For office use only
T1 ________________
T2 ________________
T3 ________________
T4 ________________

Team Control Number
1900577

Problem Chosen
C

For office use only
F1 ________________
F2 ________________
F3 ________________
F4 ________________

2019
MCM/ICM
Summary Sheet

The Gravity of the Opioid Crisis

The United States is in the midst of a national crisis due to the extreme abuse of opioids all across the country. More than ever before, users of all ages and demographics are becoming addicted. To explore future impacts of this drug epidemic, we model and characterize the spread of substance abuse.

We use NFLIS drug report data 2010-2017 to develop a multivariate analysis of the spread of drug use in and between counties of Kentucky, Ohio, Pennsylvania, Virginia, and West Virginia. We base the development of the drug spread model on three main factors:

i. Drug-use influence
ii. Current trend in drug reports
iii. Pertinent socio-economic factors

We identify the current trend of drug reports by implementing a quadratic-weighted linear regression on existing county drug report data across time. To characterize the nature of drug-use influence on a county, we define a county's drug influence factor as the density of drug reports per area. We find the origins of specific opioids and determine the drug identification thresholds based on the influence factor. Six counties crossed our determined threshold for our simulation from 2018 to 2025, indicating that the epidemic is increasing in intensity.

Using the principles of geographic gravity, we establish an inverse relationship between influence on other counties and distance as a weighted factor of their existing trend. We validate our model as a better predictor of the following year's drug reports than the data from the previous year.

We then find associations between drug reports and U.S. Census socio-economic factors over time. With the highest correlated socio-economic data, we calculate a multivariate linear regression of the drug reports based on time. This adjusted model shows a 12.2% decrease in residuals of our predicted data against the real data, compared to the baseline.

We use strategies from other parts of the United States to make improvements in accommodating drug users. Using factors such as drug education and recovery centers, we estimate a reduction in overall drug consumption.

Contents

1 Introduction

Drug abuse is an issue that the world has been plagued with for centuries. The first study of morphine addiction was done in 1875, identifying the key factors for a user to become addicted to a substance [13]. In the past several decades in the United States, the opioid epidemic has increased at an alarming rate. Currently, an average of 130 citizens of the U.S die of an opioid overdose daily [4]. Understanding the spread of this epidemic can be used to inform government policy to get control of this crisis.

1.1 Problem Summary

For the U.S. government, it is a challenge to enforce anti-drug laws, especially amid the national crisis that is happening. The Drug Enforcement Administration (DEA) wishes to see if there are factors that contribute to the spread of opioid incidents between 5 states in the eastern United States — Ohio (OH), Kentucky (KY), West Virginia (WV), Virginia (VA), and Pennsylvania (PA).

We use data analysis to build a model that **describes the spread** of opioid cases in these states, with the ability to identify any possible locations where a specific drug might have originated. We then set a **threshold** which signifies an unsafe level of drug use in the county, predicting where this will occur in the future. We then add U.S. Census data to our model, implementing **socio-economic factors** into the model. When then use this model to identify **strategies for countering the opioid epidemic**, and test the effectiveness of these strategies.

1.2 Data Sources

Our model is informed by 8 years worth of drug identification counts, 2010 to 2017, for narcotic analgesics and heroin from the National Forensic Laboratory Information System (NFLIS). We also derived geographic data from the NFLIS dataset [9]. The model is conditioned with 7 years worth of data from the U.S. Census Bureau, 2010 to 2016, that represents a common set of socio-economic factors.

1.2.1 Data Cleaning

The census data provided had missing and partially filled in data that would have been challenging to effectively utilize. We did the following to sanitize the data-set:

- Remove factors that were not measured at all (represented by the symbol (x), such as HC04_VC03)

- Remove factors that were only measured in certain years, such as computers and internet use (HC01_VC216), as trends across multiple years would be less present and more susceptible to potential outliers.
- Remove factors that had incomplete data for all the counties, often represented by the symbol "*****". Incomplete data for a factor would inhibit the creation of a proper model for drug spread, as there could be hidden trends that are not apparent because of missing data.

1.3 Existing Models

Many comprehensive studies and models of drug spread and impact have been executed in the past. Many of these models focus on the following ideas:

- Illicit drug users are broken into 3 groups: light users, susceptible users, and dealers. Each of these groups may enter any of the others through remission, death, or influence [7].
- Set a threshold quantity to how many new drug abusers an average dealer or light user will generate over their lifetime [12].
- Predict future changes by modeling the previous waves of the opioid epidemic. The first wave began in 1999 with prescription opioids, the second started in 2010 with a rise in heroin use, and the third commenced in 2013 with synthetic opioids becoming more common, as can be seen in Figure 1.

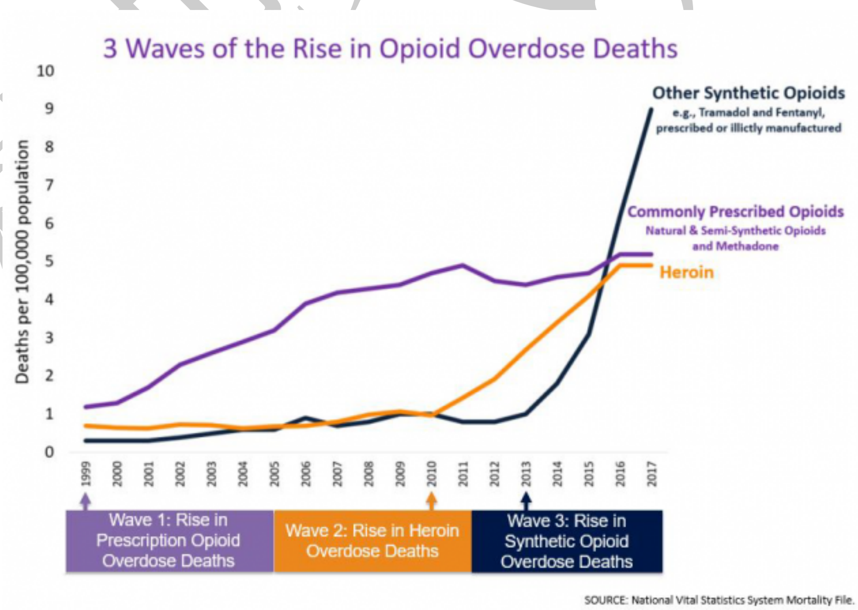


Figure 1: The three waves of the opioid epidemic [2]

1.4 Our Model

For the purpose of determining an effective model to describe the spread of opioid incidents, we characterized the **nature of drug-use influence** on each county, both externally and internally. External influence on a county is determined by the density of drug reports of nearby counties, inversely related by geographic distance. Internally, a county's influence comes from a **weighted trend analysis** of total drug reports based on category.

