# Research on Opioid drugs based on Entropy weight method

**Hengtian Yang**[*]

*China University of Mining and Technology, Xuzhou, Jiangsu, 221116, China*

*Corresponding author

*Abstract:* In this paper, we formulate multiple models to account for and predict the spread and characteristics of opioids, which considers socio-economic factors in the United States as well. We preprocess the data provided, including outlier processing and data class configuration. First, we draw the regional scatter distribution map according to the longitude and latitude of each state, and observe the characteristics of events reported between States and counties over time. Then we use the entropy method to determine the weight of each county in each state. We use SI Epidemic Model to draw the conclusion that when Drug Reports reaches 63% of Total Drug Reports State, the region begins to use specific opioids, and predicted that drug use in the Philadelphia region of PA State will first reach the threshold level in 2019.

## 1. Introduction

Opioids play a significant role in pain control. These drugs have antitussive, antidiarrheal, sedative, analgesic and euphoric effects. However, as a growing number of people choose to use synthetic and non-synthetic opioids [1], either for the treatment and management of pain (legal, prescription use) or for recreational purposes (illegal, non-prescription use), addiction problems caused by opioid abuse in the United States are increasing.

According to the study, opioid prescription drug use in the United States increased by 154% from 2010 to 2017, while the number of opiate abusers increased by 90%, which may influence important sectors of the U.S. economy as well. How to address the opioid crisis is a complex challenge for The United States [2].

Federal organizations such as the Centers for Disease Control (CDC) began to seek an appropriate strategy. We are asked to describe the spread and characteristics of opioids in and between the five states and their counties over time, and analyze if its use or trends-in-use is associated with the U.S. socio-economy. We will finally put forward a feasible strategy for the opioid crisis.

## 2. Model preparation

### 2.1 Data Visualization

According to the NFLIS data provided, we visualize the data from the angles of distribution form,

central distribution, correlation, and so on. We present the characteristics of the data in the form of graphs, and analyze the law of case transmission over time. We use hist to draw a histogram of Drug Reports in five states, which more intuitively reflects the degree of data centralization.
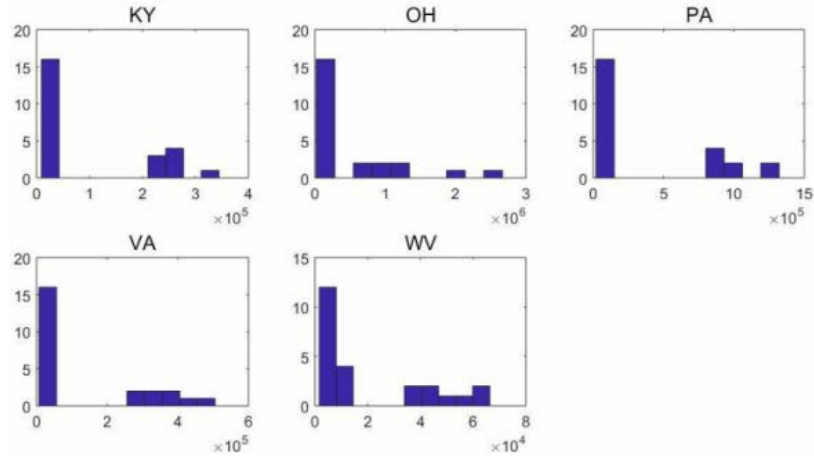


Figure 1: Columnar distribution

## 2.2 Exclusion of outliers

The method of eliminating outliers is based on Pauta criterion, which is called 3σ criterion [3].

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{1}$$

Among them, xi(i=1~628) refers to the original index data, $\bar{x}$ refers to the average value of the original index data.

