Team Control Number

1900170

Problem Chosen

C

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

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

2019 Mathematical Contest in Modeling (MCM) Summary Sheet
(Attach a copy of this page to each copy of your solution paper.)

An MCM Paper Made by Team 1900170

Summary

The main purpose of the question is to describe the spread and characteristics of the opioid crisis, after analyzing the data and modeling, we can get a curve of the quantity changes in particular locations, which is defined as spread characteristic curve, and in this way we can conclude the management of opioids.

First, we preprocess the data provided in part I including default value processing, elimination the much too small data and data synthesis and classication which is of great significance for the later calculation.

Second, in order to explore the spread and characteristics between the ve states and their counties over time, we attempt to build a spacial spread models for network(SSN) to simulate the Opioid spread between states and counties, then we can estimate the set of parameters in the SSN through simulation. With that, the obtained parameters can be used to plot the quantity changes curve of opioids in states or counties, and the solved model can be also used to analyze the spread characteristics of opioids, and to predict the birthplace of a certain drug in states and also in counties. Based on the Threshold theorem and the model above, we can work out a threshold to examine whether opioids will be out of control.

Third, after preprocessing the another seven les, we choose ten important possible variables to explore the relationship between use or trends-in-use and socio-economic factors of the U.S, after the correlation test, we determine seven basic variables. By regression analysis, we establish a multivariate linear regression model to explain the influence, then we use the selected socio-economic possible factors which associate with opioids crisis to optimize our original model.

Finally, we propose our efficient strategy for countering the opioid crisis by using the combination of results above, at the same time we test the effectiveness of this strategy by comparing the images before and after changing the parameters, and then we give a parameters success bound.

Key Words: Diffusion Model, Network Structure, Multiple Regression, Simulation Calculation, Threshold Theorem

Contents

MEMO 1

1 Introduction 2
- 1.1 Problem Restatement . . . 2
- 1.2 Our Goals . . . 2
- 1.3 Our Thinking . . . 2

2 Assumptions and Notations 3
- 2.1 Assumptions . . . 3
- 2.2 Notations . . . 3

3 Data Preprocessing 4
- 3.1 NFLIS data Precessing . . . 4
- 3.2 U.S. Census socio-economic data Processing . . . 4

4 Model Construction 5
- 4.1 spacial spread models for network(SSN) . . . 5
- 4.2 Threshold Theorem Analysis . . . 7
- 4.3 Opioid Drus Spread Characteristics between States and Counties . . . 8
  - 4.3.1 Simulated Estimation of State Model Parameters . . . 8
  - 4.3.2 Opioid Drug Spread Characteristics between States . . . 8
  - 4.3.3 Opioid Drug Spread Characteristics between Counties . . . 10
- 4.4 Regression Analysis Model . . . 12

5 Strategy for Countering the Opioid Crisis 14

6 Interference Analysis 17

7 Conclusions 17

8 Strengths and Weaknesses 18
- 8.1 Strengths . . . 18
- 8.2 Weaknesses . . . 18

References 19

Appendices 19

Appendices A Potential Origins of Opioids 19

MEMO

To: the Chief Administrator, DEA/NFLIS Database

From: Team # 1900170

Date: January 28, 2019

Subject: Solutions to the Opioid Crisis

Honorable Chief Administrator of DEA/NFLIS Database,

Currently the United States is experiencing a national crisis and seek for the possible strategy to counter it. Our team has made a comprehensive study on opioid drugs abuse based on historical data from 2010-2017 in five states(KY, OH, PA, VA and WV), including the spread and characteristics of opioids in and between the five states and their counties over time, possible locations where specific opioid use might have started, any important socio-economic factors associated with opioids abuse and our strategy for countering the opioid crisis.

Spread Characteristics and Origin of Specific Drugs: There are great regional differences in the spread characteristics of various opioid drugs between States and counties. There are counties in each state where many medicines originate, that is to say, the origin of medicines is centralized in distribution as a whole. At the same time, for a specific drug, combining the propagation characteristic curve and threshold theorem, we can predict whether the county will reach the quantity level that the government needs to worry about, and give the possible time meanwhile.

Socio-Economic Factors Associated with Opioids Abuse: After analysis the data of KV state over the time, we found that Number of newbor babies(Nnb), Higher education Rate(Her), New immigrants rate(Nir), Local birth rate(Lbr), Local population conversion rate(Lcr), Foreign-born population(Fbp), Households with teenagers(Hwt) have associated with opioids abuse. Thus, The management of opioids can start from these aspects. And the data trends of socio-economic factors in each state were similar to KY state.

Our strategy: To counter the opioid crisis, we offer you the following suggestions:

- Raise the public awareness of opioids
- Control the major counties where opioids originate to prevent it from affecting more counties.
- Develop education to improve people's education level and reduce opioid abuse.

These are our analysis of the number of opioid abuse in five regions, and I sincerely hope that it will help in the management of opioids. Thanks!

1 Introduction

1.1 Problem Restatement

Starting in the 1990s with the increase of prescription painkillers, the Opioid epidemic has become a modern health crisis.And the United States is experiencing a national crisis regarding the use of synthetic and non-synthetic opioids.Federal organizations such as the Centers for Disease Control (CDC) are struggling to save lives and prevent negative health effects of this epidemic, such as opioid use disorder, hepatitis, and HIV infections, and neonatal abstinence syndrome.

For this problem, we focus on the five U.S. states: Ohio, Kentucky, West Virginia, Virginia, and Pennsylvania and their counties.

1.2 Our Goals

Based on our understanding of the problem, we set the following goals:

- Using the NFLIS data provided, build a mathematical model to describe the spread and characteristics of the reported synthetic opioid and heroin incidents (cases) in and between the five states and their counties over time.
- Using the model above, identify any possible locations where specific opioid use might have started in each of the five states.
- Answer the question whether there are any specific concerns the U.S government should have if the patterns and characteristics identified continue. And find out that at what drug identification threshold levels these will occur and predict where and when they will occur in the future.
- Modify the model above, Try to include any important factors promoting the opioid use or trends-in-use from the U.S. Census socio-economic data provided.
- Using a combination of results above, identify a possible strategy for countering the opioid crisis. Then use the model to test the effectiveness of it and identify any significant parameter bounds that success (or failure) is dependent upon.

1.3 Our Thinking

This is a typical big data problem, so we solve it from the point of view if statistical analysis. Here is our thinking.

First, we preprocess the data provided, which includes default value processing, elimination processing of very small data and data synthesis and classification.