With the addition of socio-economical factors, our model took a **multi-variate approach**. We determined which factors had the greatest impact on the trends in drug use, and applied a predictive fit to determine how changes in these factors influence a county. This will allow us to predict how the spread of drug reports will change in the future, creating the opportunity to mitigate the opioid epidemic before it grows more out of control.

2 Background

2.1 The Opioid Epidemic

The opioid epidemic is a strange phenomenon of drug addiction where people begin taking prescription opioids for pain management related to a medical issue [5]. By nature, opioids are a highly addictive form of drug, and a user can quickly become dependent due the euphoric feeling they experience while taking them. Users quickly build up a tolerance to this drug, and require more or different types for them to continue to be effective [2]. This can cause issues, especially in medical treatment, as a user's body will no longer respond to certain forms of this drug, causing great pain.

2.1.1 Classes of Opioids

There are three main classes of opioids:

- *Opiates (non-synthetic opioids)*: codeine, morphine, opium, heroin
- *Semi-synthetic opioids*: Hydrocodone, oxycodone, buprenorphine
- *Synthetic opioids*: Fentanyl, butorphanol, methadone, propoxyphene

This distinction in type of opioid different users will use each category, which can result in varying overall trends. Although heroin is considered an opiate, it is processed from morphine and typically placed in its own category for analysis.

Strength of opioids are compared with an Oral Morphine Milligram Equivalent (MME) Conversion Factor. For example, the synthetic opioid tramadol has a 0.1 MME conversion factor, meaning tramadol is 10 times stronger than the equivalent mass of morphine. The trend is that synthetic opioids are much stronger in their effects than opiates and semi-synthetic opioids.

2.1.2 Common Reasons for Substance Abuse

There are several factors that researchers generally attribute to increasing the likelihood of addiction [6]:

- Mental Health Problems
- Career, home, school, or friendship issues
- Proximity to other drug users
- Past traumatic events

3 Nomenclature

Symbol	Definition
$G_{d,i}$	Gravity of Influence for drug d and county i
I_i	Influence factor of a county i
P_i	External influence on county i
R_E	Radius of the Earth
r	Distance along the earth in kilometers
I_{max}	Largest influence factor in the data set
i	County identifier
m	Time derivative of linear best fit in drugs vs time
$r_{i,j}$	Distance between counties i and j (m)
S_i	Set containing every county identifier excluding i
ΔDrug	Predicted change in drug reports
$\Delta\phi$	Difference of latitudes in radians
$\Delta\lambda$	Difference of longitudes in radians
ϕ_1	Radian measure of latitude 1
ϕ_2	Radian measure of latitude 2
μ_I	Median influence factor

Table 1: Variables and functions

4 Assumptions

- Illicit drug use is primarily influenced by human interaction. Like culture, drug use will spread geographically, more strongly affecting nearby locations than those far away [8].

- The drug report data is representative of overall drug usage in a county or state.
- Smaller (rural) populations will be more susceptible to change based outside influences than larger (urban) populations.
- New types of drugs generally appear in waves, which is uncertain to predict based on the drug report data provided. For the purpose of this model, we assume that no new drugs will be introduced to the population.
- Without external influence, an addicted population will continue their current trend of drug use [10].

5 Model Development

In order to quantify and predict the rate of opioid spread between these 5 states, it is necessary to justify the idea of influence. One of the largest factors that increase the likelihood of drug addiction is proximity to to other drug uses. A county, in some capacity, will be more strongly affected by the drug use of counties nearby than by those that are farther away. To quantify the density of drug use in an area, we define the **influence factor** for a county i:

$$I_i = \frac{\text{Number of Drug Reports in County } i}{\text{Area of County } i} \tag{1}$$

Assuming that the count of drug reports in an area properly represents the actual drug use in that area, we can use this factor as a standardized way to measure drug use in a county. The hypothesis that the spread of the opioid epidemic is correlated to drug reports density rather than strictly the magnitude of these values parallels work previously done in epidemiology [15].

5.1 Geographic Gravity

The next step is to determine the impact of these influences on neighboring counties. We begin with the idea of gravity. In physics, the force of gravity on an object is proportional to the product of the mass of each object divided by the distance between the two objects squared:

$$F_g \propto \frac{m_1 m_2}{r^2} \tag{2}$$

Much like how this quantity is dependent on the inverse square of the distance between to objects, the drug influence of a county on another decreases with the distance from the county as well. Yanguang Chen's work in spatial analysis in geography and social physics upholds this idea, stating that a gravitational model illustrates the interaction between counties, while a exponential decay is more suggestive of local rather than spatial interaction [8].

In order for our model to measure distance, we need to utilize external geographical data. We used the FIPS codes provided in the data to match each county with a latitude coordinate, longitude coordinate, and county land area from the U.S. Census Bureau [9]. Although these coordinates point to the center of each county (which may not be completely representative for an unusually-shaped county's location), it provides us with a method to determine distances between counties. We now use the Haversine formula to determine the distance between any two counties, which accounts for Earth's curvature to provide a more accurate value than Cartesian estimation:

$$r = 2R_E \arcsin(\sqrt{a}), \quad a = \sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\Delta\lambda}{2}\right)$$

Equipped with a measure of a county's drug influence, along with the distance between each county, we define the **External Influence Potential** on county i as:

$$P_i = \sum_{j\epsilon S_i} \frac{I_j}{r_{i,j}^2}, \quad S_i = \{j \in \mathbb{N} \mid 1 \leq j \leq n;\ j \neq i\} \tag{3}$$

5.2 Model Construction

Our opioid spread model focuses on two main ideas: a county's influence (both internal and external) along with the current trend of county drug reports with respect to time. Although it would have been beneficial to have more extensive historical data, we utilized the 8 years' worth of NFLIS to determine the current drug report trend in a county with reasonable accuracy.

