https://cdn.aarp.net/content/dam/aarp/health/conditions_treatments/2017/05/1140-pill-usa-opioids-aarp.imgcache.reva5ab0a8b1d6d8e63bb527fd5ccc8ea95.jpg

Analysis of Medicaid Prescription Data in the USA

CS435 Big Data - Fall 2018

Andre Hochmuth, Joaquín Cuomo, Josh Munoz, Mitchell Rouault

Problem Formulation

According to the CDC since 1990 there has been an increase in deaths caused by prescription opioid overdoses[1] [2]and on October 26, 2017, President Donald Trump declared the opioids addiction a Nationwide Public Health Emergency[3].
Despite the fact that prescription opioids are intended for patients, according to the American Public Health Association most of the users misusing opioids declared obtaining them from a prescription[4]. Therefore, the term 'users' instead of 'patients' is going to be used.

Opioids are powerful painkillers that bind to receptors in the brain preventing the pain signals from being triggered. Also, they release dopamine which is involved in the 'reward-system' of the brain causing euphoria. Moreover, these drugs develop tolerance which means that the users require larger doses to achieve equal effect as before[5]. Most commonly prescribed opioids are morphine and codeine (natural), hydrocodone and oxycodone (semi-synthetic), and fentanyl and methadone (synthetic). Besides all the counter effects that the drug can cause in the long term, it affects the respiratory system as it has sedative effects which may lead to respiratory failure and death[6].

The magnitude of the opioids consumption call for an urgent solution which requires

[1] "Opioid Data Analysis and Resources | Drug Overdose | CDC Injury ...." 9 feb.. 2017, https://www.cdc.gov/drugoverdose/data/analysis.html. Consulted on 27 nov.. 2018.
[2] "Overdose Death Rates - National Institute ...." https://www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rates. Consulted on 27 nov.. 2018.
[3] "The Opioid Crisis | The White House." https://www.whitehouse.gov/opioids/. Consulted on 27 nov.. 2018.
[4] "Quantifying the Epidemic of Prescription Opioid Overdose Deaths ...." 1 abr.. 2018, https://ajph.aphapublications.org/doi/10.2105/AJPH.2017.304265. Consulted on 27 nov.. 2018.
[5] "4: Opiates binding to opiate receptors in the nucleus accumbens ...." https://www.drugabuse.gov/publications/teaching-packets/neurobiology-drug-addiction/section-iii-action-heroin-morphine/4-opiates-binding-to-opiate-rece. Consulted on 27 nov.. 2018.
[6] "WHO | Information sheet on opioid overdose - World Health Organization." https://www.who.int/substance_abuse/information-sheet/en/. Consulted on 27 nov.. 2018.

information to be addressed properly. Such a comprehensive analysis should include different parameters related to opioids such as sales of prescribed and illegal drugs, number of deaths caused by opioids, revenue of pharmaceuticals due to opioids and drug abuse prevention systems and others.

Our main hypothesis will be that opioids are increasing in a dramatic fashion because they are being used as a replacement for other type of painkillers. We also will focus on evaluating trends in the amount of prescription opioid usage and determine if the 'opioids epidemic' is actually only related to opioids or if it is an overall problem for all other types of prescription drugs across the United States. Also, we will expect to see some decrease in the prescribed opioids since the National Emergency was declared.

A nationwide controversy that has sprung up in the past decade revolves around the topic of prescription drugs. It is extremely easy to get prescriptions from doctors to obtain painkillers and other drugs. This has lead to a drug epidemic in the United States and with the introduction of certain policies such as Medicaid it has become even easier to access these drugs. "Drug overdoses are now the leading cause of injury death in the United States, outnumbering both traffic crashes and gun-related deaths" (Sobriety Nation).
The Medicaid Drug Rebate Program helps offset federal and state costs of prescription drugs to Medicaid patients. A result of this program requires a record of all prescriptions to Medicaid patients to be kept. This program has been collecting records since 1991, resulting in many gigs of data, and all of it is available to the public.

The government could use this project as a tool for spreading awareness and shedding light on the issue at hand. This analysis could even be used to highlight specific trends at a per state level.

Methodology

I. Algorithms

1) Filter and Process data

The first step was to filter the original dataset and keep only the medicaments that were opioid or non-opioid painkillers, discarding all other type of drugs. Having the proper list of medicaments to filter the database was crucial to prevent any kind of skew in our analysis. Therefore, we use the CDC list of opioids[7] that is used to analyze opioid problematic, which exclude some drugs as methadone which is prescribed for opioid addiction recovery. For the non-opioid list we had to decide which non-opioid painkillers we were going to use between all possible types of drugs used as painkillers[8]:

- Steroids
- Bisphosphonates
- Antidepressants
- Drugs to prevent fits (anticonvulsants)
- Local anaesthetics
- Monoclonal antibodies
- Non steroids anti-inflammatory (NSAIDs)

Because all those drugs are used mainly with other purpose than painkiller except for the NSAIDs we decided to only use this last one in our correlation analysis. Our decision was based on two more factors[9]:

1. NSAIDs are non addictive
2. They are the most common drug prescribed for mild pain

To create the NSAIDs list we looked for a list of NSAID in the official website of US Drugs and Aliments[10]. Then we looked one by one at those drug names in the official database of the FDA[11] and saved the files as CSV.