Second, to explore the spread characteristics between the five states and their counties over time, we consider to build a spacial spread models for network(SSN) to spacial spread between nodes. And we can estimate a set of parameters through simulation. The parameters obtained can be used to plot the growth curve of opioids in states or counties, and to analyze the spread characteristics of opioids and its birthplace in states and counties. And Based on the Threshold theorem and the model above, we can work out a threshold to determine whether opioids will be out of control.

Third, after preprocessing the additional seven files, we choose 10 important variables to explore the relationship between use or trends-in-use and socio-economic factors of the U.S..we call them the basic variables. By establishing the multivariate regression model, we can find the socio-economic possible factors associated with opioids abuse. And then we modify our model above to include important factors from the data set.

Finally, we propose our strategy for countering the opioid crisis using a combination of your Part 1 and Part 2 results above. And then use our model to test the effectiveness of this strategy.

2 Assumptions and Notations

2.1 Assumptions

Due to lack of necessary data and limitation of our knowledge, we make the following assumptions to help us perform modeling. These assumptions are the premise for our subsequent analysis. In fact, the more data we collect, the more accurate result that the model will catch.

- While considering the opioids flow, the impact of topography is ignored. Therefore, we only need to consider the drug flow in two-dimensional space.
- Simplify the model by assuming that a person will not stop taking opioids without outside intervention after starting to take opioids.
- Assumed that the difference in the probability of opioid infection among susceptible individuals of different ages is not taken into account.
- Assumed that the influence of the duration of opioid consumption on the spread of opioids is not taken into account.

2.2 Notations

The primary notations used in this paper are listed in Table 1. Other notations will be explained later.

Table 1: Notations

Symbol	Definition
$\alpha_{i(t)}$	Increase rate of opioid cases for county itself
β	The transformation factor of adjacent counties
d_{ij}	The distance between location i and j
D_{ij}	The diffusion rate of area i and j
$D(x)$	Opioid Diffusion Rate in One-Dimensional Direction
$N(x,t)$	The number of opioids of X positions at t time
$f(N,x)$	The average growth rate of opioids at x position

3 Data Preprocessing

For data-analysis problem, there are usually some incomplete, abnormal data and neglectable data in the large amount of raw data, which may seriously affect the efficiency of modeling and the accuracy of conclusions. So it is quite important to preprocess the data.

3.1 NFLIS data Precessing

- Replace data in columns "TotalDrugReportsCounty" and "TotalDrugReportsState" with columns "TotalOpioiddDrugReportsCounty" and "TotalOpioidDrugReports State". We are required to describe the spread and characteristics of the reported **synthetic opioid and heroin** incidents (cases), but the data provided in these two columns are the sum of all drugs which we don't need.
- Eliminate data on some negligible drugs. Noting that there are some substances with very few occurrences in the data set,we eliminate them, for the reason of inadequate and worthless roles they play for our modeling.
- introduce the longitude and latitude information of each county and state and present it intuitively in the form.

3.2 U.S. Census socio-economic data Processing

- Exclude indicators with missing data for some years.
- Data Synthesis and Classification. There are approximately a hundred variables in the dataset provided. Not all of them are used in our model. Thus, we sort the data we need into a new dataset. And we find that there is a hierarchical relationship between the values of variables, which makes it very convenient for us to select the variables that may be relevant.
- Selection of indicators statistics. The ratio of the cumulative value of indicators

to the number of counties in a state:

$$\hat{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}$$

4 Model Construction

4.1 spacial spread models for network(SSN)

To explore the spread and characteristics of the reported synthetic opioid and heroin incidents (cases) in and between the five states and their counties over time, we intend to build a network model to simulate the spread of opioids.

The pre-processed MCM_NFLIS_Data.xlsx table can provide a reference value for the number of opioid cases in each region every year.

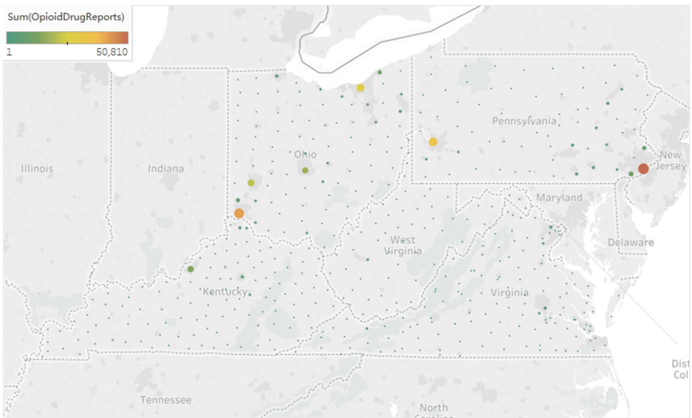


Figure 1: Network Diagram of Node and Number of Opioid Cases

Figure 2: Interstate Network Flow Map

For each basic node, the number of opioid cases corresponding to each node on the map can be regarded as a scalar field in the discrete vector space. For simplification, we only consider the two-dimensional space. Therefore, we can use the two-dimensional discrete flow-diffusion equation[1] to simulate the interaction between the top nodes, that is, the opioid diffusion model can take the following form:

$$IPOC = DIC + SGFC \tag{2}$$

Where:

IPOC is Increasing proportion of opioids in Nodes, DIC is Dynamic impacts between Nodes and SGFC is Self-growth factors in Nodes.

First, in order to obtain the specific propagation equation, we assume that the spread of opioids is caused by a large number of micro-human factors (like population or circle-driven effect). Then we can know that these small steps form a description of the overall diffusion movement. For a spatial dimension, if propagating along a straight line, the opioid diffusion model can be transformed into the following equation:

$$\frac{\partial N}{\partial t} = \frac{\partial}{\partial x}(D(x)\frac{\partial N}{\partial x}) + Nf(N,x) \tag{3}$$

Where:

- $N(x,t)$ is the number of opioids of X positions at t time.
- $D(x)$ is the diffusion rate.
- the function $f(N,x)$ is the average growth rate of opioids at x position.