To find this trend, we perform a weighted trend analysis for each county. We apply a linear fit to the total county drug reports with respect to time using a quadratic weight of $(year - 2009)^2$ for each point up to the current year. This allows us to take into account a county's past drug report use, while also placing a much higher significance on more recent trends in the data. The time-derivative of this linear fit m_i is calculated for every year and county for which there were at least two years' worth of drug report data.

In order to combine influence factors with current trends in the data, we introduce a normalization coefficient for the influence values. This coefficient serves to normalize these influence values, along with providing a basis for how much the external influence potential will affect the change in drug reports for that county. Based on models of social physics, a county with a large influence will be much affected much less by outside influences than a smaller county, thus we base the scale on the ratio between the median and current influence factor values [8]. This leads to the formula for the change in drug reports for county i in a given year, seen in Equation 4.

$$\Delta\text{Drug}_i = \left[\left(\frac{\mu_I}{I_i I_{max}}\right)\sum_{n \epsilon S_i}\frac{I_n}{r_{n,i}^2} + 1\right] * m_i \tag{4}$$

The ΔDrug is calculated in MATLAB and iterated over for every county for every time step.

5.3 Extensions of the Model

In order to apply this model to specific drugs or drug categories, rather than solely the total drug growth in a county, we can recalculate the influence factor values using the drug report count for that specific drug, rather than the combined total value. The trend values, m_i, would also need to be recalculated in a similar way. However, the smaller drug report values when broken down by type of opioid is a significant hurdle for our model as it is susceptible to low numbers of counties reporting the drug.

5.4 Model Validation

To test our model, we predict drug report values for a year one time step into the future, and compare it to existing data for that year. For example, we use 2014 data to predict 2015 drug report values and then compare the prediction to the actual 2015 drug reports. A reference line with slope one going through the origin represents the ideal predictor, where every predicted value equals the actual value in the following year.

Qualitatively, we can see if the predictions fall close to the line and whether the predictions trend high or low. Quantitatively, we use two measures to judge the quality of the prediction. A least-squares line of best fit for the prediction versus the actual data produces a slope with values close to one. A slope lower than one informs us that our model predictions trend too low and a slope higher than one says our model predictions trend too high. We also calculate the residuals from the points to the reference line of slope one. The sum of residuals allows us to give a magnitude of the closeness of the prediction.

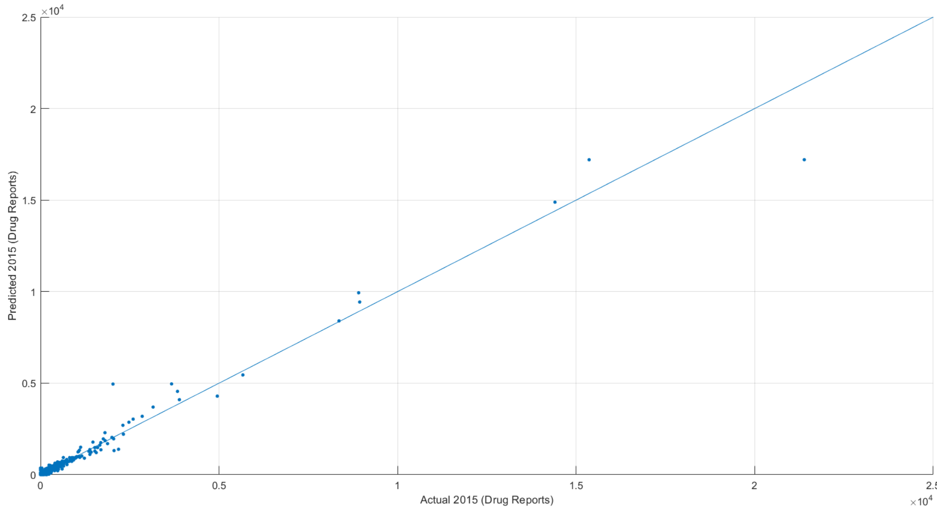


Figure 2: County drug report data for 2015 versus model prediction based off 2014

We need to establish a baseline to determine if our predictions were effective. The best choice is to compare the past year to the next year that is being predicted. Our model computes ΔDrug_i and adds it to the past year's drug reports. If our model is to be effective, it would show an improvement over the baseline measure.

As illustrated by Figure 2, our model predicts better than the baseline in the year 2015, with best fit slopes of 0.9276 and 0.9262 respectively. The maximum our model could possibly improve the slope is by $1 - 0.9262 = 0.0738$. Our model improves the slope by 0.0014 which is a 1.9% increase for 2015, and shows an improved slope from the baseline for the years 2013 through 2017 as well. However, the prediction was only marginally better than the baseline. The larger numbers of counties with drug reports in the later years helped our model better predict the directional trend. Likewise, The baseline sum of residuals for the 2015 example is 47408 drug reports. Our model has a reduced sum of residuals at 45539. This is a 3.94% decrease.

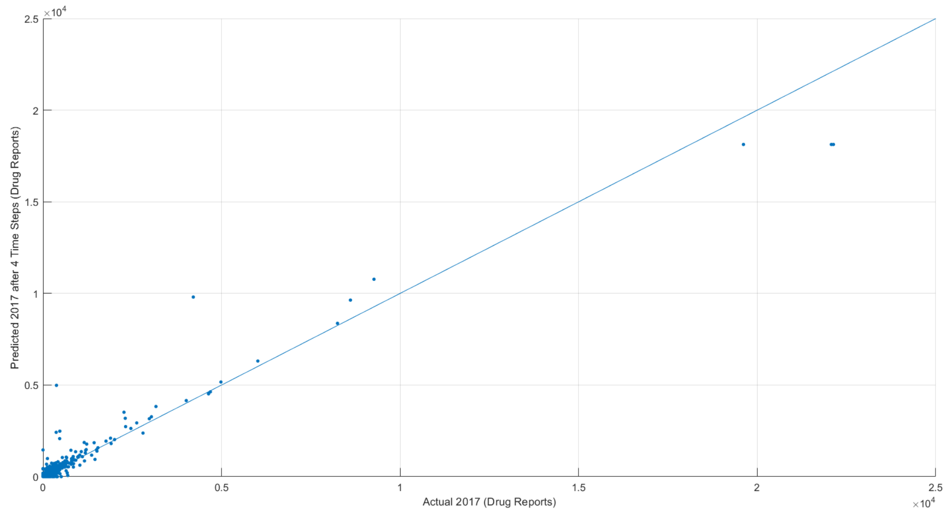


Figure 3: Model Prediction of 2017 Drug Reports after 4 Time Steps

The next step is to increase the number of time steps our model predicted out from one year. We simulate four years, starting with 2013 drug report data and predicting 2014 through 2017 drug reports, basing each subsequent time step on the previous prediction. This is again compared to the baseline of comparing the 2013 drug reports to the drug reports of the years our model predicted. As expected, the model decreased in accuracy every time step, which can be seen in Figure 3. Regardless, our model beats the baseline at every time-step. By 2017, the predicted 2017 from 2013 data versus the real 2017 data had a linear best fit slope of 0.913 while the baseline (Figure 10) had a slope of 0.786 (16.6% increase). The sum of residuals from this simulation is 82835 drug reports compared to 93656 drug reports for the baseline (11.5% decrease).

