Opioid Overdose Analysis

Mid-Atlantic Opioid Task Force

Aviv Farag
af3228@drexel.edu

Joseph Logan
jrl362@drexel.edu

Daniel Turnbow
djt96@drexel.edu

Silas McKee
sm4543@drexel.edu

Abstract – This paper describes the study of opioid overdose in Pennsylvania in which related datasets were pulled from Census, CDC and PAGOV. Those datasets were read and merged in Python to create new features to explore the overdose epidemic. For better and clean representation of the information that was acquired and pre-processed, Tableau was utilized to create interactive maps for both County and State levels.

Keywords – Tableau, Opioid, Python, API, Census, CDC, PAGOV

GitHub Repository – https://github.com/Joeyloganpython/Capstone

Dashboards – https://joeyloganpython.github.io/Capstone/

♦

1. Introduction

In 2021 deaths due to drug overdoses in the United States hit an all-time high of 100,000 [1]. Opioid overdoses are a major issue leading to many of those deaths. One of the most sinister elements of the current opioid crisis is the addition of fentanyl. Fentanyl a highly concentrated opioid, which is sometimes laced into other drugs, which can be dangerous for those who ingest them.

Many of the publicly available datasets, such as those from the CDC, or state level datasets, are challenging to work with due to incomplete data, very large datasets, or differences in data collection practices. Our project goal is to bridge those gaps by providing actionable analysis, dashboards, and visualizations that public health officials or volunteer organizations can use to serve their communities.

1.1.Fentanyl

Fentanyl is a highly concentrated, synthetic opioid [2], and is sometimes laced into other drugs, often causing those who take them to ingest a profoundly higher dose than intended. Fentanyl is extremely potent, which means that even small absolute amounts can lead to an overdose, especially for users who have not developed a tolerance to opioids.

Fentanyl is similar to morphine, though 50 – 100 times more potent [2]. In 2022 the DEA has issued a public warning for fentanyl related overdose death, stating that fentanyl poisoning and deaths are at an all-time high. Fentanyl is cheap to illegally manufacture and can come in pill, eye drop and paper form. The pills can be made to look like other prescription medication.

1.2. Naloxone

According to drugabuse.gov [3], a site maintained by the National Institute on Drug Abuse: Naloxone is a medication designed to rapidly reverse opioid overdose. It is an opioid antagonist meaning that it binds to opioid receptors and can reverse and block the effects of other opioids. It can very quickly restore normal respiration to a person whose breathing has slowed or stopped as a result of overdosing with heroin or prescription opioid pain medications.

The brand Narcan is a form of naloxone administered as a nasal spray. It is packaged in a carton containing two doses to allow for repeat dosing if needed. Narcan is an easy-to-administer form of the medication, and the training does not take long

2. Datasets

This study examines data on both the state and county level. Therefore, data was harvested from the CDC [4] (Centers for Disease Control) for all the states and from PA.gov [5] for counties within Pennsylvania. Furthermore, census data [6] was used to incorporate state and county population data into our features to enable meaningful population-based comparison across geographic areas (e.g., overdose deaths per 10,000 people per state)

2.1. Data Acquisition

Data was pulled from three primary sources: PA.gov, the CDC, and the Census. The PA.gov and CDC datasets were all available via Socrata APIs [7] which provide access to a variety of open data resources from governments as well as other organizations.

The Census data is accessed via the Census microdata API [8]. A guide for calling the API can be found at Census Microdata API User Guide [9].