---

[7] "Data Resources | Drug Overdose | CDC Injury Center." 19 oct.. 2018, https://www.cdc.gov/drugoverdose/resources/data.html. Consulted on 27 nov.. 2018.
[8] "Other drugs for pain control | Coping with cancer | Cancer Research UK." https://www.cancerresearchuk.org/about-cancer/coping/physically/cancer-and-pain-control/treating-pain/painkillers/types-of-painkillers/other-drugs. Consulted on 27 nov.. 2018.
[9] "Pain Relievers: MedlinePlus." 29 ago.. 2018, https://medlineplus.gov/painrelievers.html. Consulted on 27 nov.. 2018.
[10] "Drug Safety and Availability > FDA Drug Safety Communication: FDA ...." 26 feb.. 2018, https://www.fda.gov/Drugs/DrugSafety/ucm451800.htm. Consulted on 27 nov.. 2018.
[11] "National Drug Code Directory - FDA." 4 feb.. 2016, https://www.accessdata.fda.gov/scripts/cder/ndc/. Consulted on 27 nov.. 2018.

For filtering both opioids and nonopioids (only NSAIDs type) from our initial dataset, we used distributed cache to load these tables (listing the types of drugs we wanted to keep). Once loaded in cache we were able to match them with the medicaid database and discard the rest.

The inputs are:

- Medicaid_database.csv (list of amount of all prescribed drugs per year)
- opioid/opioids_list.csv (list of opioids)
- nonopioid/aspirin.csv, ibuprofen.csv, etc. (lists of NSAIDs)

As opioids and non-opioids are not datasets from the same source the parsing is not the same, but we merged both into the same program. Therefore, there is an argument to tell if we are processing opioids or non-opioids. Also, there is an argument to specify if we want to preserve state field to do a statewide or nationwide analysis.

2) Correlational Analysis

For proving our hypothesis that opioids are replacing non-opioids we thought of computing different correlations:

Given:

$Opioids:\ o_1,\ o_2,\ \dots,\ o_N$

$NSAIDs:\ n_1,\ n_2,\ \dots,\ n_M$

Number of states: S (in Nationwide analysis S=1)

We wanted to compute the following correlations nationwide and statewide:

1. $corr(o_i, n_j)\ \ \forall i \in [0, N] \wedge \forall j \in [0, M]$

   The correlation between all opioids and all non-opioids is the one with most heavily computation requirements as the output has a size in the order of $N\ x\ M\ x\ S$. We need to be careful with the results as many drugs might have zero values for many years and those will have a high positive correlation. In this analysis high negative correlations might be more representative.

2. $corr(\sum_{i=0}^{N} o_i\ , \sum_{j=0}^{M} n_j\ )$

   The correlation between the sum of all opioids and all non-opioids can be more statistically accurate than the previous one, but some skew might be still present due to the difference between N and M.

3. $corr(topK(o)_i\ , topK(n)_j)\ \ \forall i \in [0, K] \wedge \forall j \in [0, K]$

To overcome the previous scenery we thought on computing the correlation between the top K opioids and the top K non-opioids. Top K are calculated summing amounts of the same drugs across the entire time frame and keeping those with the greatest accumulated prescribed amount.

4. $corr(\sum_{i=0}^{K} topK(o)_i \ , \sum_{j=0}^{K} topK(n)_j)$

We could average previous correlations or compute the correlation between the sum of the top K opioids and nonopioids to obtain a unique value to prove or reject our hypothesis.

For programming the correlation we used the cartesian product[12] pattern and then the mathematical formula of correlation. The correlations were done both normalizing the dataset to percentage and without normalization.

All correlation (abs value) above 0.4 are statistically significant due to the sample size of 27 years. We use |corr|>0.7 to consider there are highly correlated.

3) Trend Analysis (National and Statewide)

The trend analysis will be useful both for viewing clearly the correlation results but more important for detecting any other potential hypothesis or conclusion. In order to detect trends in opioid and non-opioid filtered datasets we will plot all of the drugs by their total reimbursement amount each year from 1992-2017. Years will be the independent variable and total reimbursement amount will be the dependent variable on our plot. Their will be one plot for opioids and one for non-opioids. Each plot will have many lines for the different drugs. Since we are not concerned with seeing trends with specific drugs, we can plot these lines without a label/key as long as we can tell if it is an opioid or nonopioid. We can also plot the top 10 to see if it better represents trends since some drugs may have unusual behaviors that are not related to the whole.

We can then average the total reimbursement amount for each year for all the drugs in each type (opioid/non-opioid) and plot these values to come up with a single line graph that sums up the national trends of opioids and nonopioids. We cannot compare the two drug groups though due to inconsistency factors between the two. Regardless, we can look at each group separately and analyze trends from the plot of the total average reimbursement amount for all drugs in each category (opioid/non-opioid) over time.

We will also come up with a linear regression model that can accurately model the data of the average total reimbursement amounts for each drug category. We will then assess the

[12] "hadoop-map-reduce-patterns/CartesianCommentComparison.java at ...." https://github.com/geftimov/hadoop-map-reduce-patterns/blob/master/src/main/java/com/eftimoff/mapreduce/joins/cartesianproduct/CartesianCommentComparison.java. Consulted on 27 nov.. 2018.

accuracy of the linear regression model and how well the data fits a linear regression model. In the case that the data does fit a linear regression model, we can also read values from the linear regression model that tell us other information such as the correlation of total units reimbursed every year. In this case it ends up being a positive correlation for both, which can be attributed to a population increase in the United States.