5.5 Application

Based on the characteristics of the data and our model, the epicenters of the opioid epidemic are in counties with cities that consistently have a high influence factor. The influence of these counties radiates outwards and influences the counties nearby. However, the real issue is counties that already have high numbers of drug reports continue to have high numbers of drug reports for years. The U.S. government should also be concerned specifically with the synthetic opioids like fentanyl and its analogs. The data suggests that synthetic opioids are spreading at a faster rate than the natural opiates. Synthetics have a strength level 400 to 6000 times higher than that morphine, leading to an much higher rate of overdose [3].

5.5.1 Predicting Points of Origin

Using an extension of our model, we are able to determine the location of where a specific drug was most likely to have started. Equation 5 introduces the notion of **Gravity of Influence** for any specific drug d in county i.

$$G_{d,i} = \sum_{n \epsilon S_i} \frac{I_{i,d} I_{n,d}}{r_{n,i}^2} \tag{5}$$

For any drug, the county with the maximum G_d value in the initial outbreak year has the highest likelihood to be its origin of use. As an example, we will look at the upsurge of 3-Methylfentanyl, which first appears in the NFLIS data in 2016. Erie County, Ohio has the highest G_d value in 2016 for 3-Methylfentanyl, indicating a high probability that this is where the drug started. Looking at the spread of 3-Methylfentanyl drug reports in Figure 4 supports this hypothesis, as the region of northern Ohio (the location of Erie county) has the most significant concentration of this drug in the year of outbreak.

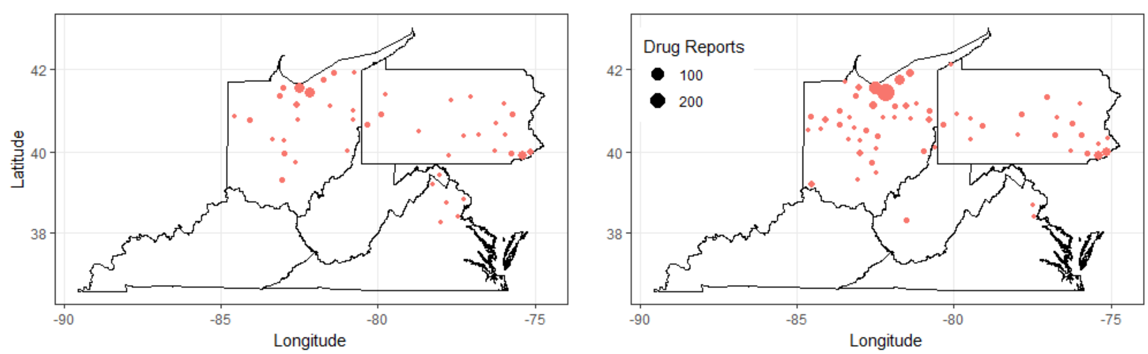


Figure 4: 2016 (Left) and 2017 (Right) county drug reports for 3-Methylfentanyl. Map data from GADM [1]

5.5.2 Drug Identification Threshold

For this model to be of use to help with the opioid epidemic, we determine a threshold which signifies a point of concern for the government. We know that there already exist significant issues with opioid overdose in high drug use counties. So, we chose our threshold to be the 90th percentile of influence values from 2017. By basing our threshold on influence values, this ensures that each county has their own individualized drug report threshold scaled by size.

Using our model, we simulated the spread of total drug use through the year 2025. In 2017, where we began this simulation, there were already 11 counties above the threshold (Lake and Cuyajoga County, OH; Philadelphia, Delaware, and Beaver County, PA; Franklin, Montgomery, and Hamilton County, OH; Arlington, Roanoke, Fairfax City County, VA). Six more counties rose above the threshold when simulating out to 2025. Unfortunately, none of the counties which were already above the threshold in 2017 sank below it again. However, this is understandable, as the opioid epidemic is modeled to expected to continue spreading without external intervention.

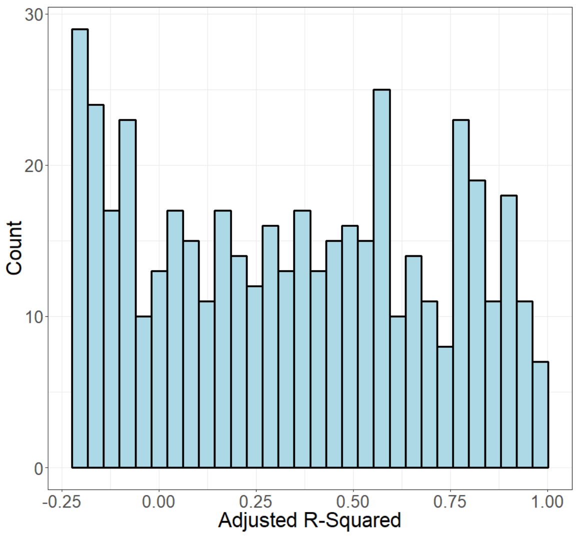
Year	Counties
2018	Fairfax, VA and Kenton, KY
2020	Henrico, VA
2021	Stark, OH
2023	Summit, OH
2024	Lorain, OH