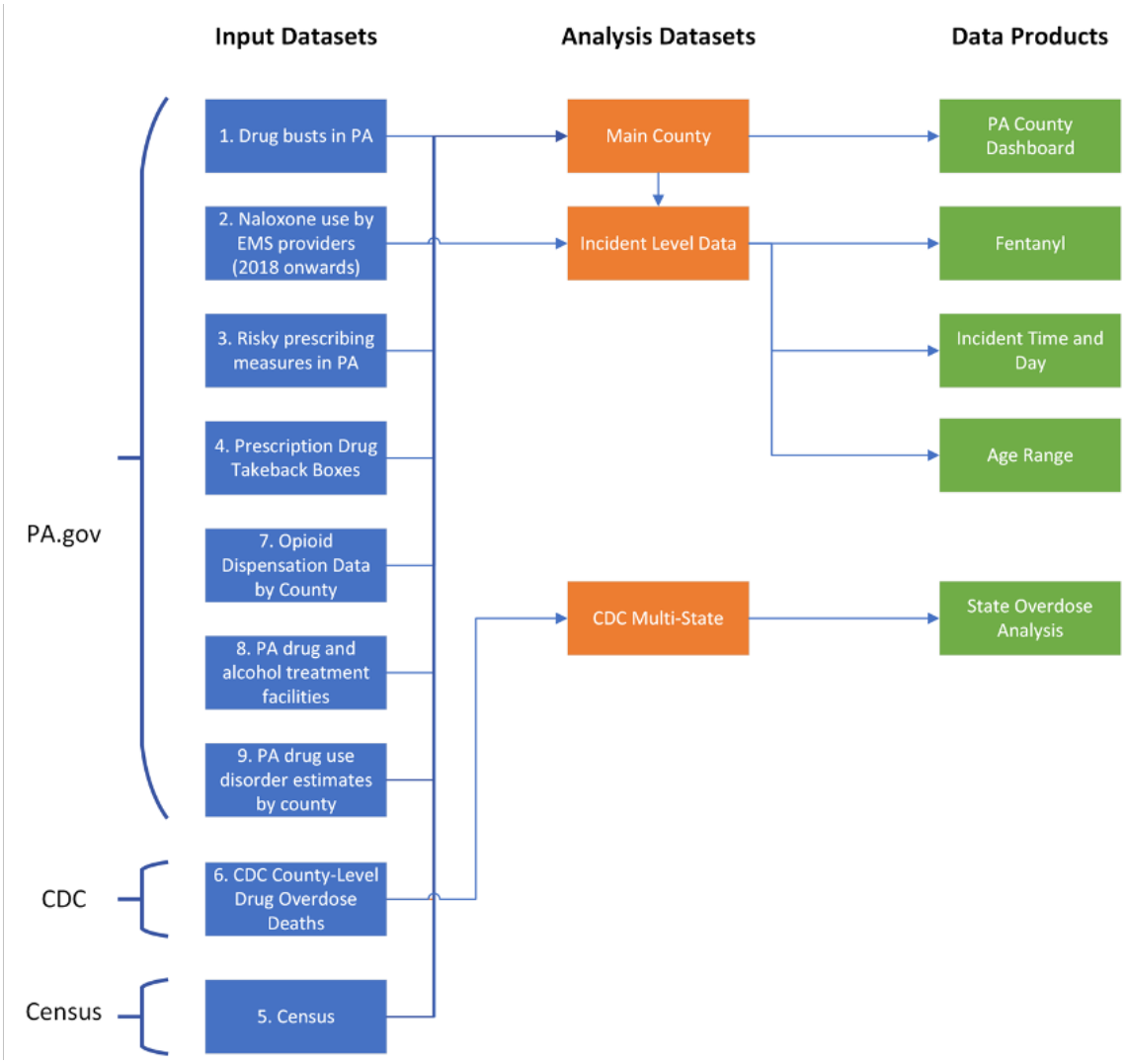

Figure 1- Data Preprocessing Diagram

2.2. Data Preprocessing

Figure 1 shows the preprocessing procedure. The left side (input datasets) contains a list of all datasets that were pulled and used in this study, and they were grouped by their source (PA.GOV, CDC, and Census). The middle column (Analysis Datasets) is where the datasets were combined, cleaned, and analyzed to produce the Data Products in the last column.

3. Exploratory Data Analysis

This study examines several rich datasets from different sources. To efficiently analyze the data in both County and State levels, Dashboards were created using Tableau Public. Those dashboards are shared on a website that was developed using GitHub Pages.

3.1. Basic Metrics

PA.GOV dataset was filtered by the drug that was used, so that the new dataset contains record related to opioids. Then, information regarding the variables was printed as shown in the figure below.

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 19235 entries, 2 to 36256
Data columns (total 10 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Incident ID    19235 non-null  int64
 1   County         19235 non-null  object
 2   Narcan Admin   19235 non-null  int64
 3   Survive        19235 non-null  int64
 4   Gender Desc    19235 non-null  object
 5   Incident Date  19235 non-null  object
 6   Day            19235 non-null  object
 7   Incident Time  19235 non-null  object
 8   Age Range      19235 non-null  object
 9   Year           19235 non-null  int64
```

Figure 2 - Variables Information

This dataset contains 19,325 rows and there are no missing values. Next, univariate analysis was executed by creating histogram and listing unique values. County unique values are as follows:

```
array(['Delaware', 'Chester', 'Beaver', 'Bucks', 'Philadelphia',
       'Washington', 'Cumberland', 'Northumberland', 'Montgomery', 'Pike',
       'Armstrong', 'Carbon', 'Centre', 'Bradford', 'Dauphin', 'Lehigh',
       'Blair', 'Allegheny', 'Erie', 'York', 'Lebanon', 'Monroe',
       'Franklin', 'Lancaster', 'Berks', 'Mifflin', 'Westmoreland',
       'Crawford', 'Mercer', 'Luzerne', 'Susquehanna', 'Elk', 'Lycoming',
       'Snyder', 'Lackawanna', 'Lawrence', 'Wayne', 'Juniata', 'Perry',
       'Northampton', 'Adams', 'Cambria', 'Schuylkill', 'Clearfield',
       'Tioga', 'Columbia', 'Potter', 'Fulton', 'Wyoming', 'Butler',
       'Somerset', 'Fayette', 'Indiana', 'Jefferson', 'Clarion',
       'Bedford', 'Greene', 'Huntingdon', 'Union', 'McKean', 'Montour',
       'Forest', 'Warren', 'Clinton', 'Venango', 'Sullivan'], dtype=object)
```

Figure 3 - County Unique Values

There are 67 counties in PA, but there are only 66 listed in Figure 3, meaning there is one missing county which is Cameron. Distribution analysis of Gender, Year, Survive and Naloxone Administered is shown in Figure 4 below.

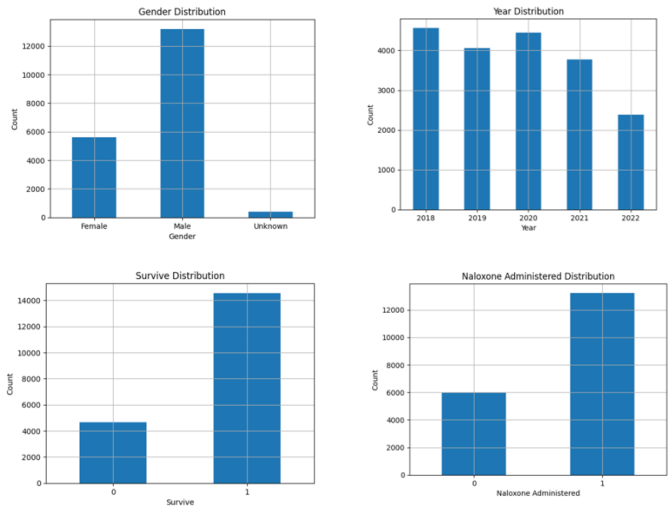

Figure 4 – Histogram of Gender (Top Left), Year (Top Right), Survive (bottom Left), and Naloxone Administered (Bottom Right)

Gender shows that more males are more involved in overdoses, and that the number of incidents is consistent throughout the years. Note that the year 2022 hasn't over yet, so the data is limited, and therefore much lower than other years. Survival rate is high (~14,000) compared to around 4,000, but this is probably due to Naloxone administration which is also high (~12,000) compared to around 6,000 cases without Naloxone.

3.1. Naloxone Effectiveness