If each node is considered as a regional location on a two-dimensional plane, then x is a block, and the partial differential is replaced by a number of gradients. In many cases, the radial symmetric solution is usually used to describe the propagation. Therefore, the problem can be simplified as having only one spatial variable r(r represents the distance from the starting point). If the diffusion coefficient is constant D, then the equation is:

$$\frac{\partial N}{\partial t} = \frac{D}{d_{ij}}\frac{\partial}{\partial d_{ij}}(d_{ij}\frac{\partial N}{\partial x}) + Nf(N,d_{ij}) \tag{4}$$

However, for this problem, opioid spread between regions should be regarded as a discrete space or a discrete space-time model, thus the simulation form of discrete space continuous time is as follows:

$$\frac{\partial N_i}{\partial t} = \sum_{j=1}^{n} D_{ij}(N_j - N_i) + N_i f_i(N_i) \tag{5}$$

Where:

- $N_i(t)$ is the number of opioids in area i at t time.
- D_{ij} is the diffusion rate of area i and j.
- function $f_i(N_i)$ is the growth rate of opioid cases in area i.

Based on this, we use a two-dimensional diffusion equation to describe the interaction between counties. We divide five states into two-dimensional networks. For each county, we label (i, j) points on the two-dimensional plane as a basic node, and he number of reports of specific opioids in its area as a scalar field in discrete space. It is easy to know that the spread between counties (regional points) should be inversely proportional to the distance between counties, and assume that the growth rate of each region itself remains unchanged.

$$\frac{\mathrm{d}N_{i(t)}}{\mathrm{d}t} = \alpha_{i(t)}N_{i(t)} + \beta_i \sum_{i \neq j} \left(N_{i(t)} - N_{j(t)}\right) \frac{1}{d_{ij}} \tag{6}$$

Where:

- We define $(\frac{N_{i,j}-N_{i-1,j}}{d_{i-1,j}}) + (\frac{N_{i,j}-N_{i+1,j}}{d_{i+1,j}}) + (\frac{N_{i,j}-N_{i,j-1}}{d_{i,j-1}}) + (\frac{N_{i,j}-N_{i,j+1}}{d_{i,j+1}})$ as $\beta_i \sum\limits_{i \neq j} \left(N_{i(t)} - N_{j(t)}\right) \frac{1}{d_{ij}}$.
- α parameter is the increase rate of opioid cases for county itself(It is a parameter that changes with time).
- β parameter is the transformation factor of adjacent counties

4.2 Threshold Theorem Analysis

Based on the Threshold theorem[2] and **the Equation 6** above, we have the following derivation:

Let:

$$\frac{\mathrm{d}N_{i(t)}}{\mathrm{d}t} = \alpha_{i(t))}N_{i(t)} + \beta_i \sum_{i \neq j} \left(N_{i(t)} - N_{j(t)}\right) \frac{1}{d_{ij}} \tag{7}$$

Then we can get:

$$\alpha_i + \beta_i \sum_{i \neq j} \frac{1}{d_{ij}} - \beta_i \sum_{i \neq j} \frac{N_j/N_i}{d_{ij}} > 0 \tag{8}$$

Thus the drug identification threshold levels N_i is:

$$N_i > \frac{\xi}{\delta + \gamma} \tag{9}$$

Where:

- We define ξ as $\sum\limits_{i \neq j} N_j/d_{ij}$.
- We define δ as $\frac{\alpha_i}{\beta_i}$.

- We define γ as $\sum_{i \neq j} \frac{1}{d_{ij}}$.

When the number of drug cases exceeds this threshold, $\frac{\mathrm{d}N_{i(t)}}{\mathrm{d}t} > 0$, the number of drug cases will gradually increase, and the spread will continue to spread outward.

4.3 Opioid Drus Spread Characteristics between States and Counties

4.3.1 Simulated Estimation of State Model Parameters

After data processing, we can estimate a set of roughly corresponding coefficients $\alpha_{i(t)}$ and β by simulation.

Table 2: Parameter tables of $\alpha_{i(t)}$ and β

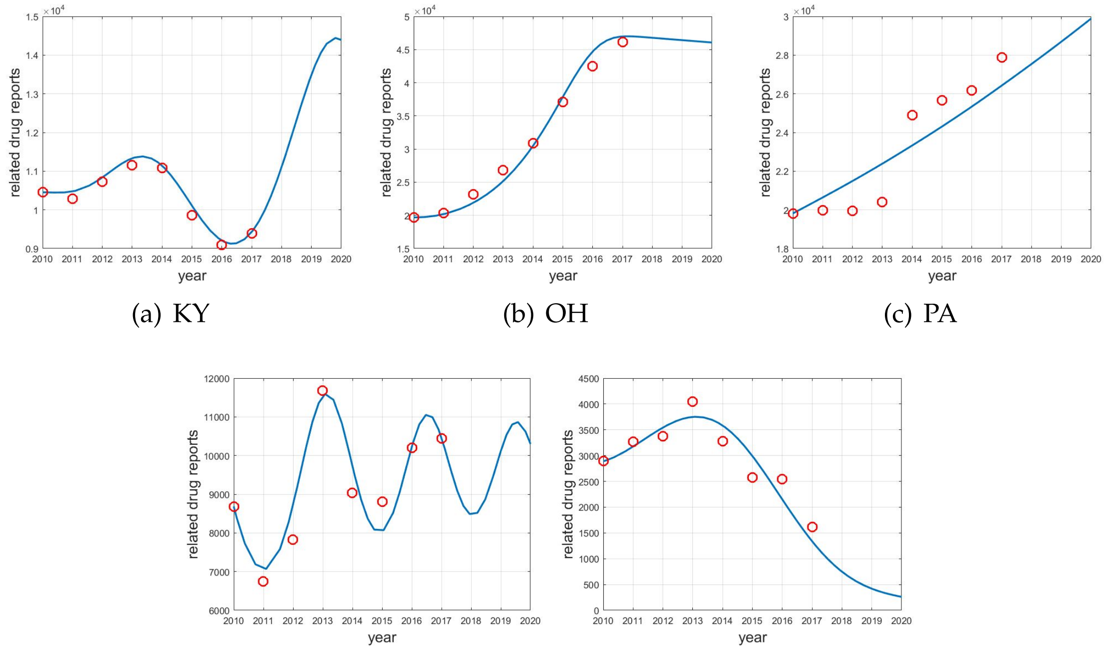
	$\alpha_{i(t)}$	β
KY	$0.032t \cdot sin(1.28t^{0.96} - 0.9)$	1.06
OH	$0.058e^{-0.00001t^{6.5}}$	1.13
PA	$0.05e^{-0.1t^{0.06}}$	1.09
VA	$0.41cos(t^{1.25} - 2.6) \cdot e^{-0.05}$	1.08
WV	$0.19t \cdot (-0.35e^{0.1t} + e^{-0.3t})$	1.11

4.3.2 Opioid Drug Spread Characteristics between States

According to a set of roughly corresponding coefficients $\alpha_{i(t)}$ and β above, we adopt the descriptive statistics method to show the State related Drug Reports of each state.