Table 2: Counties rising above threshold in simulation from 2018-2025

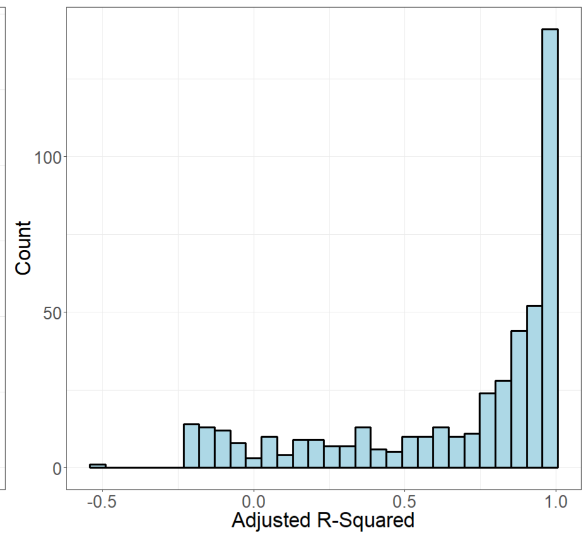
5.6 Important Socio-Economic Influences

Outside of influence factor, there are other components that may lead to the spread of drug use. As part of the included data, we were given U.S. Census data for a wide range of socio-economic factors. In order to quantify which factors best correlated with trends in drug use, we found the R^2 values of the linear fit for the difference between the total county drug reports and the value of each factor with respect to time. Then, we calculated the median value of these R^2 values for each specific factor across all counties.

Using these median values as a baseline, we can examine which factors have the closest trends to total county drug reports. By plotting the distribution of the correlation values for a each factor, two main trends emerge, which are depicted in Figure 5. For factors with a high median value, the distribution is heavily skewed towards one end of the scale of possible R^2 values (-1 or 1), indicating there is a high number of counties with which these factors could model the trend in drug use. For low median values, the distribution is nearly uniform, indicating that there is little correlation between this factor and the trend in drug use, and that it would not be a good candidate for the model.



Low median R^2 value



High median R^2 value

Figure 5: Trends in R^2

Table 3 indicates the factors with the highest median R^2 values located in the cleaned data along with their Census metadata codes.

Factor	**Metadata Code**	**Median R^2 Value**
Population of > 1 year residency	HC01_VC117	.861
Civilian Population with Veteran Status	HC01_VC98	.790
Estimate of Associate's Degrees	HC01_VC89	.744

Table 3: Factors and their median R^2 values

5.7 Adjustments

To take into account these correlated socio-economic factors into our model, we turn to a multivariate approach to find the current trend of drug reports with respect to time. Instead of using a weighted linear regression to find the value of m_i, we perform a polynomial regression fit on the total drug reports with respect to time and our most correlated socio-economic factors. Equation 6 indicates the form of this equation fit for county i with socio-economic factors x_1 through x_n in year y.

$$\text{Total Drug Reports}_i = m_1 y + m_2 x_1 + \cdots + m_{n+1} x_n \tag{6}$$

From this, we obtain the value $m_{i,y}$ for county i in year y by inputting the values for the change in each of the factors with respect to time at year y.

$$m_{i,y} = m_1 + m_2\frac{dx_1}{dt}(y) + \cdots + m_{n+1}\frac{dx_n}{dt}(y) \tag{7}$$

The overall equation for the opioid spread model remains the same, with the new factor $m_{i,y}$ replacing the previous trend values. After trying different combinations of implementations the factors in Table 3 into our model, we found the model that best fit the data was that using just the factor with the greatest median R^2 value (HC01_VC117). This makes sense as it was the factor that was, on average, best representative of the trends in drug reports for each county.

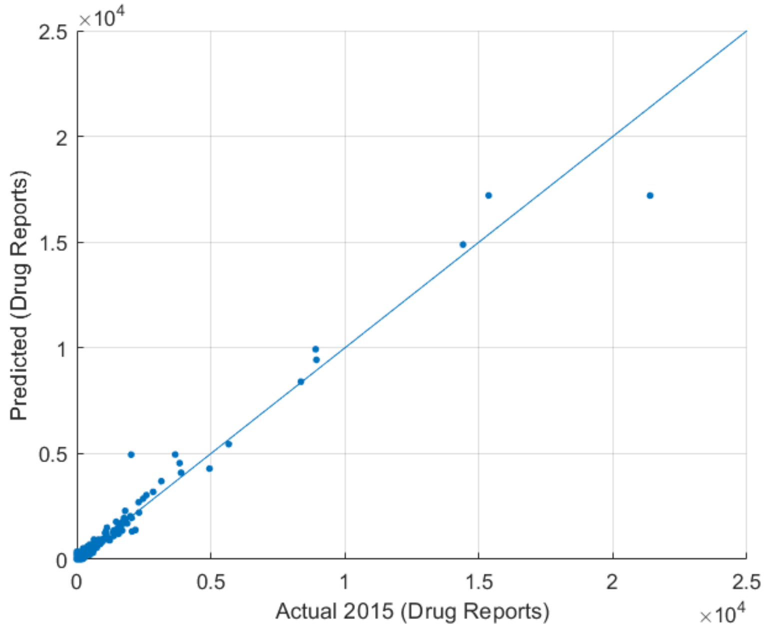


Figure 6: County drug report data for 2015 versus non-adjusted model prediction based off 2014

With the adjusted model, as shown in Figure 7, we examine how the addition of the socio-economic factor influenced our predictions. After using the adjusted model to predict 2015 drug report spread based off of 2014 data, the slope of the best-fit line stayed only slightly increased (to .9288 from .9276). But, the sum of the residuals with the adjusted model was 41631, which is a 12.2% decrease from the 47408 drug reports in the baseline, and a 3.9% decrease from the 45539 drug reports in the original model (Figure 6).

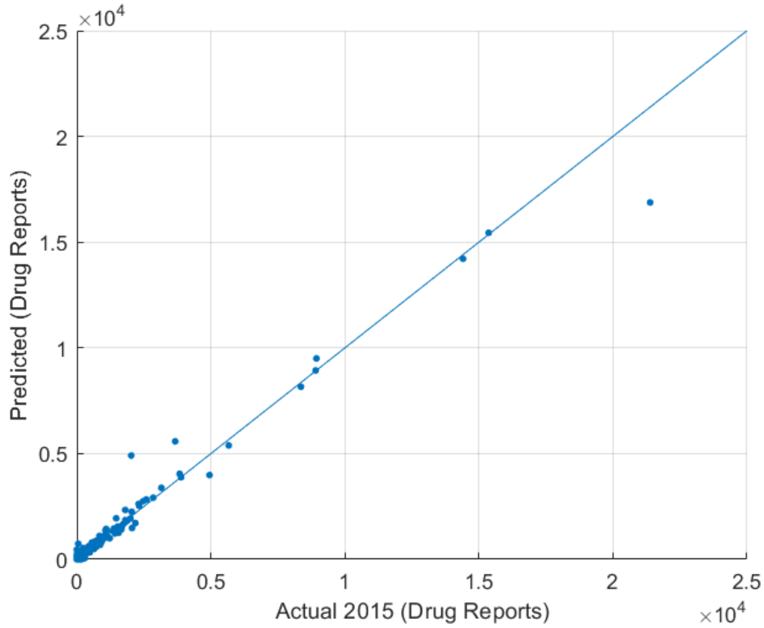


Figure 7: County drug report data for 2015 versus adjusted model prediction based off 2014

6 Results

6.1 Part I: Origin Prediction and Threshold

The model can predict the point of origin of a specific drug based on each county's gravity of influence (G_d). For example, the origin of 3-Methylfentanyl was found to be Erie County, Ohio in 2016. This was determined by finding the maximum G_d out of all the counties for this specific drug.