Naloxone effectiveness over age ranges, and age range analysis appears on Figure 5. The graph on the top shows survival rate both with (Orange) and without (Blue) naloxone, across different age ranges. The lower bin graph shows the amount of incident for each group. Most overdose cases are in the 30 to 39 age group with more than 6,000 cases. Also, the upper graph clearly shows the impact of naloxone on survival rate which is a steady line above 80% survival rate compared to around 50% without naloxone.

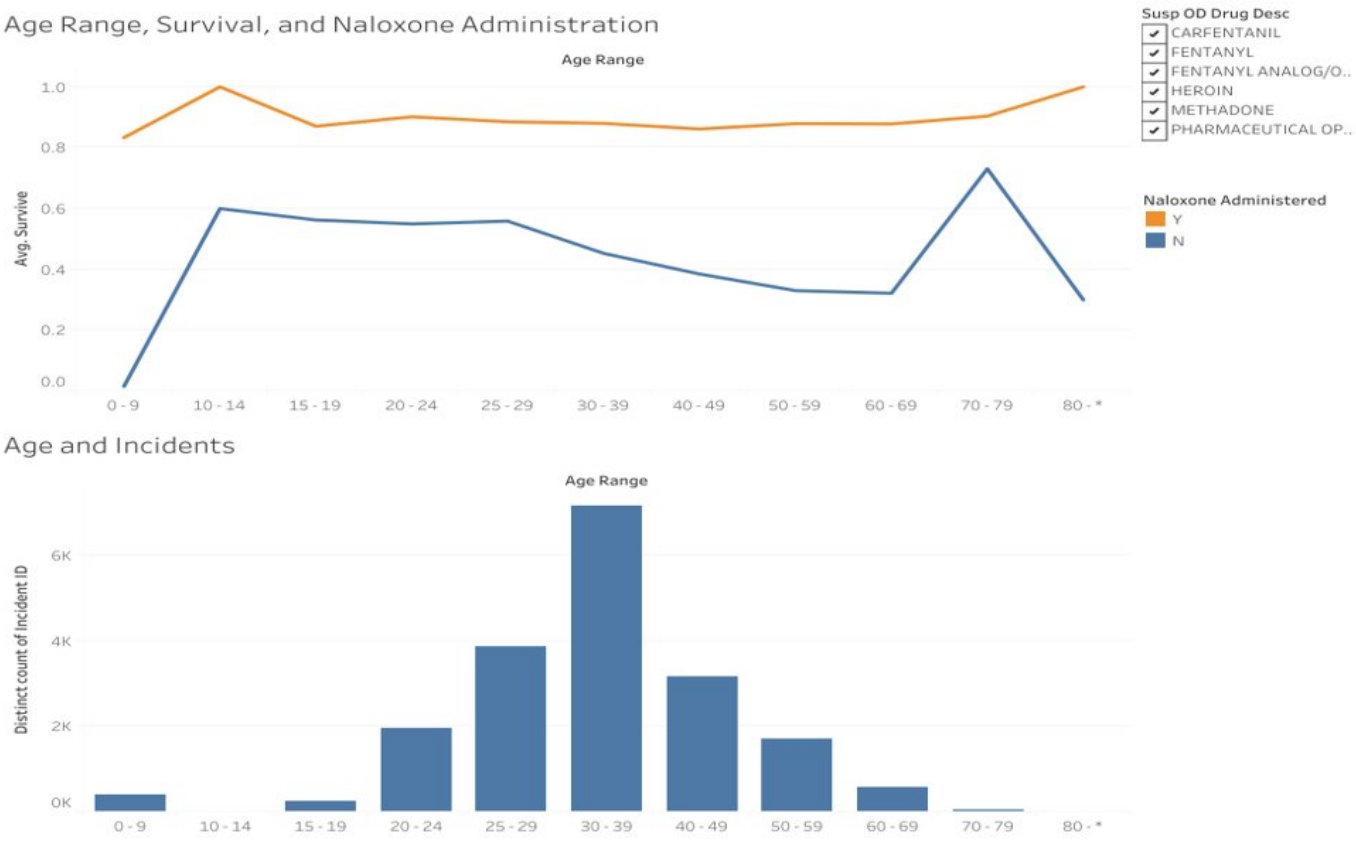

Figure 5 - Survival Rate and Incident Count across age groups

3.2. Gender

Another feature that was examined is the gender among age ranges which appears in Figure 6. Here Female is highlighted in purple, Male in Green, and Kids (Unknown) in Red. Males experience overdoses more often than females, and there is a great number of youngsters (0-9 years old) involved (around 500).

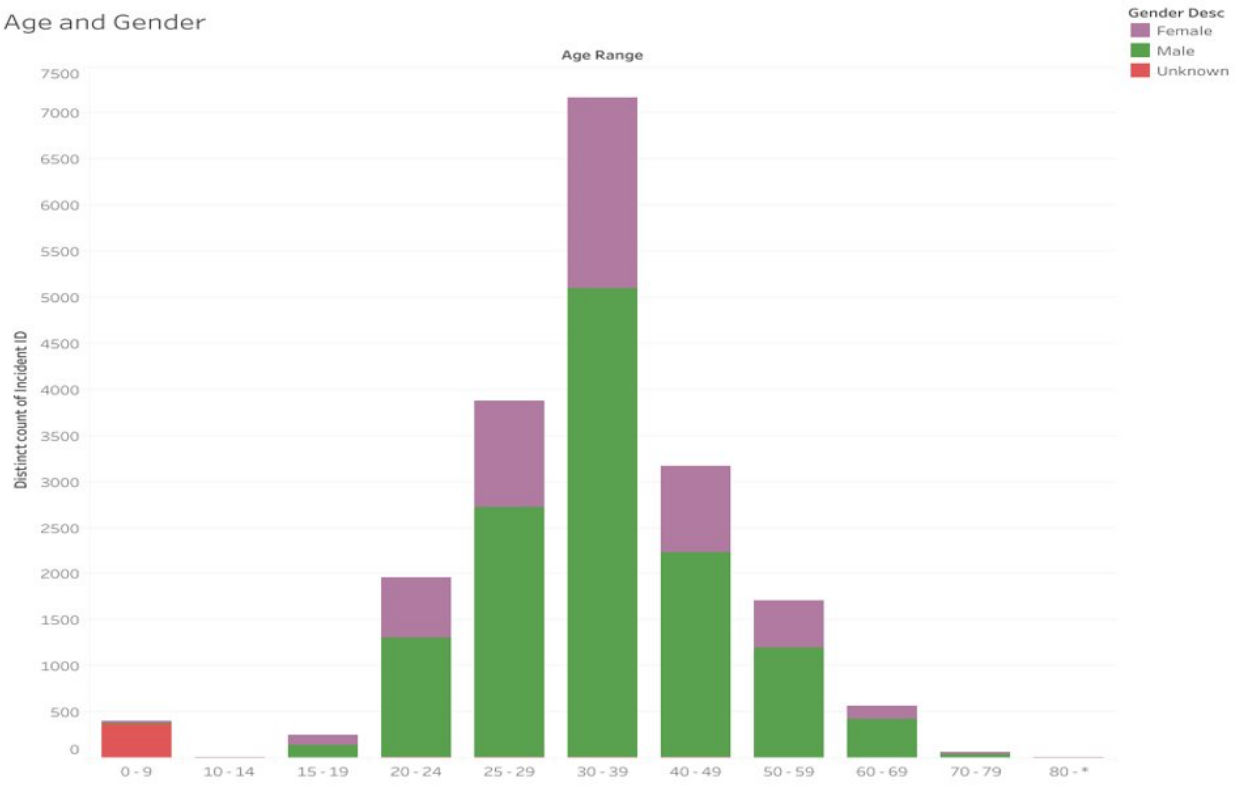

Figure 6- Age and Gender

3.3. Fentanyl Overdose

One type of opioid is Fentanyl, and its impact on survival rate without naloxone is shown in Figure 7. The x-axis represents ranges of age groups, y-axis is survival rate, where orange line represents Fentanyl overdose cases, and blue line represents other opioids.

Fentanyl is an important factor, especially if the victim is older. When naloxone is not given and the overdose is from fentanyl there is only a 30% survival rate, compared to a 48% survival rate when naloxone is not given, but it's another type of opioid.

3.4. Time and Day of Overdose Incidents

The datasets contain the date and time of the incidents which is useful for analyzing when overdoses are more likely to happen. This could be beneficial for both Emergency Medical Services and the Police to arrive faster at a scene and potentially reduce death rate.

The top graph in Figure 8 shows the amount of overdose incidents for every hour in a day, whereas the bottom graph shows both the average Narcan administered and the average survival rate.