By analysis the Trend Map of State related Drug Reports, we work out the the characteristics spread of the reported synthetic opioid and heroin incidents (cases) in and between the five states:

- For KY state.The number of overall opioid cases in KY has increased and then decreased in the past eight years. In 2010, the growth rate has gradually increased, and reached a maximum in 2013, but then began to decline year by year. It can be seen that the spread of drugs in the state has obviously been effectively controlled, showing a state of drastic reduction in recent years, but there is a rebound trend in 2017. It is predicted that if the government does not take any effective measures at this time, the states opioid cases will gradually rising, the state's opioid flooding will lose control when $N_{i(t)}$ is reached.
- For OH state. The opioids in the OH state have been rising for the past eight years, but they have a downward trend after reaching the maximum in 2017. If the government does not take effective measures at this stage, the growth trend may continue.



(a) KY (b) OH (c) PA

(d) VA (e) WV

Figure 3: Trend Map of State related Drug

- For PA state. The number of opioid cases in PA has shown an overall upward trend in the past eight years, and its growth rate is highly variable, but it has not produced a significant reduction in the number of drug cases. It may be that the local government has not regulated it. It may also be that the suppression measures did not produce significant effects. It is predicted that the number of cases in the state will continue to grow, that is, the number of drugs in the state will always be higher than the threshold. Therefore, the state government obviously needs the concern whether the opioid spread will continue to increase, and effective measures should be taken to reduce its internal internal rate of propagation, which is to reduce its own growth rate.

- For VA state. The number of opioid cases in VA has shown an oscillating change, at the same time, we can see the volatility of the curve is decreasing, so we can assume that the government has started to take appropriate measures to curb this phenomenon. However, due to the changing nature of the county itself, we can still predict that the county will reach a small peak again around 2023. The government needs to have such concerns and take effective measures to prevent the county from returning to above the threshold by 2021.

- For WV state. The number of WV state cases has increased and then decreased in the past eight years. In 2010, the growth rate has gradually increased, and reached a maximum in the year of 2013, but then began to decline year by year. It can be seen that the management of the government has taken effect. If it can continue to persist the policy, opioids can be depressed within an acceptable range.

4.3.3 Opioid Drug Spread Characteristics between Counties

For each type of opioid, we use its reports in various counties from 2010 to 2017 as the independent variables. Through the model, we can get the possible birthplace of each drug. Because some drugs have only appeared several times, we then don't discusse these drugs for the particularity of their data.

In particular, Heroin's data with particularly high heat of spread has been very large compared to other drugs over the years. Its spread characteristics are very obvious and have good research significance. We have obtained the the diffusion characteristics of heroin through the model. Then we pick some typical examples to analyze them:

For KY state. Through the model, we can find that it is very likely that it is one of the origins of heroin abuse in KY state. First, the number of its own drugs is much larger than that of the adjacent counties in the past years. For example, the difference on number of drugs between BULLITT and OLDHAM counties which have obvious influence relationship is less than 100 in 2010 to 2017. Their quantity changes show a short-term reactivity. At a low level, it is obvious in 2-3 years. At that time, JEFFERSON County was in the state of growth and outward diffusion. Combining with the low level of initial drug crisis in BULLITT and OLDHAM counties, the stability of the two counties was strong, indicating that the spread of heroin would also be significantly hindered. JEFFERSON County itself started to grow at a low level. After rapid growth, the growth rate began to decline in 2015. It is predicted that the number of heroin cases in this county will begin to decrease year by year under some existing restraining factors, the dissemination will weaken year by year, and the government can also decrease their worries to the county.

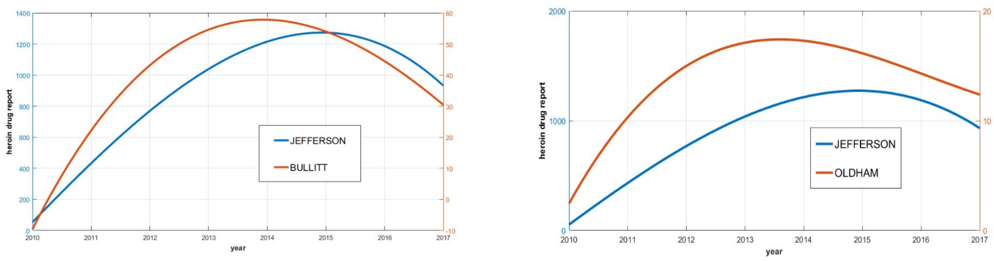


Figure 4: Possible birthplace in the KY state

For OH state. In OH state, we draw the conclusion that one of the most likely heroin producing counties is Hamilton. The change of Hamilton County compared to its neighboring counties Butler and Warren is more stable and gentle, indicating that it is less disturbed by the outside world. Its growth is obviously affected by its own factors, as a birthplace, it accords with the stable spread of heroin. Compared with Hamilton, the amount of heroin is also quite different, and the characteristics of external interference factors are obvious. This shows that the spread of heroin in the surrounding counties also shows a rapid rise and disappearance, which might thank to the government's effective and timely interference. At the same time, as the number of cases in the county has shown a downward trend in 2017, which below the threshold, it may be that the government has taken appropriate measures. We can predict that the drug spread in the county has been controlled, the number of drugs will begin to decrease gradually, without causing greater concern to the government for the time being.

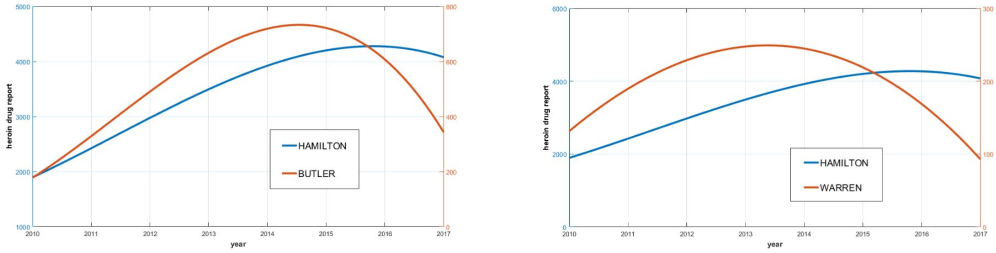


Figure 5: Possible birthplace in the OH state

For PA state. We get the spread characteristic curve of each county through the model, and predict its origin in ALLEGHENY. Compared with Washington County, Allegheny County has similar characteristics of change in the past few years. There is an instant type in its spread characteristics, which indicates that its spread is rapid and the time interval is short. At the same time, we can see the peak value of the line of Allegheny, which appeared earlier than Washington County, and its spread also shows a time lag. It can predict that the Opium crisis in Washington is transmitted from Allegheny County.