## 2.3 Dimension Reduction

After obtaining the ideal data source, it is necessary to find out the correlation between each index in the data and the use of opioid drugs. Firstly, each index is regarded as a latitude, and the correlation between each latitude and dependent variables is analyzed separately [4]. According to the correlation analysis method, the original data and the cleaned data are imported into SPSS software respectively, and the scatter plot analysis and correlation analysis are carried out. The results are as follows:
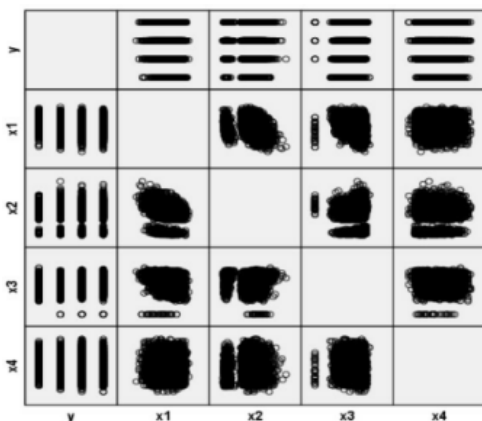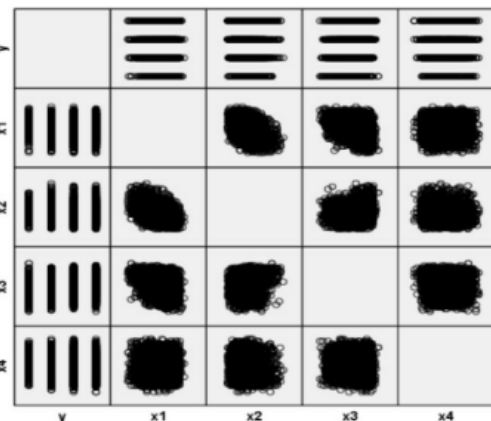


Figure 2: Scatter plot before preprocessing      Figure 3: Scatter plot after preprocessing

57

## 3. Dissemination and characteristics of drugs

We found the longitude and latitude coordinates of each county in the United States from GitHub and projected them onto maps of five states, drawing scatter maps of regional distribution [5].
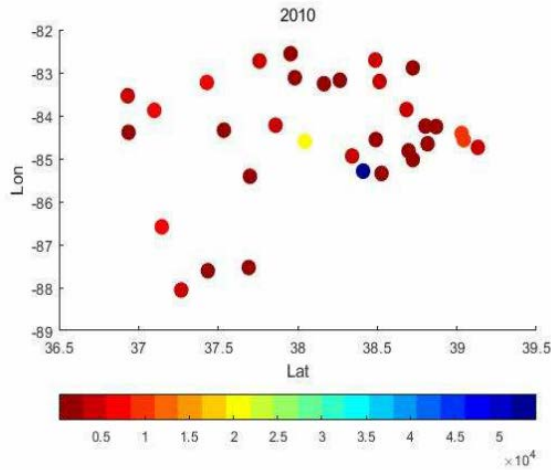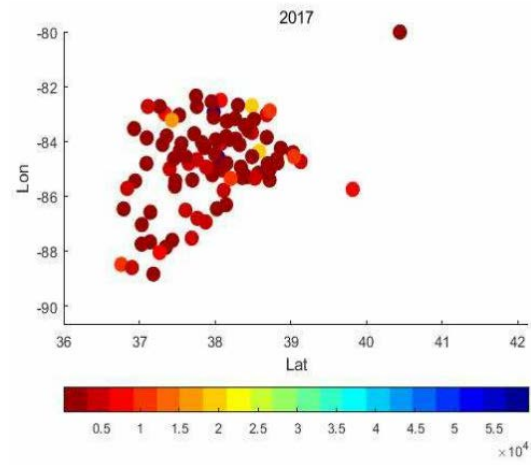


*Figure 4: Year 2010 Scatter plot*  *Figure 5: Year 2017 Scatter plot*

According to the trend of the line chart drawn, we draw the conclusion that the spread and characteristics of opioids and heroin and between the five states and their counties over time.
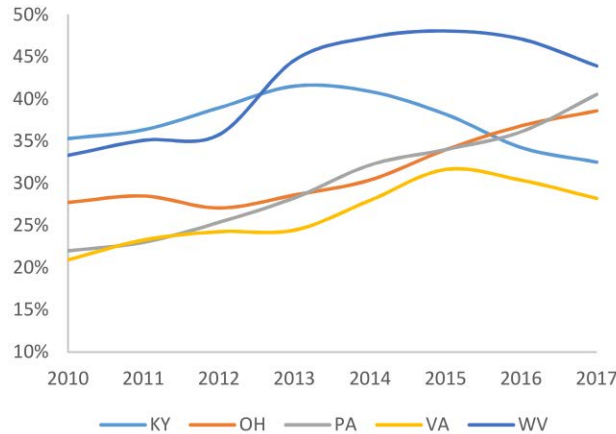


*Figure 6: Drug Reports*

According to the distribution of time series, the Total Drug Reports State of OH in the five continents is increasing, and its Drug Reports are also increasing. According to the order of time, Total Drug Reports State and Drug Reports in KY and WV states are in a certain range and tend to be stable.