More incidents occur during the evening hours, around 1,000 compared to 400 in the morning, suggesting that many of the overdose victims have full time jobs/other responsibilities. This reinforces the hypothesis of many of these overdose events possibly being fentanyl laced into other drugs.

Although there are more overdoses in the evening, there is a lower % chance of survival and lower Naloxone administration in the morning hours. This may indicate that victims are found unresponsive in the morning.

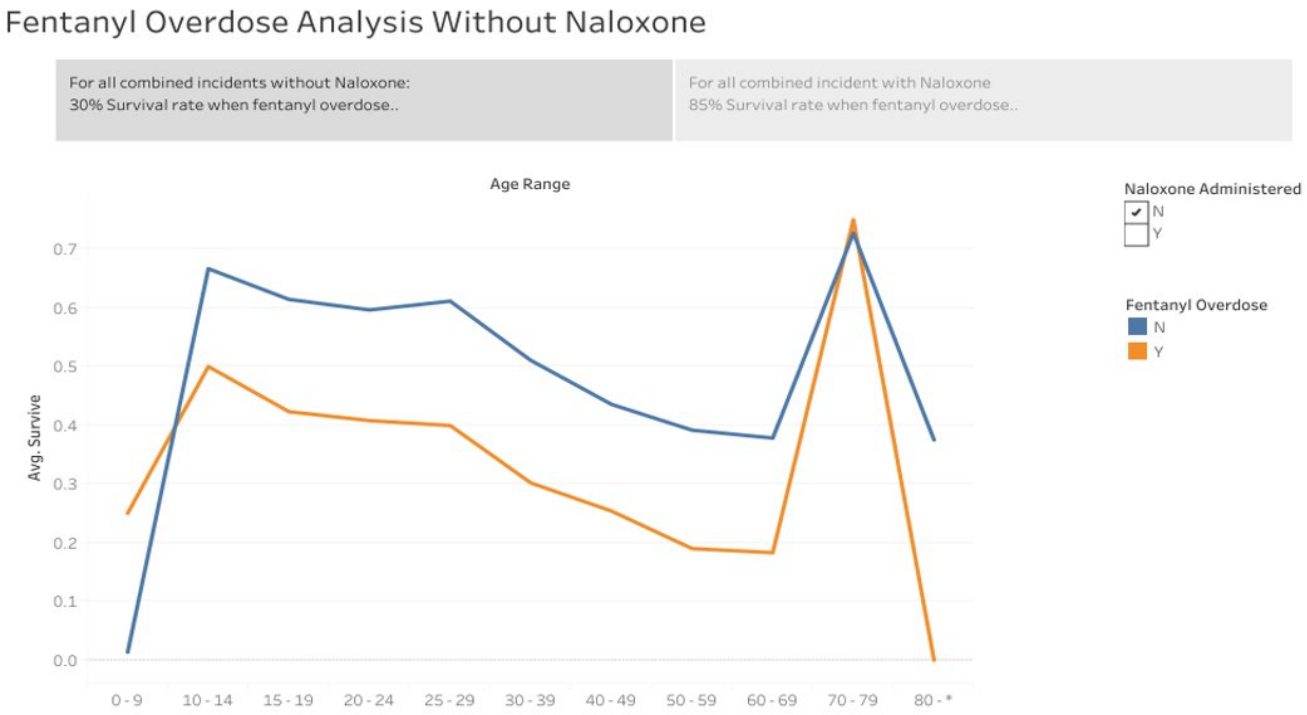

Figure 7 - Fentanyl Overdose Analysis without Naloxone

3.5. Day of week

Day analysis is shown in Figure 9, where it shows a bar chart that counts the number of overdose events for every day of the week. One can clearly see that more overdose events occur on Friday or the weekends than the rest of the week.

3.6. Month of year

The spring and summer months have many more overdose incidents (~1,800 incidents) than the colder winter and fall months (~1,300 incidents) as can be seen in Figure 10. These findings are consistent with other studies that have looked at seasonality in overdoses (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9124554/).

The observed seasonality in opioid overdose incidents is consistent with the hypothesis that opioid overdose events are correlated with changes in behaviors that differ between the warmer and colder months. An example might be increased incidences of partying and non-opioid drug use leading to opioid overdoses due to fentanyl contamination of other drugs.

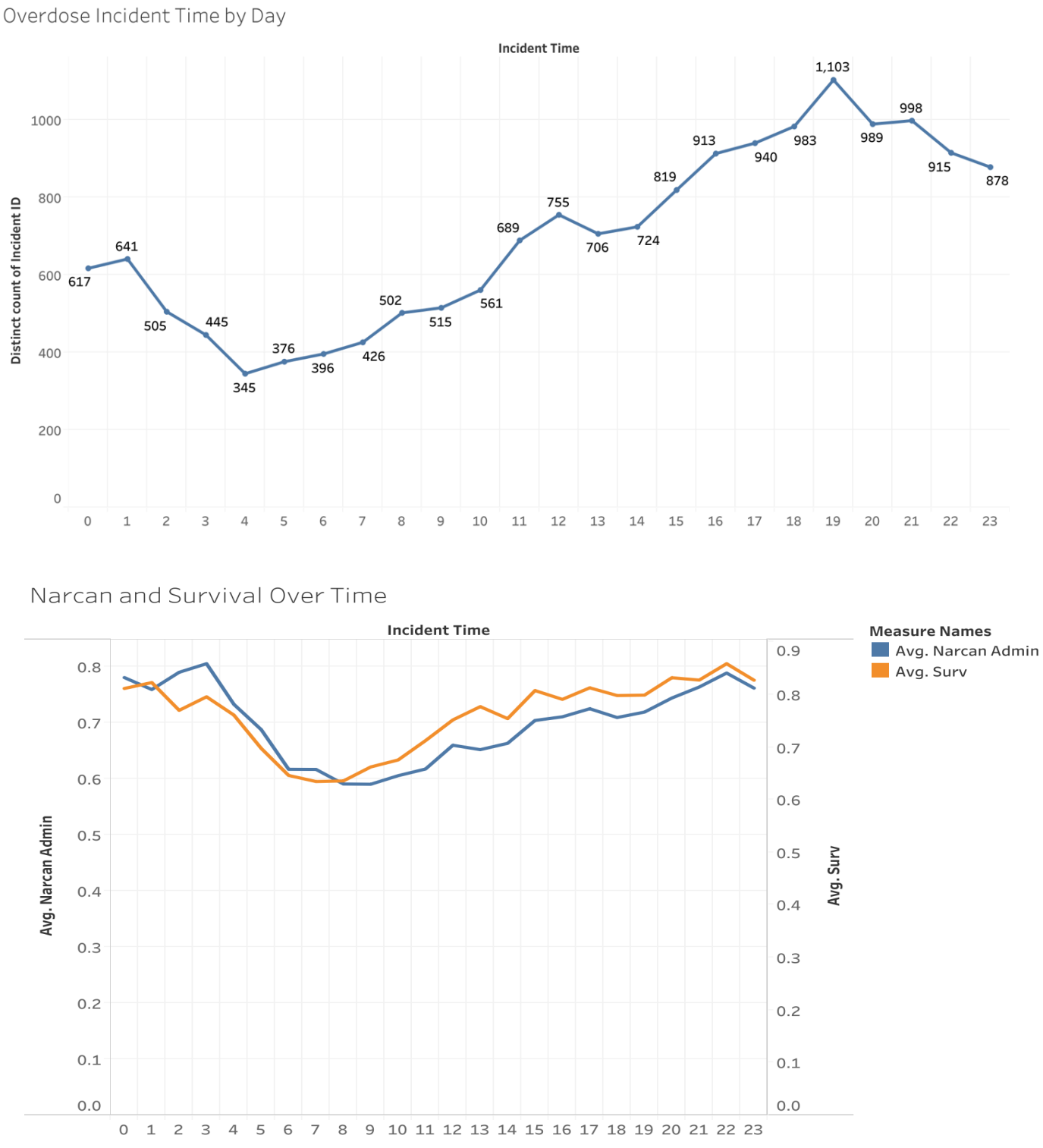

Figure 8 - Top graph overdose incident time and bottom graph shows survival rate and average Narcan administered.

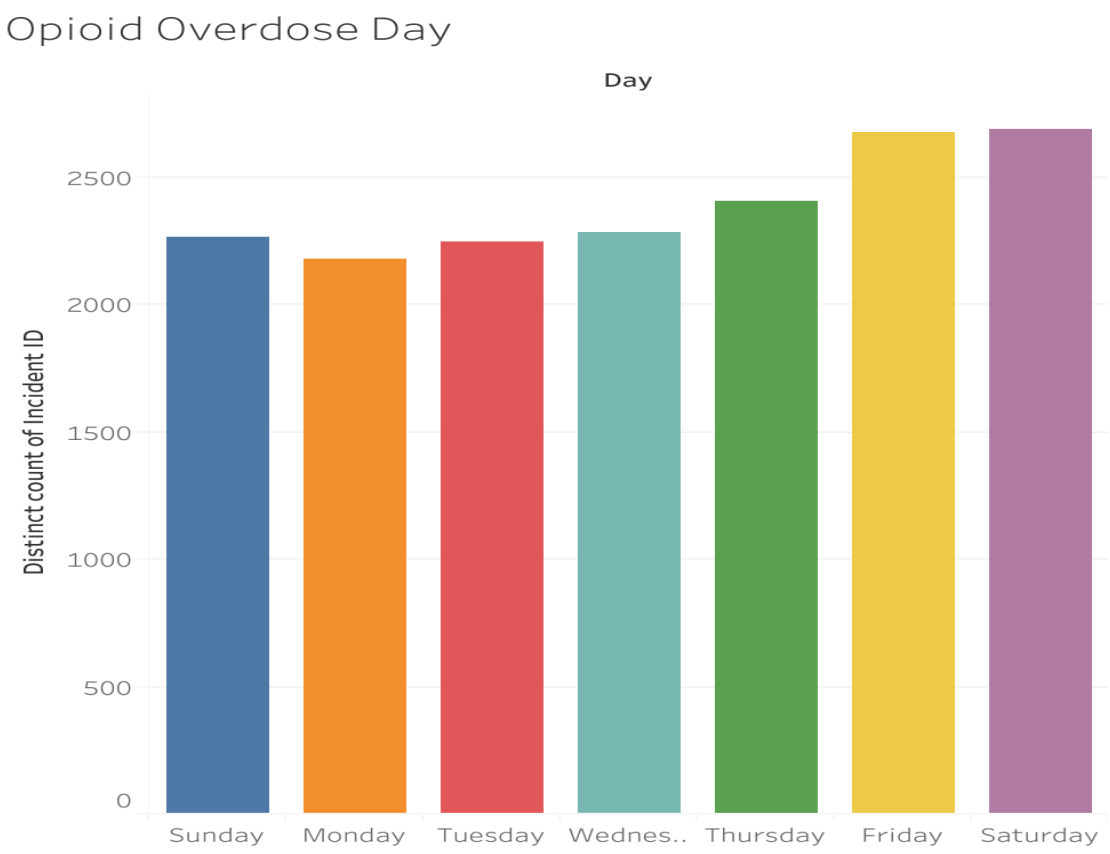

Figure 9 - Opioid Overdose Count by Day

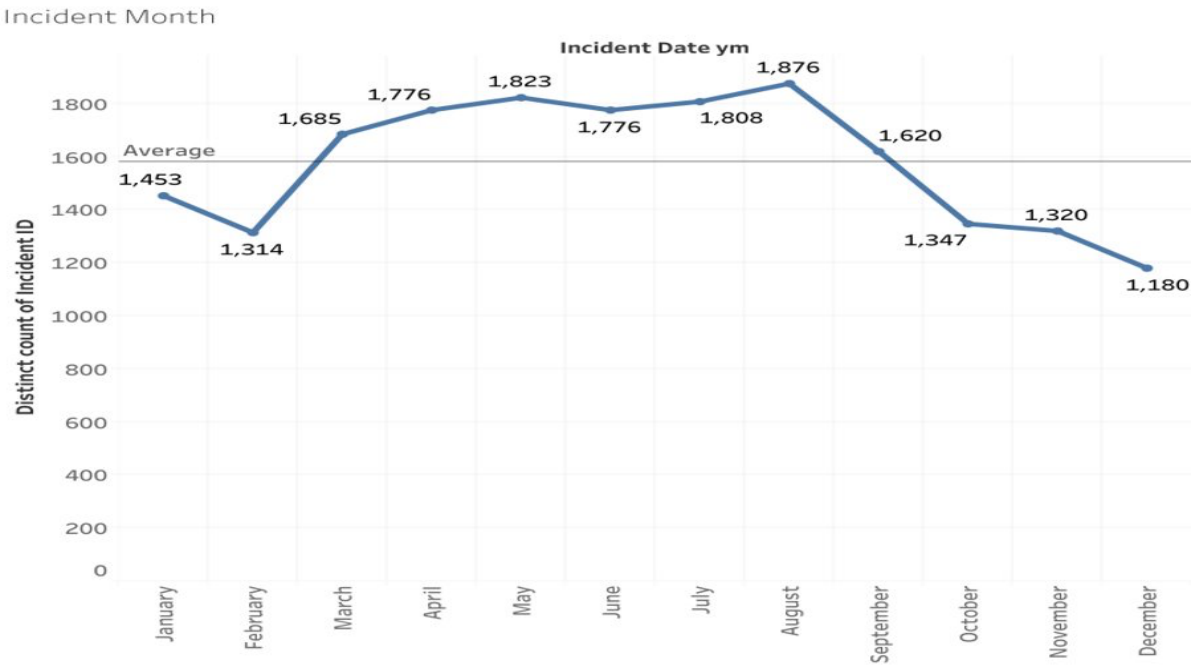

Figure 10- Incident Count by Month

3.7 Month of year with Naloxone

Survival remains strong throughout the months when Naloxone is used (Orange line) as shown in Figure 11 below. When Naloxone is not used (Blue line), there is a decrease in survival in the colder months.