For another Westmoreland County adjacent to Allegheny, the differential equation of Westmoreland County is inverted U-shaped, which leads to the increase of the number curve first, then decrease, and the deceleration gradually increases. So we can predict that the number of Westmoreland County will continue to decline and stabilize at a low level. This may be due to the effective control that the government has begun to carry out. Affected by Allegheny, Westmoreland's spread shows a pattern of fast grow ande fast decrease.The number of Al counties uctuates from threshold level to threshold level. While the number of Allegheny counts changes back and forth around the threshold. And the range of change tends to increase, which leads to the oscillation of its quantitative characteristics, but it can be predicted that the number of drugs in this area will reach a new peak around 2021. Therefore, the above analysis shows that the government needs to take more effective measures to restrain this growth and reduce the possibility of returning to the threshold level.

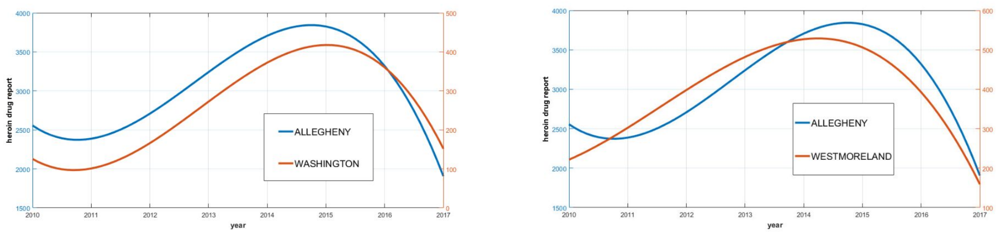


Figure 6: Possible birthplace in the PA state

For VA state. Henrico County is similar to Chesterfield County, and the fluctuation of Henrico County is obviously stronger than Chesterfield County. At this time, the mutual spread between the two counties reflects a positive feedback. At the same time, the changes between Henrico County and Hanover County show a mutual mobility at a certain growth rate. Because the initial level of Henrico County is much different from Hanover County. And the initial level of Hanover is very low, of which number started from zero.It can be predicted that Hanover should be affected by Henrico, so Henrico County is more likely to be the birthplace. At the same time, it can be predicted that both Hanover and Henrico may show an increasing trend as a whole under the influence of the increasing mobility of Henrico and Hanover. Under the positive feedback of Henrico and Chesterfield, Chesterfield will also show a similar growth trend as HE, which can be achieved by the model. It is predicted that Henrico and Chesterfield will reach a new peak in about 2025, and Hanover will continue to rise because of its number above the threshold. There may be a short-term peak in 2020.

Because heroin forms an uncontrollable and interactive spread chain in these counties and their surrounding areas, the local government needs to take measures to control the chain-like spread quickly and reduce the interaction factors among counties, so that the spread of heroin can be effectively suppressed.

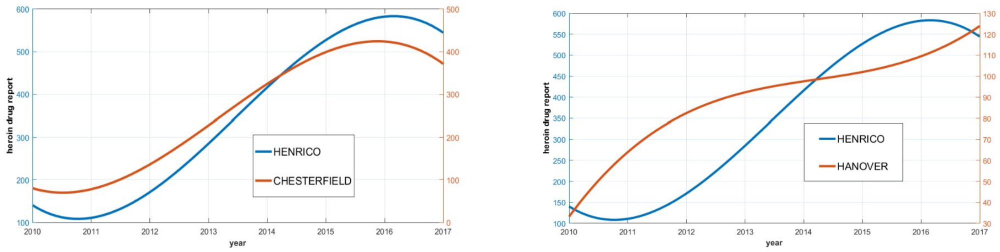


Figure 7: Possible birthplace in the VA state

4.4 Regression Analysis Model

After data preprocessing, we choose 10 important variables to explore the relationship between use or trends-in-use and socio-economic factors of the U.S.. we can call them the basic variables. And meanwhile we select the number of annual state reports on opioids as a dependent variable.

And after data analysis and comparison, we found that the data trends of socio-economic factors in each state were similar in 10-16 years, so here we select KY state as a representative for data analysis.

Then, establish the following multivariate linear regression model:

$$y = b_0 + \sum_{i=1}^{10} b_1 x_1 + \varepsilon, \varepsilon \in N(0, \sigma^2) \tag{10}$$

Calculating the data based on the equation 10, We find that the coefficients before Ahs,Nes and Tpp are 0, which is irrelevant and need to be eliminated.

Next, we calculate the data again with the remaining variables based on the equation 10(But change the upper bound of i to 7).

The results of regression coefficient is showed in table 4. We found that the coefficients of the Nnb, Lbr, Lcr, and Fbp terms are all positive, indicating that the number of

Table 3: 10 Basic variables

Basic variables	Meaning	Unit
Number of newborn babies(Nnb)	Based on the number of women 15 to 50 years old who had a birth in the past 12 months	Number
Higher education Rate(Her)	Derived from the proportion of people who have undergone college or higher education to the total population.	Percent
New immigrants rate(Nir)	Derived from the sum of the proportion of people who did not live in the county or abroad a year ago	Percent
Local birth rate(Lbr)	Derived from the proportion of the number of babies born locally to the total number of local people.	Percent
Local population conversion rate(Lcr)	Ratio of population naturalized U.S. citizen to Foreign-born population	Percent
Average household size(Ahs)	Derived directly from raw data	Number
Foreign-born population(Fbp)	Number of foreign births	Number(thouand)
Non-English native speakers(Nes)	Population whose mother tongue is not English	Percent
Total populatio(Tpp)	Total population	Number
Households with teenagers(Hwt)	Households with one or more people under 18 years	Percent