In order to analyze statewide trends in drug reimbursement values, we will sum all the drugs for each state per year. We will then plot these results using a choropleth map in using Plotly in Python. This will take up one map per year, so we will only keep the years 2000, 2005, 2010, and 2015.

II. Framework

A. Hadoop

Large-Scale framework running on a cluster of 8 nodes.

B. Python Pandas Library

Python Pandas can process large amounts of data depending on the computer hardware being used. Data can also be vectorized or read in chunks to save memory. Python Pandas can be used to work around using large-scale frameworks in certain cases

C. Jupyter Notebooks (Python)

Jupyter Notebooks were used to showcase data modeling and analysis once finalized data was created using Hadoop. For example, the Top K program implementation gets us the top drugs to work with in a much smaller dataset. Multiple libraries were used in Python, including MatPlotLib, PlotLy, StatsModels Api, etc.

III. Comparative Analytics

A. Heat Maps Displaying Various Correlations

B. Plot displaying Opioid to Opioid Correlation Nationwide.

IV. Data Modeling

1) Choropleth Map of United States Total Drug Reimbursement Values By State
2) Various Graphs of Trends over Time on the National Level (Opioid and Nonopioid)

Dataset

Medicaid database

https://www.medicaid.gov/medicaid/prescription-drugs/state-drug-utilization-data/index.html

There are a total of 20 fields in this dataset. The field names to be used in the project include utilization type, state, product code, year, quarter, product name, and number of prescriptions. Utilization type field description is a unique record identifier. This field would be used as a key. Product code is a four digit segment of the National Drug Code (NDC) representing a product. The other fields are self explanatory. Some of the data is missing certain fields such as number of prescriptions. For several legal reasons, some data is not available to the public. For instance, there is no data on the amount of prescriptions for morphine. This will not be a problem because the majority of the data is complete.

http://www.netstate.com/states/tables/st_population.htm
This dataset was used to get state populations. These state populations were used to normalize data when creating visualizations in order to account for the differences in state populations. The category of "percentage of total U.S. population" was used.

Drug lists

Official NDC > https://www.accessdata.fda.gov/scripts/cder/ndc/index.cfm
CDC opioids > https://www.cdc.gov/drugoverdose/resources/data.html
NSAIDs > https://www.fda.gov/Drugs/DrugSafety/ucm451800.htm
These databases were used to retrieve a list of opioids and nonopioids.

State Population Estimates

https://www.census.gov/data/tables/time-series/demo/popest/intercensal-2000-2010-state.html

https://www.census.gov/data/tables/2017/demo/popest/state-total.html

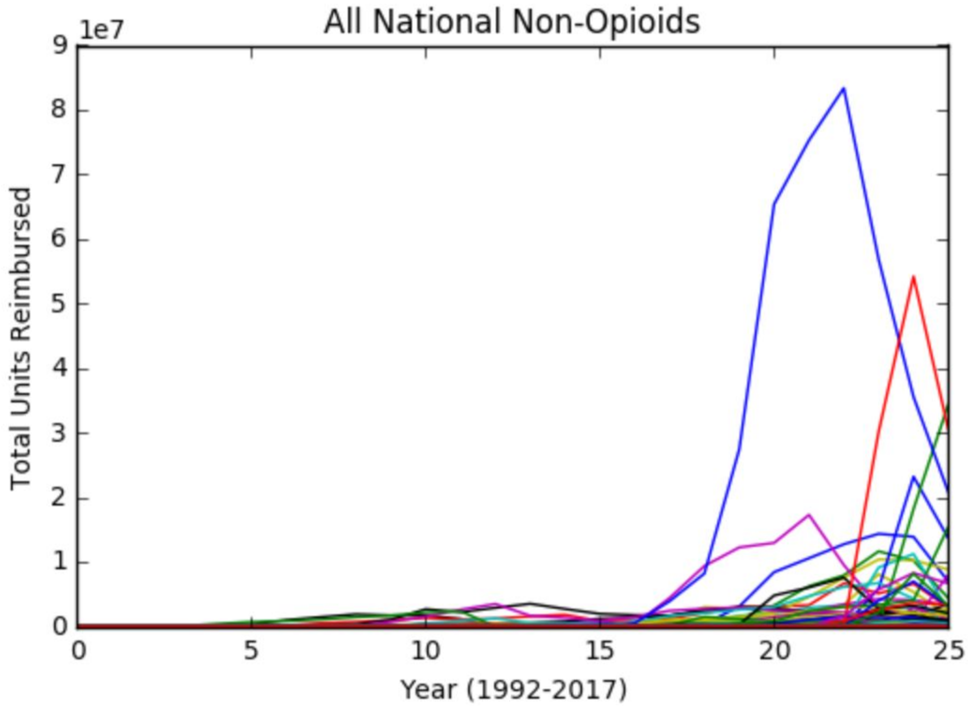
These databases were combined to create one file of state population data from 2000-2017.