The drug identification threshold level that we determined was the 90th percentile of influence values for counties in 2017. There are already 11 counties above this threshold, and the model predicts that cross this threshold from 2018 to 2025. Unfortunately, none of the counties currently above the threshold are predicted to go below it without changes in governmental policy or actions against opioid addiction.

6.2 Part II: Socio-Economic Adjustments

We found that population of users with residency of greater than 1 year, civilian population with veteran status, and number of Associate's degrees were the socio-economic factors that most closely resembled the drug report data. By conditioning our model based on the most influential factor, the model prediction residuals decreased by 12.2% compared to the unconditioned model.

6.3 Part III: Strategy for Countering the Crisis

In the past several years, there have been numerous efforts to curb the opioid epidemic. Simply arresting drug users does not help the problem, as this simply makes recovery more difficult for these individuals and it is more challenging to find a job with a criminal record. A data-driven solution to this problem is necessary to see a significant change.

From part I, we see that counties with large influence values are not easily affected, but counties with low influence values can grow exponentially under the right conditions. Early detection of the emergence of new drugs up would be an effective strategy for countering crisis with the correct approach.

From part II, we notice that there are several socio-economic factors that seem to be strongly associated with trends in drug use. One of these factors is the number of Associates Degrees. This may be due to there being less opportunities, challenge finding jobs, or other struggles that are not as present with higher-level college degrees. A solution to this could be to implement a program to incentivize Associate graduates to pursue a 4-year degree upon graduation. By allowing the graduated to continue on in their education, they may be able to get more opportunities than they otherwise would with their associates degrees [14]. Another strongly associated factor was the population of veterans in a county. This could be because veterans have an increased chance of having some form of past trauma, which is a known cause of drug addiction [6]. A possible solution to this would be to offer an addiction recovery course could offer a low cost recovery alternative or supplement, and veterans could be eligible to participate in this course at no or reduced cost.

One problem with the current opioid crisis is that counties who already use the most drugs will continue to use the most over time. In order to mitigate this, the government should must changes into the counties with the most annual drug reports. By implementing programs that are self elective, like recovery and educational centers, populations will have access to help they need. And, these programs have already been shown to be effective. The Chinese Needle Exchange Program and Methadone Maintenance Treatment centers saw a significant reduction in users after the program had been implemented [11]. These programs were so effective in the private sector in China, the government decided to open four of their own in 2004, to which they expanded to 558 by 2008. Implementing centers like these in the United States would likely follow a similar trajectory of improvement.

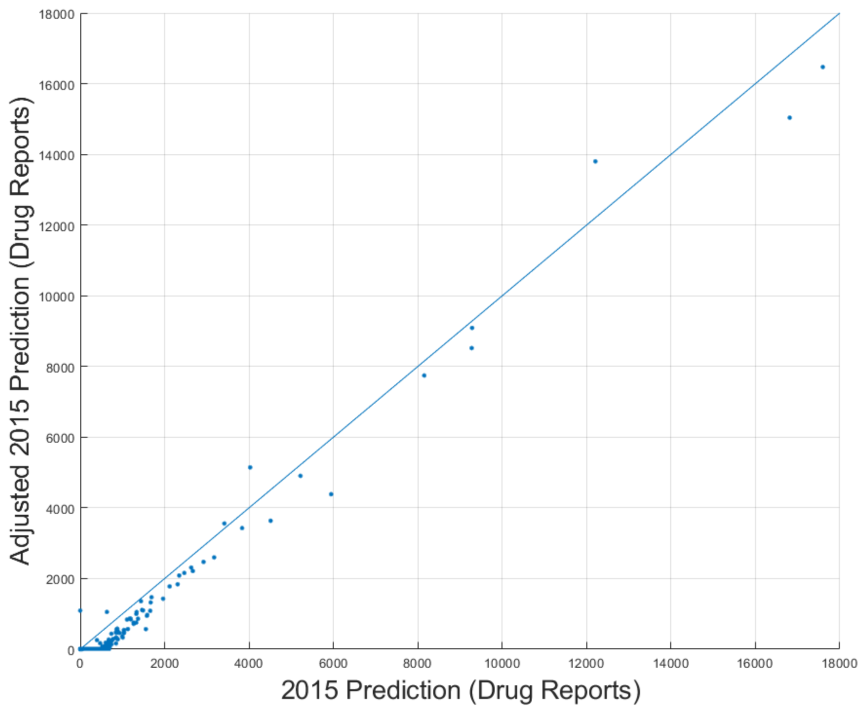


Figure 8: Effect of drug recovery centers if they were implemented in 2015

Figure 8 shows the possible outcome of implementing drug recovery centers in our modeled counties. We estimated the effect on the trend slope to predict the effects of our policy change. With this adjustment, we can see that the overall change in drug reports is predicted to decrease, implying an overall decrease in addiction and abuse.

7 Sensitivity Analysis

7.1 Variation of the Scale of Drug Reports

We tested our model's sensitivity to high m_i values, which simulates a quickly increasing epidemic. We doubled the 2015 drug report values, keeping the 2014 drug report values the same and recalculating the linear fits and m_i values. Our model had an excellent linear regression slope of 1.01. The residual was only slightly higher at 60186 drug reports. This low sensitivity is likely due to the quadratic weighted best-fit line, which puts a higher weight on changes closer to the present.