3.8 Naloxone Administration per Month

The warmer months tend to have a higher percentage of Naloxone administration compared to the colder months as shown in Figure 12. July at 70.5% as the highest compared to January 66.1% as the lowest.

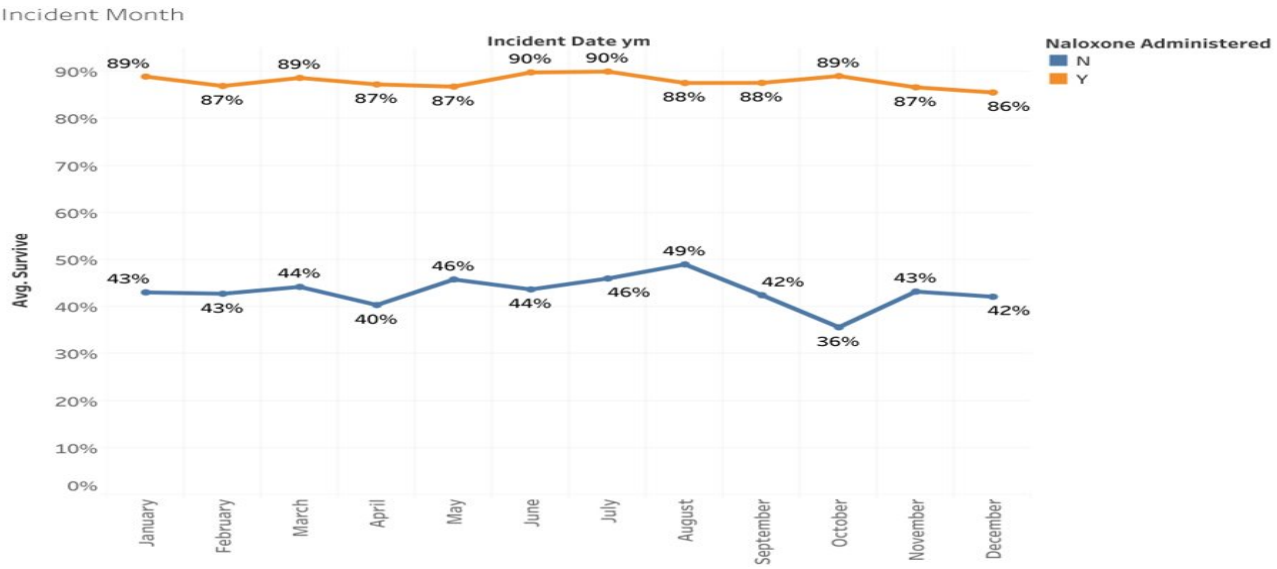

Figure 11- Survival Rate per month with and without Naloxone

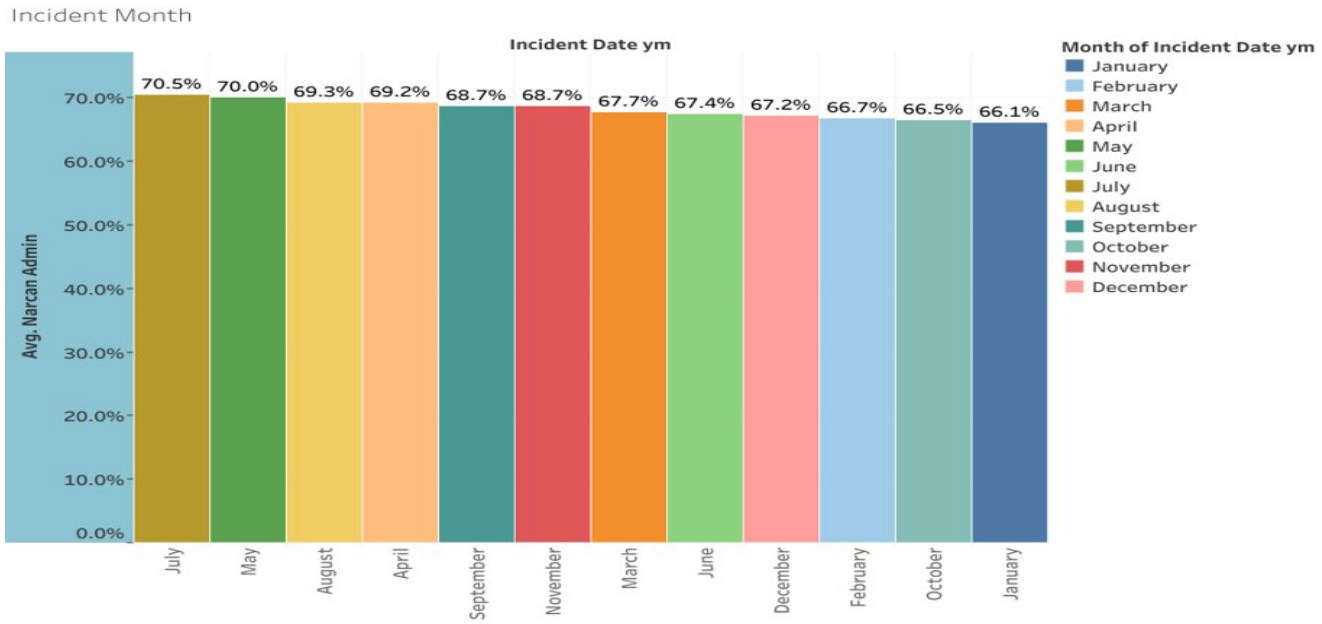

Figure 12 - Naloxone administration percentage per month

3.9 Seasonality of overdoses

The seasonality of overdoses occurring in the warmer months is mostly consistent across years. State death from overdose analysis all drugs. The CDC dataset had the counts of deaths by drug overdoses by year. However, not every state specified the drug type. We took this dataset and combined it

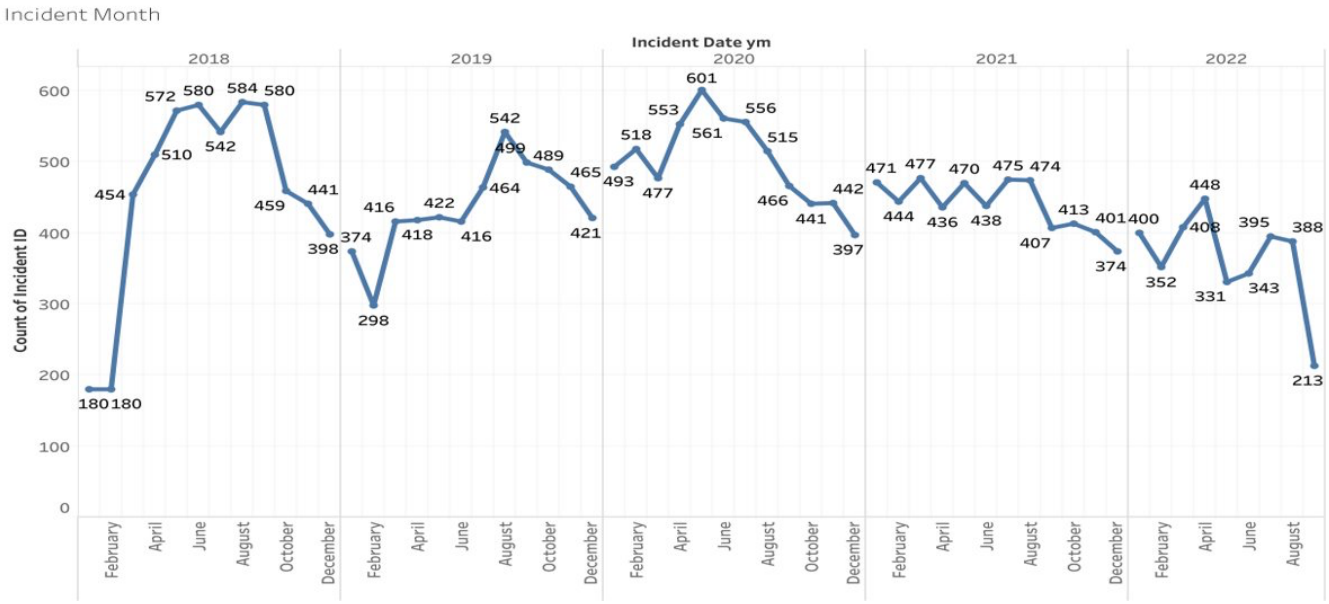

Figure 13 - Seasonality of overdoses

4. Results and Discussion

This study examines several datasets from three sources (PAGOV, CDC and Census), which were merged, and aggregation were made. The final dataset was uploaded to Tableau Public to create various visualization as presented in EDA section.

The visualizations revealed several insights. First, Naloxone is very effective keeping survival rate at ~80% or higher compared to ~40% when Naloxone was not administered.

Second, age range analysis shows that 30-39 age range group had most overdose incidents (~6,000), and males were more involved in overdose incidents.

Third, Fentanyl was found to be more deadly than other opioids, reducing survival rate by 20%. This suggests that the type of the drug is a potential feature in a machine learning model.

Fourth, day analysis shows that overdose incidents occur more often during the weekends (Friday and Saturday), and time analysis shows that more incidents occur at night. These findings suggest that many of the overdose events are related to people holding a full-time position.