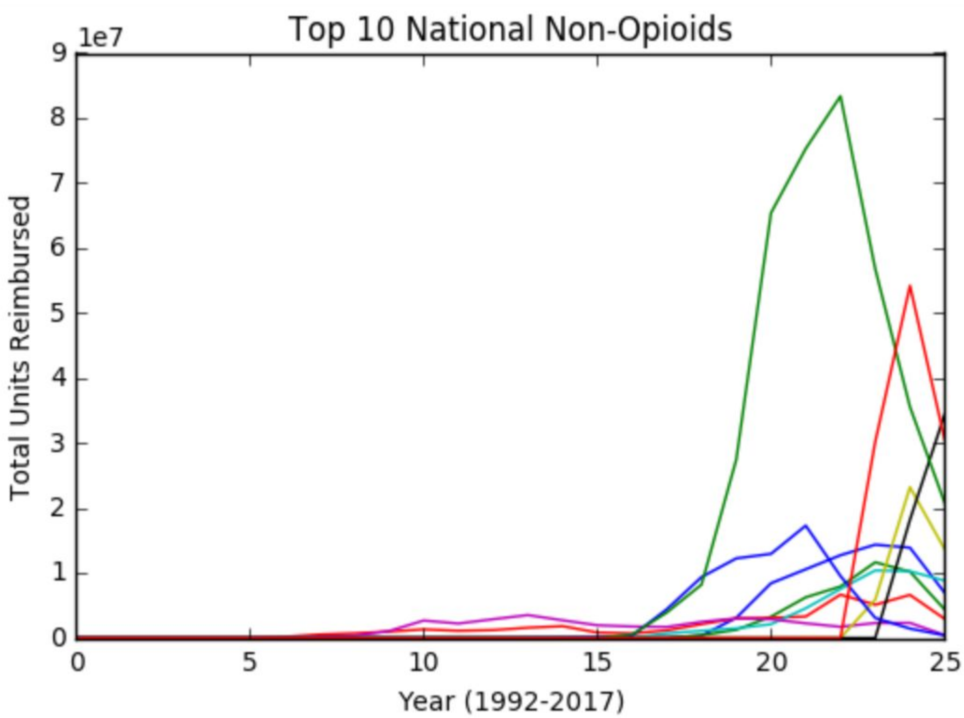
Discussion and Analysis

I. Results

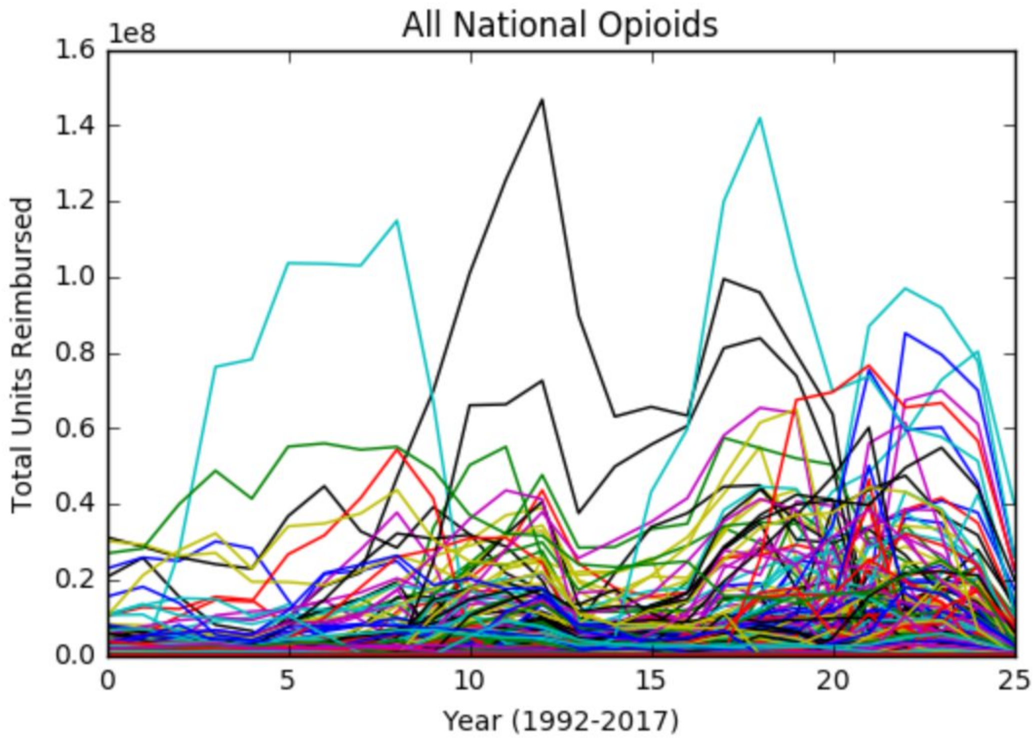
The results of our analysis clearly display a spike in drug usage in the last decade. To our surprise non-opioids have surged since 2008. Even more surprising, the total of non-opioid painkillers shot past the total amount of opioids in 2015 and continues to grow higher. The national average of all opioids and the top 10 national opioids have been in decline since 2016. We also found high positive correlations between the topK opioids and topK non-opioids creating evidence that these drugs are following similar trends. We discovered that there is higher correlation between opioids and nonopioids within the same state as opposed to outside states. In fact, there was often no correlations between seperate states. Nationwide, opioids had many large correlations to other opioids, and while these correlations may be unreliable, there were few inverse relationships. This would suggest most opioids are following the same pattern.