We then doubled both the 2014 and 2015 drug reports and recalculated m_i. This models high drug report levels but increasing at approximately the same rate as the original data. The residual was higher at 120190. The prediction trends lower than the expected 2015 data with a slope of 0.890. This shows that our model generally robust to changes in the scale of drug reports.

7.2 Variation of Socio-economic Factors

We doubled the value census data for residents who have lived in the same place more than a year (the most representative socio-economic factor). The new 2015 drug report prediction had a linear regression slope of 0.971 but had a much larger residual of 106250 drug reports. Our model is highly sensitive to dramatic changes in socio-economic factors.

8 Conclusion

The use of opioids continues to spread across the country as a lucrative and alluring product for both the poor and the wealthy. Despite the best efforts of the government, this spread is challenging to contain and restrict. As many opioids are misused prescription medications, they are available everywhere. Once in the illegal possession of a user or dealer, however, they are far more restricted in mobility.

Instead of being able to travel internationally to any pharmacy funded by multi-billion dollar conglomerates, the movement of opioids is limited to a sphere of influence. Like culture, drugs are more likely to spread nearby than far away. Once a drug has been spread, it is hard to stop, as its users become addicted and are willing to pay more to get their fix. This has been a problem for decades, and it will only grow worse as time progresses. In order to reduce the mortality and addiction rates, the government must have a greater impact in helping the people that have become addicted, whilst dissuading others from starting.

As states and counties try to regulate and rehabilitate opioid users, they will need to keep adjusting their tactics to the people. Socio-economic factors such as veteran status and length of residency can help predict the trend drug incidents. By implementing programs that are both proactive and reformative, there is a way by which citizens can recover from addiction. Placing facilities and education in the hands of the users helps them to choose self improvement over incarceration. New drugs are more harmful and powerful then ever and it is up to the people and the government to change the nature of drug abuse.

8.1 Model Strengths

- Our model **quantifies the nature of drug-use influence** through gravitational geography, and illustrates how it affects a county both internally and externally.

- By using a weighted trend analysis of total drug reports, we **account for historical data** while placing a **heavier influence on recent developments**.
- A multivariate approach to the model characterizes the influence of related socio-economic factors on the spread of drug use.
- The model can calculate the trend in drug use for every drug and drug type for any county given initial data, and is easily **verifiable against a baseline**.
- The model is **easily expandable** to other states and counties. Having drug report data for the neighboring states of our five states would increase model accuracy.

8.2 Model Weaknesses and Limiting Assumptions

- Since the model is based on only the NFLIS data for 5 states in the Northeastern United States, some county influence values are not being properly calculated, specifically those which are close to borders with other states. As such, the model will underestimate the effect that outside influence will have on these counties, and the model will likely underestimate the trend in drug use. This can be improved by expanding the intial data to contain more counties and states.
- The U.S. Census data that was provided only had a subset of possible socio-economic factors. There are other factors not included in this data that are often more indicative to the spread of drug, such as average household income or homeless population. As such, our model was not able to account for changes in these factors.
- The model needs a substantial amount of data to properly predict the trend in drug use. As such, trends in drug use in counties with less or incomplete data were not able to be effectively represented.
- Since the data provided was on a yearly basis, the model was limited by discrete time steps and lacks a more continuous evaluation for improved accuracy.
- Waves of new drugs are hard to predict with the dataset provided, so the model does not account for the introduction of new types of drugs into these counties. This will cause our model to underestimate level of upwards trend in drug use.

9 Memo to the DEA/NFLIS Chief Administrator

Date: January 29, 2019
To: Chief Administrator, DEA/NFLIS Database
From: MCM Team #1900577
Subject: Opioid Crisis Insights and Recommendations

Chief Administrator,

In an effort to address the opioid crisis, we have modeled how it has spread over time. From the heroin crisis in 2010 in Philadelphia county to the recent outbreak of the 3-Methylfentanyl in Erie county in 2016, drugs continue to be a nationwide issue every year. One of the largest offenders in the origin of new drugs is Ohio. Due to this, Ohio also has one of the highest mortality rates in the United States. However, as we continue to model ways to find where drugs start, it becomes easier to react to them.

It is a grave concern if a county's influence factor rises above the 90th percentile (as measured in 2017). Simulating the spread through 2025, we identified 6 counties that will pass abover the threshold. Above this threshold, it will be very difficult to reduce the drug reports as it will influence the counties around it and stay above the threshold.

There is still hope, though, as we identified several socio-economic factors that correlate with the drug reports. People who have lived in a county for greater than 1 year, the populations of veterans, and people with associates degrees are all groups whose trends are heavily associated with those of drug reports. Creating programs to help residents be positively involved in the community, providing veterans with job opportunities while taking care of their mental health, and encouraging the county's residents pursue higher education are all ways to help fight the opioid epidemic.

With new technology comes an ease of communication, and dealers can be further disassociated with buyers, increasing sales. Synthetic opioids are highly alluring and are 400 to 6000 times stronger than their natural counterparts like morphine. With increase in strength comes an increase in mortality. Another problem with opioids, is that many of the people who are misusing them are acquiring them from their friends or relatives. This means that regardless of geography, drugs will always be accessible to those who choose to take them. It is a different form of war on drugs.

Currently, drug sentencing varies greatly. Some convictions are are low as a $500 fine, while others are as much as 15 years in prison with a $25000 dollar fine for the first offense. Although very harsh punishments do cut back on drug use, they also cost the state a significant amount in incarceration costs for the convicted. Other states have begun implementing a policy by which the county offers a course to first, and sometimes second offenders, which allows them to try to recover from their addiction on their own or as a part of their sentence. Education and recovery saves the state money for offenders, and offers them an opportunity to change their habits. Programs like this have already been shown to be effective in other counties across the United States and the rest of the world.

Insights:

- Counties where drugs begin are the often the same as where they are the most harmful.
- Counties will always have drugs available to them and so it is an not expectation that one would be able to eradicate them all.

Results:

- The count of veterans, long term residents, and junior college graduates have the highest association to the trend in drug incidents per year
- Pennsylvania has the highest mortality rate from drugs
- Erie county has been a recent hot spot for new drugs, such as 3-Methylfentanyl

Recommendations:

- Begin programs to help recovering addicts, similar to the Chinese Clean Needle exchange.
- Make drug effects and use education available to the population at as little cost as possible.
- Provide drug users opportunities to rehabilitate over state funded incarceration.

Best,

MCM Team #1900577

References

[1] GADM. https://gadm.org/.