Finally, monthly analysis indicates that there is seasonality of overdoses, meaning there are more incidents during Spring and Summer. These findings could be beneficial for emergency services, such as police, in order to prepare for busy hours, and therefore reduce the time it takes to arrive at an overdose event which can potentially reduce death rate.

5. Conclusions

5.1 Recommendations

Based on our analysis, we recommend public health officials create a campaign to spread awareness and training for recognizing signs of an overdose, and for using Naloxone on a victim. Further out analysis suggests that many overdose events could be from fentanyl, public health officials should conduct trainings in places where these accidental overdoses are likely to happen. For example, security at nightclubs and outdoor events.

Additionally, we found a large discrepancy in Philadelphia County between the EMS dataset and PA.gov dataset where Philadelphia had a tendency to rely on EMS for overdose encounters over police. Therefore, we recommend that police are trained on a given Naloxone, and that they can respond to overdose events more often.

5.2 Future Work

We are interested in expanding our Opioid Overdose analysis to include machine-learning based prediction of several variables of interest including:

1. Survival (Incident-Level)
2. Naloxone Administration
3. Opioid Overdose Events (County-Level)
4. Drug Overdose Events (State-Level)

Using the incident-level data, we can predict both Survival and Naloxone Administration. Because Naloxone use is so highly predictive of survival, it may be more informative to understand what factors lead to high or low values for Naloxone administration.

The county and state level data can also be used to predict the number of overdose events that will occur in the future. If effective, this analysis would provide insight into upcoming demand for services and interventions aiming to reduce opioid overdose lethality.