## 4. Potential location of opioids

## 4.1 Establishment of Entropy Method Model

Taking the quantity of drugs as the evaluation index, the weight distribution is calculated by using the method of entropy value. Set up the original data matrix X with *j* drug indexes and *i* regions. For a certain index, the greater the difference between the standard values, the greater the role of the index

in the comprehensive evaluation, and vice versa. If the index values of a certain index are all equal, then the index does not work in the comprehensive evaluation. In information theory, there exists a functional relationship:

$$H(x) = -\sum f(x_k) \ln f(x_k) \tag{2}$$

The absolute values of the two are equal and the symbols are opposite. The more discrete the index value of a certain index is, the smaller it will be, and the more information it provides, so the weight should also be greater; conversely, the smaller the degree of discreteness between the indicators, the smaller the information entropy, the smaller the information it contains, and the smaller the weight should be.

## 4.2 Solution of the Model

Step1: Calculate the proportion of the index value in $i$ year under $j$ index $p_{ij}$.

$$p_{ij} = \frac{x_{ij}}{\sum x_{ij}} \tag{3}$$

Step 2: Calculating the entropy of index.

$$e_j = -(1/\ln m)\sum p_{ij} \ln p_{ij} \tag{4}$$

Step 3: Calculating the difference coefficient $g_j$ of index $j$. The smaller the entropy value, the greater the difference between the indicators, the more important the indicators are $g_j=1-e_j$.
Step 4: Defining the weight $a_j$ of item j:

$$a_j = g_j / \sum g_j \tag{5}$$

Step 5: Calculating index degree $D_i$ of area $i$:

$$D_i = \sum a_j p_{ij} \tag{6}$$

By calculating, it can be concluded that the counties with higher weights in each state are as follows:

*Table 1: Weight of countries*

| Counties | JEFFERSON | FAYETTE | FAYETTE | MONTGOMERY | PHILADELPHIA |
|----------|-----------|---------|---------|------------|--------------|
| Weight | 0.105 | 0.046 | 0.171 | 0.152 | 0.133 |
| Counties | ALLEGHENY | HENRICO | CHESTERFIELD | KANAWHA | BERKELEY |
| Weight | 0.105 | 0.082 | 0.096 | 0.062 | 0.067 |

After analyzing the distribution entropy weights of these drugs in different counties, we can see that the counties where specific opiates have been used are mainly Jefferson and Fayette in KY, Hamilton and Mongtoery in OH, Philadelphia and Alegheny in PA, Henrico and Chesterfield in V A, Kanawha and Berkeley in WV.

## 5. Conclusion

This paper focuses on the spread and characteristics of opioids and takes into account the socio-economic factors in the United States. First of all, this paper preprocesses the provided data, including exception handling and data class configuration. The regional dispersion distribution map is drawn according to the longitude and latitude of each state, and the characteristics of events reported between

states and counties are observed over time. Then we use the entropy method to determine the weight of each state and each county. Using the SI epidemic model, we concluded that when drug reports reached 63% of the state's total drug reports, the area began to use specific opioids.

## References

[1] Ni Ning, Zhang Shaoquan, Chen Xiaoyun, Zhang Xiaoyu. Audit risk Assessment based on Entropy weight method and Fuzzy Analytic Hierarchy Process [J]. Software, 2018, 39 (10): 254-259.

[2] Ren Xue. Studies on some problems of three kinds of infectious disease models [D]. Northeast Forestry University, 2018.

[3] Xu Chonggang, Hu Y uanman, Chang Y u, Jiang Y an, Li Xiuzhen, Burencang, he Hongshi. Sensitivity Analysis of Ecological Model [J]. Journal of Applied Ecology, 2004 (06): 1056-1062.

[4] Qiao Jia-jun. Application of improved Entropy method in Evaluation of Sustainable Development ability in Henan Province [J]. Resource Science, 2004 (01): 113-119.

[5] Ding Guoqing, Ji Tongyuan, Xu Liang, Qin Negen. Research on Technical State of High pile Wharf based on Analytic Hierarchy Process and Fuzzy Comprehensive Evaluation method [J]. Hong Kong Technology, 2018 55 (06): 100-103.