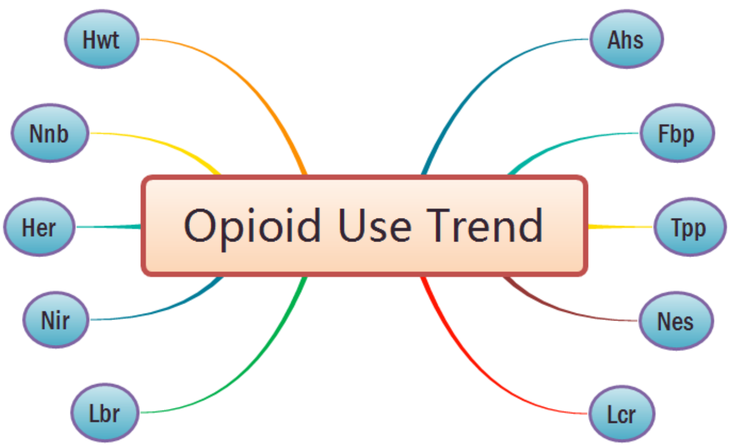


Figure 8: 10 possible related variables

drug reports will increase as the number of new populations increases, the local population conversion rate increases, and the number of people born in foreign countries increases, that is, these factors are positively correlated. The Her, Hir, and Hwt coefficients are negative, indicating that as the level of education received increases, the proportion of migrants increases, and the number of families in the home increases, the use of opioids decreases, that is, these factors are negatively correlated.

Table 4: Regression Coefficient Table

Regression Coefficient	Coefficient Estimation Value
Nnb	136.5895
Her	-9597.4947
Hir	-6549.9746
Lbr	5828.1000
Lcr	247.0401
Fbp	15522.502
Hwt	-10577.0315

Then we add 7 related variables of socio-economic variables above to modify your model from Part 1 to and get the following models:

$$\frac{\mathrm{d}N_{i(t)}}{\mathrm{d}t} = \alpha_{i(t)}N_{i(t)} + \beta_i \sum_{i\neq j} \left(N_{i(t)} - N_{j(t)}\right) \frac{1}{d_{ij}} + \sum_{m=1}^{7} k_m \varepsilon_m N_{i(t)} \tag{11}$$

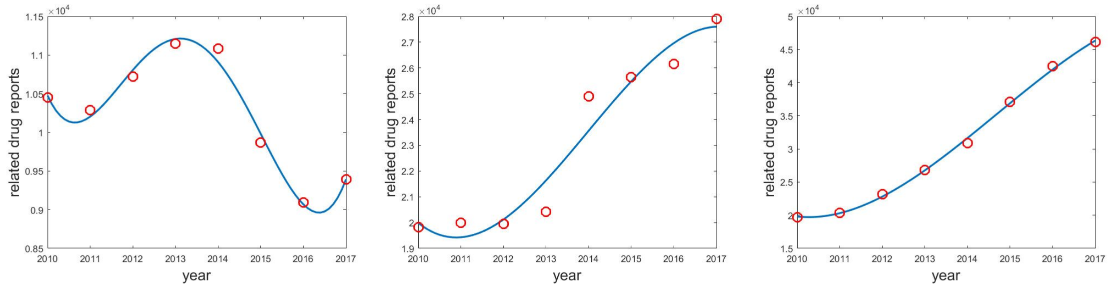
where:

- k_m is he coefficient of the corresponding variable.
- ε_m is the 7 variables selected.

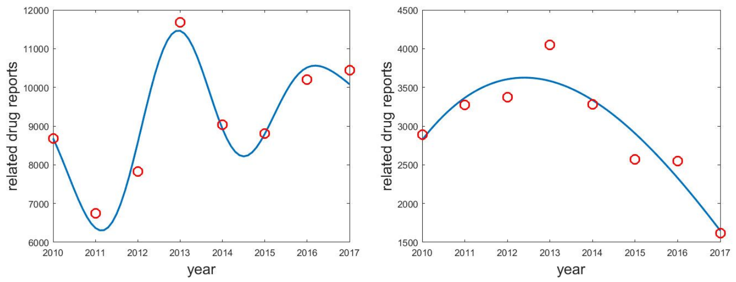
The function image drawn by the improved model is shown in the figure 9. Compared with the original model, it can be found that the improved model is more consistent with the data of 2010-2017, which shows that the model is more accurate and more consistent with the actual situation.

5 Strategy for Countering the Opioid Crisis

In the second part of the regression analysis, we can find that the improvement of the education level of the local population can obviously and effectively form a negative effect on the number of drug reports in the region. Combining with the existing



(a) KV (b) OH (c) PA



(d) VA (e) WV

Figure 9: Corrected Trend Map of State related Drug

investigation of the opioid crisis, we have come up with a very effective countermeasure in theory, that is, through the public propaganda of the crisis, we can raise the public's awareness of the opiate crisis and make it bigger. People's understanding of relevant knowledge can be approximated by the fact that the knowledge level of the population has been improved, thus improving the education level of the local population. Based on this, we expand the corresponding coefficients of the factors of population education level in the second question model to a certain extent, that is, the model becomes:

$$\frac{\mathrm{d}N_{i(t)}}{\mathrm{d}t} = \alpha_{i(t)}N_{i(t)} + \beta_i \sum_{i\neq j}\left(N_{i(t)} - N_{j(t)}\right)\frac{1}{d_{ij}} + \Theta k_1 m_1 N_{i(t)} + \sum_{m=1}^{6} k_m \varepsilon_m N_{i(t)} \tag{12}$$

Take a = 1.05, that is, assuming that the overall education level of citizens can be increased by 5% after implementing our countermeasures. Take the communication between states as an example, the images of WV,OH and VA states are as follows:

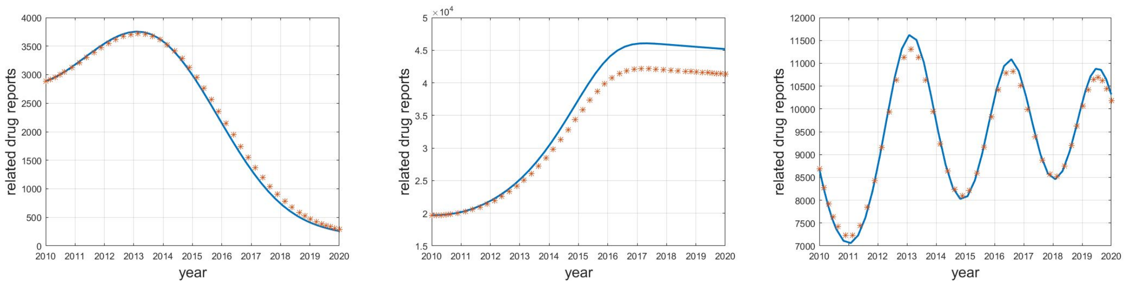


Figure 10: Changed drug reports graphs in WV, OH, VA States

It can be found that after the education level of the local population, only the peak can be reduced in one state, that is to say, the degree of spread can be reduced, but the number of cases has not been significantly reduced in some years.