During this project we prepared a preliminary incident-level dataset for machine-learning analysis. Many of the insights we gleaned from the exploratory data analysis can be leveraged in to create informative features for machine-learning. An example includes computing a variable that indicates whether multiple drugs were involved in the overdose incident

In the future, we also aim to use clustering algorithms on the county-level data to identify similar counties and explore whether this similarity measure affects the target variables of interest.

References

1. Center for Disease Control and Prevention (CDC), *U.S. Overdose Death In 2021 Increased Half as Much as in 2020 – But Are Still Up 15%*, May 11, 2022, https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2022/202205.htm.

2. United States Drug Enforcement Administration (DEA), *Fentanyl*, April 2020, https://www.dea.gov/factsheets/fentanyl.

3. National Institute on Drug Abuse (NIDA), 2022, January 11. Naloxone DrugFacts. Retrieved from https://nida.nih.gov/publications/drugfacts/naloxone on 2022, November 26.

4. CDC, Provisional County-Level Drug Overdose Death Counts, October 26, 2022. https://www.cdc.gov/nchs/nvss/vsrr/prov-county-drug-overdose.htm

5. Pennsylvania State Police, Overdose Information Network Data CY January 2018 - Current Monthly County State Police, December 27, 2018, https://data.pa.gov/Opioid-Related/Overdose-Information-Network-Data-CY-January-2018-/hbkk-dwy3

6. Census, State Population Totals 2010 - 2019, December 16, 2021. https://www.census.gov/data/datasets/time-series/demo/popest/2010s-state-total.html#par_textimage_1873399417

7. SODA Developers, Socrata API, https://dev.socrata.com.

8. Census Microdata API, https://api.census.gov/data.html

9. Census, Census Microdata API User Guide, July 22, 2021. https://www.census.gov/data/developers/guidance/microdata-api-user-guide.How_To_Create_API_Queries.html