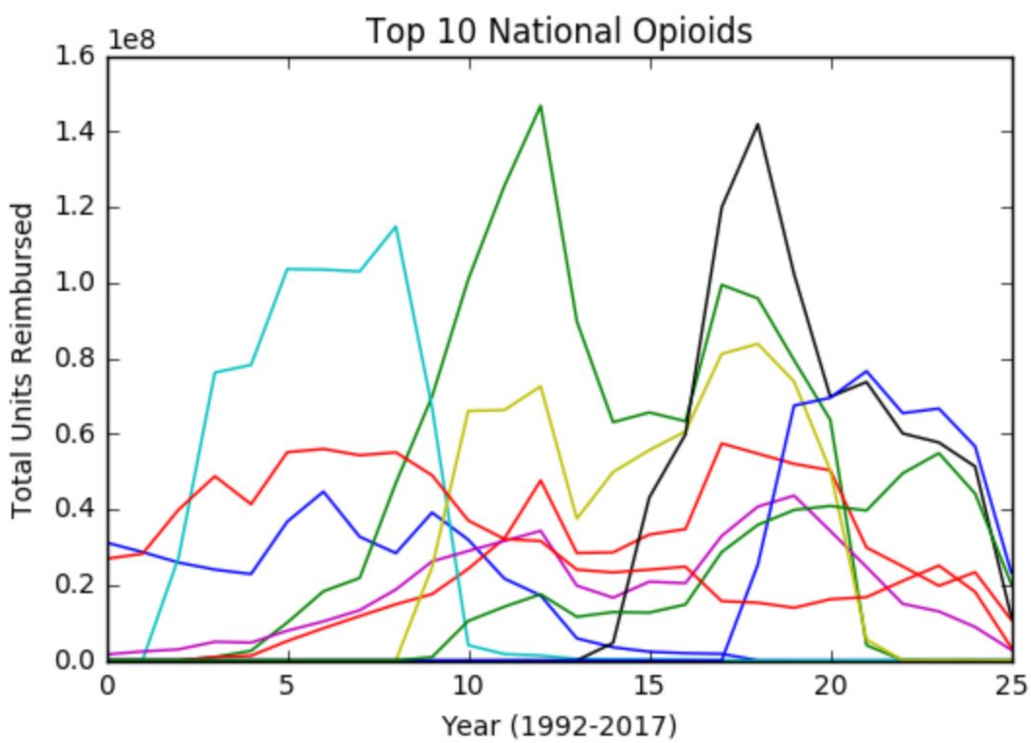
National Trends of Opioids and Nonopioids[13]