However, according to the original mode of spread and the characteristics of communication after the implementation of countermeasures, the state is expected to stabilize at a low level in the end, so we can think that we proposed to increase the propaganda so as to make the degree of spread intense. The countermeasure of weakening is effective.

At the same time, for WV, OH states, the number of cases can be significantly reduced, the spread characteristics are weakened to a certain extent, with very good results.

The improvement of the education level of the approximate local population produced by this strategy may also have an unknown impact on the spread factor β between counties. Therefore, we revise the model to read as follows:

$$\frac{\mathrm{d}N_{i(t)}}{\mathrm{d}t} = \alpha_{i(t)}N_{i(t)} + \vartheta\beta_i \sum_{i\neq j}\left(N_{i(t)} - N_{j(t)}\right)\frac{1}{d_{ij}} + \Theta k_1 m_1 N_{i(t)} + \sum_{m=1}^{6} k_m \varepsilon_m N_{i(t)} \tag{13}$$

On the basis of Θ=1.05, we gradually adjust the value of ϑ. We find that when ϑ =1.45 or so, the optimized communication characteristics begin to fail. The overall number of cases has risen as a whole for the number of cases after taking measures. The intensity of spread in part of the time will even be higher than the original curve. Obviously, when the given scheme is implemented, the spread factor between each state has increased by 45% and by 45%. At that time, the plan failed. Because the spread factors between states are small and each state is similar, the permissible range of success is large. For some counties, the influence of spread factors on each other is strong, that is, when the spread factors are large, we can know that a small range of changes in spread factors will result in ineffective schemes. Therefore, in the implementation of the strategy, we must consider the impact of the implementation of the program on mutual communication, but under certain conditions, the program can be regarded as an effective countermeasure.

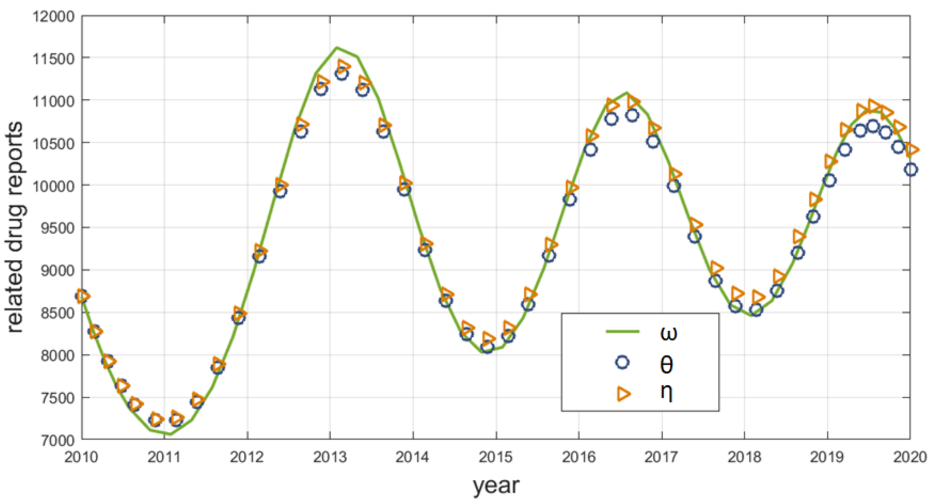


Figure 11: Changed drug reports graphs in WV, OH, VA States

6 Interference Analysis

In order to analyze the uncertainty of SSN model, we add each of the parameters a distraction, which is a random variable that obeys standard normal distribution, and the we can get a inspection equation:

$$\begin{aligned}\frac{dN_{i(t)}}{dt} &= (\alpha_{i(t))} + \xi)N_{i(t)} + (\beta i + \xi i)\sum_{i \neq i}(N_{i(t)} - N_{j(t)})\frac{1}{d_{ij}} \\ &+ \sum_{7}^{m=1}(k_m + \xi_i)\varepsilon_m N_{i(t)}, \xi \sim N(0,1)\end{aligned} \tag{14}$$

Then we can arrived at a figure which shows their characteristic of quantity under the interference.

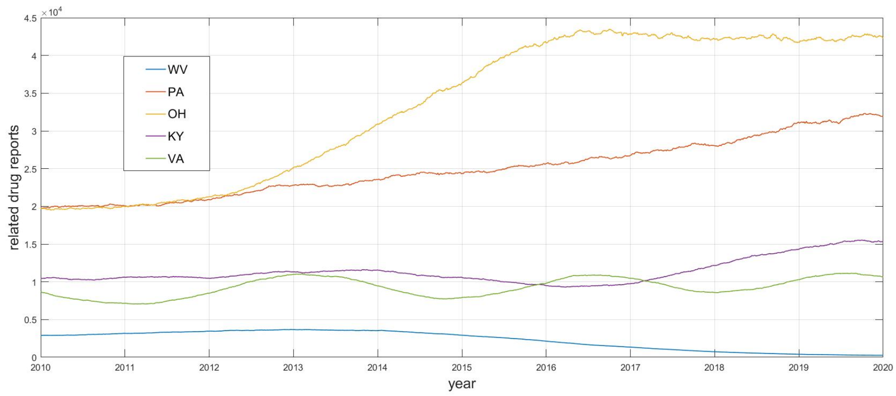


Figure 12: Interference analysis Graph

7 Conclusions

We are asked by the Chief Administrator, DEA/NFLIS Database to identify a possible strategy for countering the opioid crisis. After performing data analysis and modeling, we have finished the task successfully.

First, Using Reaction-Diffusion Model of Bilevel Network, we find the spread and characteristics of the reported synthetic opioid and heroin incidents (cases) in and between the five states and their counties over time is characterized by oscillation or rise. After that, We have known any possible locations where specific opioid use might have started in each of the five states by using the model we build. Then, We raised the concerns the U.S. government should have if the patterns and characteristics continue. We utilize infectious Disease Threshold Theorem to find the drug identification threshold levels successfully.

Second, we construct a Multivariate Linear Regression Analysis Model by choosing 10 possible variables derived directly or indirectly from dataset provided. And finally select 7 related variables associated with the opioid crisis.