[2] Data Resources | Drug Overdose | CDC Injury Center, November 2018. https://www.cdc.gov/drugoverdose/resources/data.html.

[3] Drug Overdose Deaths | Drug Overdose | CDC Injury Center, December 2018. https://www.cdc.gov/drugoverdose/data/statedeaths.html.

[4] Understanding the Epidemic | Drug Overdose | CDC Injury Center, December 2018. https://www.cdc.gov/drugoverdose/epidemic/index.html.

[5] National Institute on Drug Abuse. What is the scope of prescription drug misuse? https://www.drugabuse.gov/publications/research-reports/misuse-prescription-drugs/what-scope-prescription-drug-misuse.

[6] National Institute on Drug Abuse. The Science of Drug Use: Discussion Points, February 2017. https://www.drugabuse.gov/related-topics/criminal-justice/science-drug-use-discussion-points.

[7] Afsaneh Bakhtiari and Alexander Rutherford. Mathematical Modelling to Evaluate Measures and Control the Spread of Illicit Drug Use. In Evangelos Kranakis, editor, *Advances in Network Analysis and its Applications*, volume 18, pages 257–268. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012. http://link.springer.com/10.1007/978-3-642-30904-5_11.

[8] Yanguang Chen. The distance-decay function of geographical gravity model: Power law or exponential law? *Chaos, Solitons & Fractals*, 77:174–189, August 2015. https://linkinghub.elsevier.com/retrieve/pii/S0960077915001587.

[9] US Census Bureau Geography. 2017 U.S. Gazetteer Files Record Layouts. https://www.census.gov/geo/maps-data/data/gazetteer2017.html.

[10] William M. Holmes. *Using Propensity Scores in Quasi-Experimental Designs*. SAGE Publications, Pg. 231, June 2013. Google-Books-ID: 7Z92AwAAQBAJ.

[11] Jianhua Li, Toan H Ha, Cunmin Zhang, and Hongjie Liu. The Chinese government's response to drug use and HIV/AIDS: A review of policies and programs. *Harm Reduction Journal*, 7:4, March 2010. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2847555/.

[12] Steady Mushayabasa and Gift Tapedzesa. Modeling Illicit Drug Use Dynamics and Its Optimal Control Analysis. *Computational and Mathematical Methods in Medicine*, 2015, 2015. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4707024/.

[13] Institute of Medicine (US) Committee on Opportunities in Drug Abuse Research. *Drug Abuse Research in Historical Perspective*. National Academies Press (US), 1996. https://www.ncbi.nlm.nih.gov/books/NBK232965/.

[14] Marisa A. Scala. Going Back to School: Participation Motives and Experiences of Older Adults in an Undergraduate Classroom. *Educational Gerontology*, 22(8):747–773, January 1996. https://doi.org/10.1080/0360127960220804.

[15] Patrick M. Tarwater and Clyde F. Martin. Effects of population density on the spread of disease. *Complexity*, 6(6):29–36, July 2001. http://doi.wiley.com/10.1002/cplx.10003.

Appendix

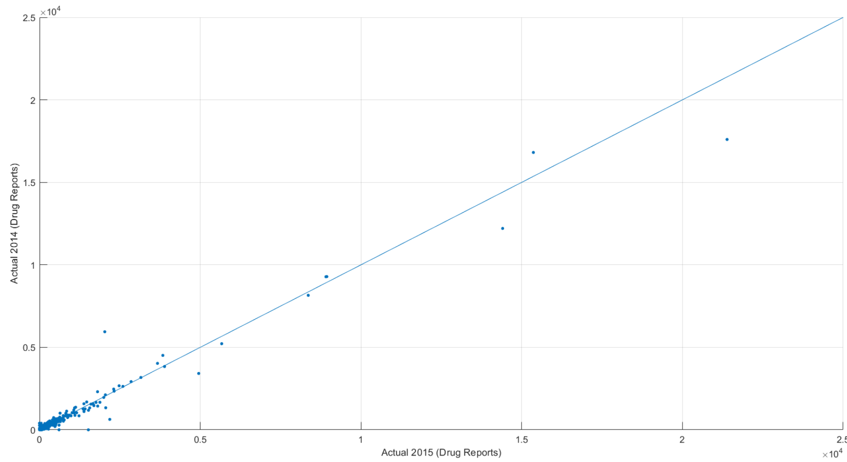


Figure 9: Baseline Test using 2014 data as the predictor for 2015

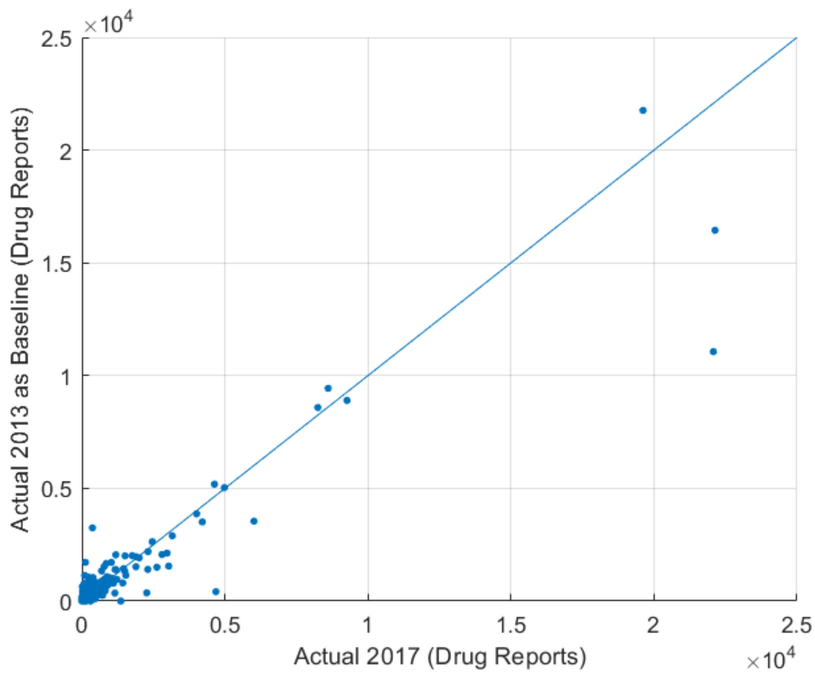


Figure 10: Baseline of 2013 Drug Reports after 4 Time Steps