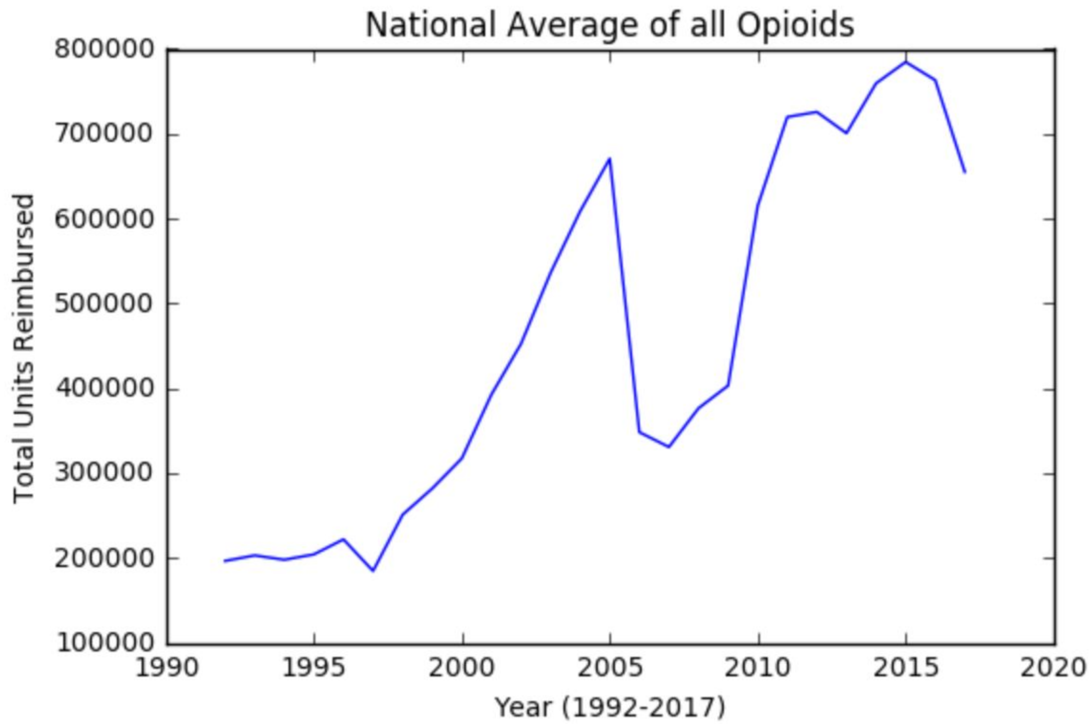


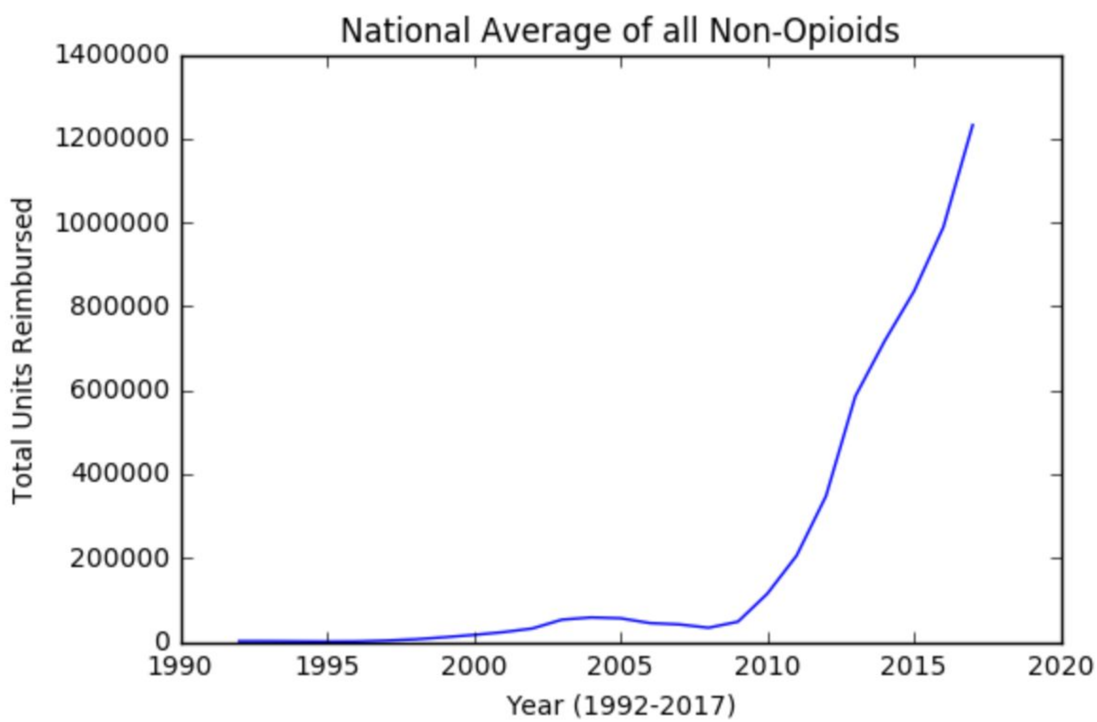


[13] "Jupyter Notebook Tutorial | plotly." https://plot.ly/python/ipython-notebook-tutorial/. Se consultó el 28 nov.. 2018.


All National Opioids
1e8
Total Units Reimbursed
Year (1992-2017)
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25


Top 10 National Opioids
1e8
Total Units Reimbursed
Year (1992-2017)
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25