Finally, we put forward our strategies for countering the opioid crisis by based o the previous analysis. To test the effectiveness of policies, we test the effectiveness of this strategy by comparing the images before and after changing the parameters. And it is found that the policy proposed by us has a certain effect on reducing the prevalence of opioids.

8 Strengths and Weaknesses

8.1 Strengths

- **Data Preprocessing.**When faced with big data problem, the data processing is really important. Through this step, we greatly improve the quality of the data. Thus, it is more efficient and convenient for us to solve the problem.
- **Good Expansibility.** Not only do we consider its individual changing factors, but also the spatial interaction factors. And after the optimization of the model, it also adds relevant social factors, so it has flexibility in adjusting.
- **Based on Mature Theory.** Based on mature theory, its theoretical derivation is simple and its complexity is low.
- **The Reference of Differential Equations.**So it can approximate the trend of non-linear data well. Further more, from the interference analysis, it can be seen that the model is stable after it is established.
- **Strong Objectivity.** The coefficients of the model are calculated instead of subjective assignment, which reduces the additional errors.

8.2 Weaknesses

- **Simplifying assumption**.Simplified assumptions are adopted for a solvable model, so the result may slightly digress from the ground truth. cause extra error of our models.
- **Default Propagation from High Density to Low Density.** The interaction part of the model is derived from divergence, and the final result may be slightly deviated from the actual situation.
- **Too little yearly data.** The data solved by this model are only obtained from eight years'data processing, and the amount of data is small, which may deviate from the actual situation.
- **Time-varying Parameters.**In practice, the transformation factors and socio-economic impact parameters may be time-varying parameters, which may lead to errors in the model results.

References

[1] Britton N F. Reaction-Diffusion Equations and Their Applications to Biology. *Salt Lake City: Academic Press*, 1986.

[2] Vincenzo Capasso, Gabriella Serio. A generalization of the Kermack-McKendrick deterministic epidemic model. *Mathematical Biosciences*, 1978:43-61.

[3] Murray, J. D., Mathematical biology. *Springer, Berlin*, 1993.

[4] Legrand J, Grais RF, Boelle PY., Mathematical biology. *Springer, Berlin*, 1993.

[5] Marek R., Straub J., Understanding the dynamics of Ebola epidemics. Epidemiology and Infection, 2007;135 (4): 610-21.

[6] Blair H. Smith. *Opioid epidemic the global spread explained*, 2018.

[7] Robin Ghertner and Lincoln Groves, Ph.D. THE OPIOID CRISIS AND ECONOMIC OPPORTUNITY: GEOGRAPHIC AND ECONOMIC TRENDS.

[8] Caminero J.A., Sotgiu G., Zumla A., Migliori G.B., Best drug treatment for multidrug-resistant and extensively drug-resistant tuberculosis. Lancet Infect Diseases, 2010;10: 621-629.

[9] Prideaux B, Via LE, Zimmerman MD, The association between sterilizing activity and drug distribution into tuberculosis lesions, 2015; 21: 1223-1227.

Appendices

Appendices A Potential Origins of Opioids

*The empty form indicates that the substance has no origin in the state.

SubstanceName	KY	OH	PA	VA	WV
Fluorofentanyl	GRAYSON	MONTGOMERY			
3,4-Methylenedioxy U-47700			ALLEGHENY		
Isobutyryl fentanyl		HAMILTON			WOOD
Acetyldihydrocodeine			PHILADELPHIA		
4-Methylfentanyl		LUCAS			
Tetrahydrofuran fentanyl		CLARK	YORK		
U-51754		CUYAHOGA			
p-Fluorofentanyl		CUYAHOGA			HARRISON
o-Fluorofentanyl		GREENE,MARION		STAFFORD	
Furanyl/3-Furanyl fentanyl					WOOD, NICHOLAS , MORGAN
Thebaine		STARK	PIKE		
Dihydrocodeine		LICKING		HENRICO	
Acetylcodeine			PHILADELPHIA		

Figure 13: Table of Substance_1

SubstanceName	KY	OH	PA	VA	WV
Fluorofentanyl	GRAYSON	MONTGOMERY			
3,4-Methylenedioxy U-47700			ALLEGHENY		
Isobutyryl fentanyl		HAMILTON			WOOD
Acetyldihydrocodeine			PHILADELPHIA		
4-Methylfentanyl		LUCAS			
Tetrahydrofuran fentanyl		CLARK	YORK		
U-51754		CUYAHOGA			
p-Fluorofentanyl		CUYAHOGA			HARRISON
o-Fluorofentanyl		GREENE,MARION		STAFFORD	
Furanyl/3-Furanyl fentanyl					WOOD, NICHOLAS , MORGAN
Thebaine		STARK	PIKE		
Dihydrocodeine		LICKING		HENRICO	
Acetylcodeine			PHILADELPHIA		

Figure 14: Table of Substance_2

SubstanceName	KY	OH	PA	VA	WV
Fluorofentanyl	GRAYSON	MONTGOMERY			
3,4-Methylenedioxy U-47700			ALLEGHENY		
Isobutyryl fentanyl		HAMILTON			WOOD
Acetyldihydrocodeine			PHILADELPHIA		
4-Methylfentanyl		LUCAS			
Tetrahydrofuran fentanyl		CLARK	YORK		
U-51754		CUYAHOGA			
p-Fluorofentanyl		CUYAHOGA			HARRISON
o-Fluorofentanyl		GREENE,MARION		STAFFORD	
Furanyl/3-Furanyl fentanyl					WOOD, NICHOLAS , MORGAN
Thebaine		STARK	PIKE		
Dihydrocodeine		LICKING		HENRICO	
Acetylcodeine			PHILADELPHIA		

Figure 15: Table of Substance_3

SubstanceName	KY	OH	PA	VA	WV
Fluorofentanyl	GRAYSON	MONTGOMERY			
3,4-Methylenedioxy U-47700			ALLEGHENY		
Isobutyryl fentanyl		HAMILTON			WOOD
Acetyldihydrocodeine			PHILADELPHIA		
4-Methylfentanyl		LUCAS			
Tetrahydrofuran fentanyl		CLARK	YORK		
U-51754		CUYAHOGA			
p-Fluorofentanyl		CUYAHOGA			HARRISON
o-Fluorofentanyl		GREENE,MARION		STAFFORD	
Furanyl/3-Furanyl fentanyl					WOOD, NICHOLAS , MORGAN
Thebaine		STARK	PIKE		
Dihydrocodeine		LICKING		HENRICO	
Acetylcodeine			PHILADELPHIA		

Figure 16: Table of Substance_4