdep.2017.04.030.

[3] D. Ciccarone, "Fentanyl in the US heroin supply: A rapidly changing risk environment." The International Journal on Drug Policy. 2017 Aug; 46:107-111. DOI: 10.1016/j.drugpo.2017.06.010.

[4] P. Dilokthornsakul et al., "Risk factors of prescription opioid overdose among Colorado Medicaid beneficiaries." J Pain. 2016; 17 (4): 436 - 443. DOI: 10.1016/j.jpain.2015.12.006.

[5] L. Giommoni, A. Aziani and G. Berlusconi, "How do illicit drugs move across countries? A network analysis of the heroin supply to Europe." J Drug Issues 2017; 47 (2): 217 - 240. DOI:10.1177/0022042616682426.

[6] R. McGinnis, L. Sztandera, R. Vadigepalli, and J. Ruane, "Analyzing the relationships between city opioid deaths and socioeconomic factors" J Opioid Mgmt. 2021; 17(5): 363-382. unpublished.

[7] S. Nechuta, B. Tyndall, S. Mukhopadhyay and M. Mcpheeters, "Sociodemographic factors, prescription history and opioid overdose deaths: A statewide analysis using linked PDMP and mortality data." Drug Alcohol Depend. 2018; 190: 62 - 71. DOI: 10.1016/j.drugalcdep.2018.05.004.

[8] US Cluster Mapping. Harvard Business School. Retrieved June 12, 2021 from https://clustermapping.us/region

[9] B. Zedler et al., "Risk factors for serious prescription opioid-related toxicity or overdose among veterans' health administration patients." Pain Med. 2014; 15 (11): 1911 - 1929. DOI:10.1111/pme.12480.

 ISBN: 978-1-61208-891-4