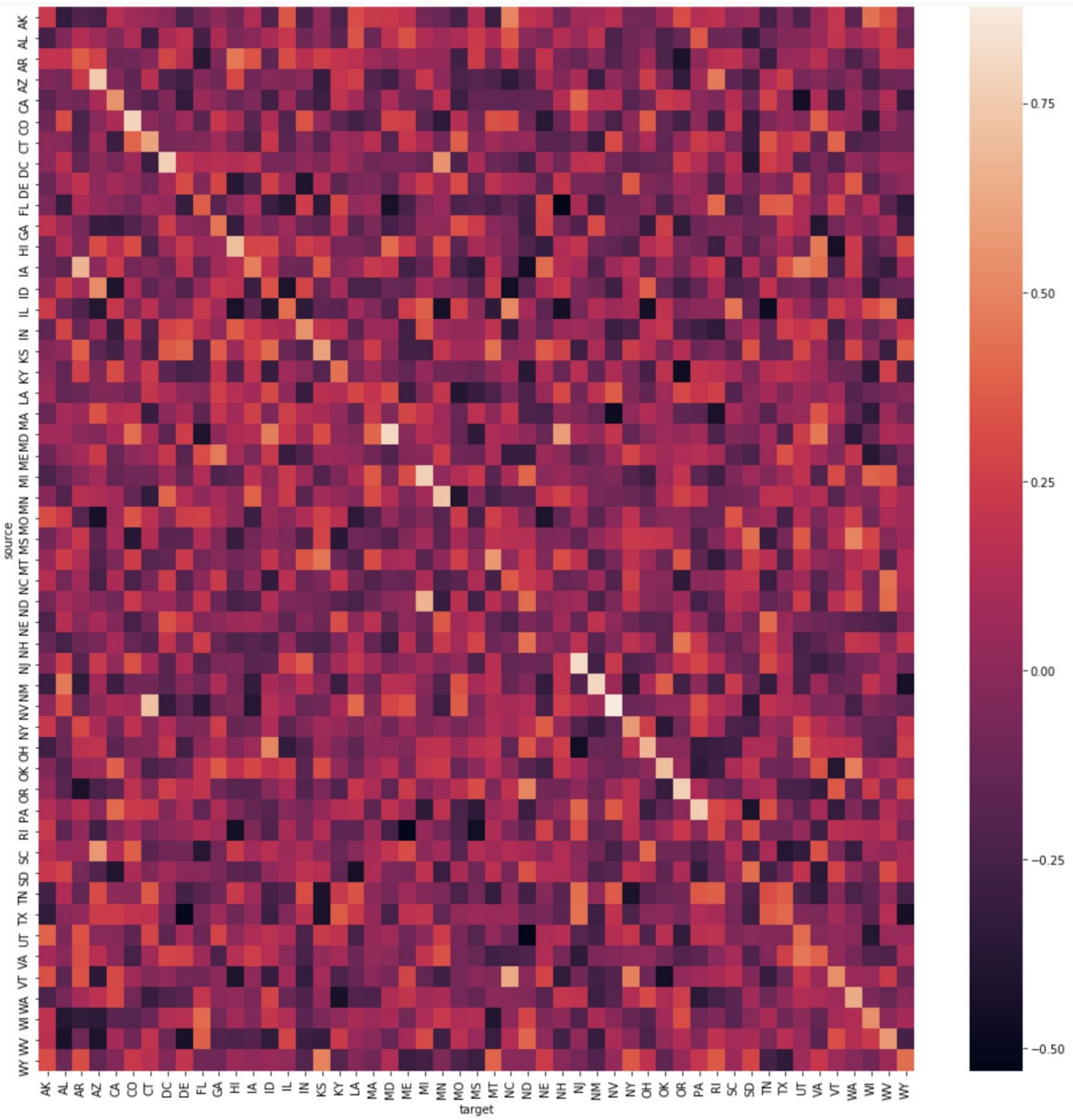




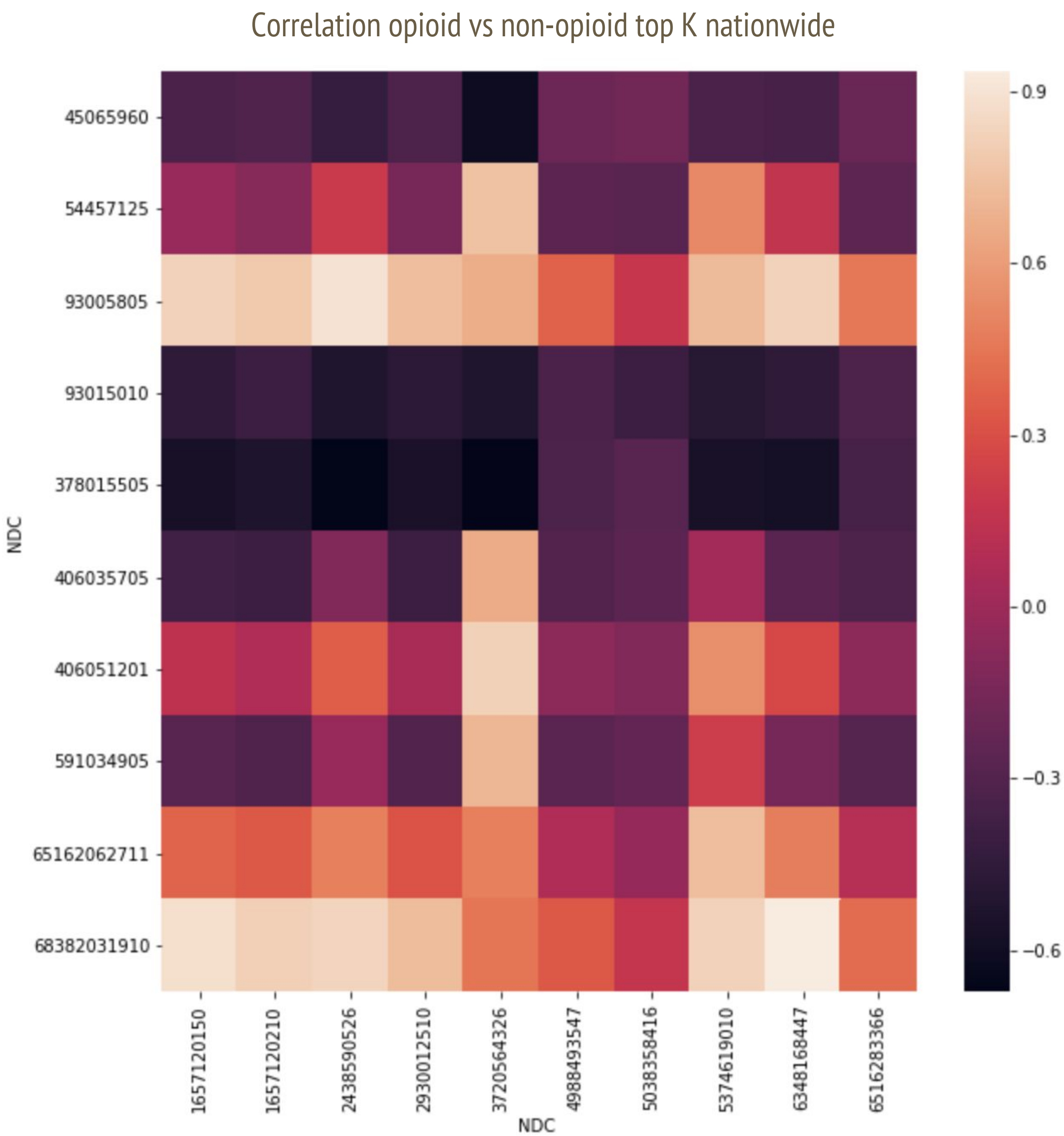
In this graph an overall increase in non-opioid can be seen after 2010 which coincides with an increase in opioids shown in the previous graph of the national average of opioids in 2010.

Correlations Analysis

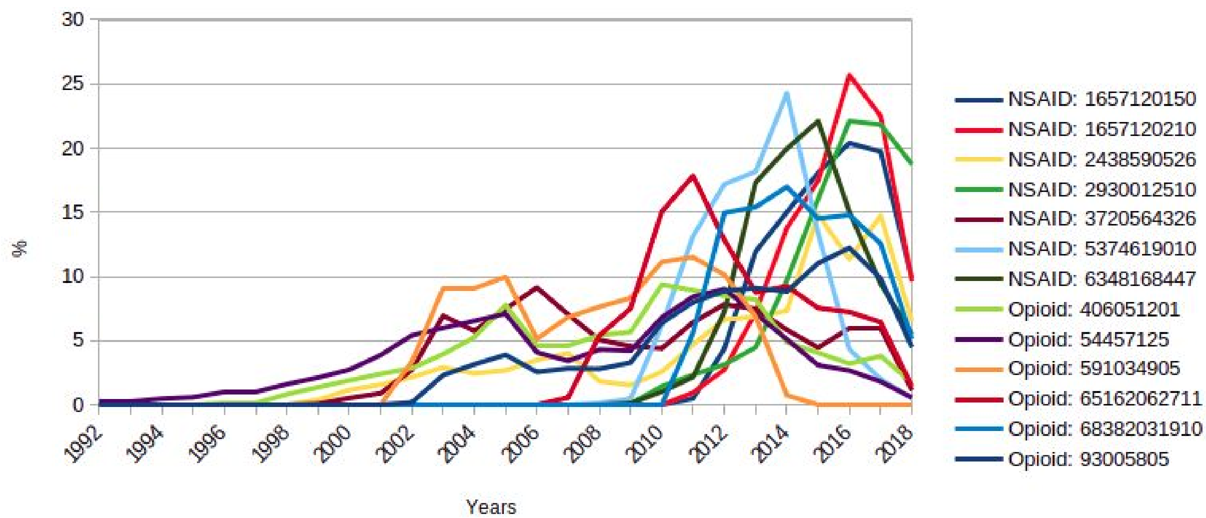
Correlation opioid vs non-opioid acum. statewise



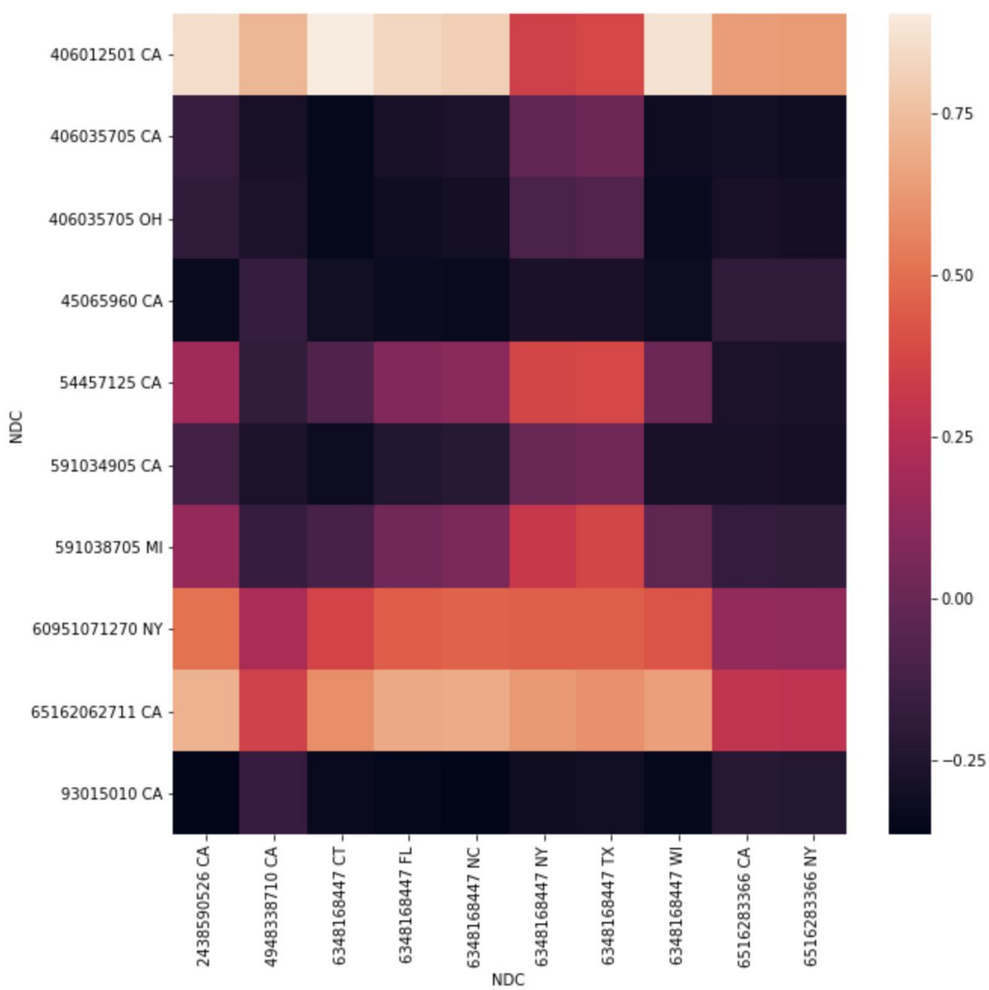
The above heat map indicates that there is a high positive correlations within each state between opioids and non-opioids, but not between different states. This can suggest that there might be missing some standardization to be able to compare across states, or just that the behaviour varies greatly among states.



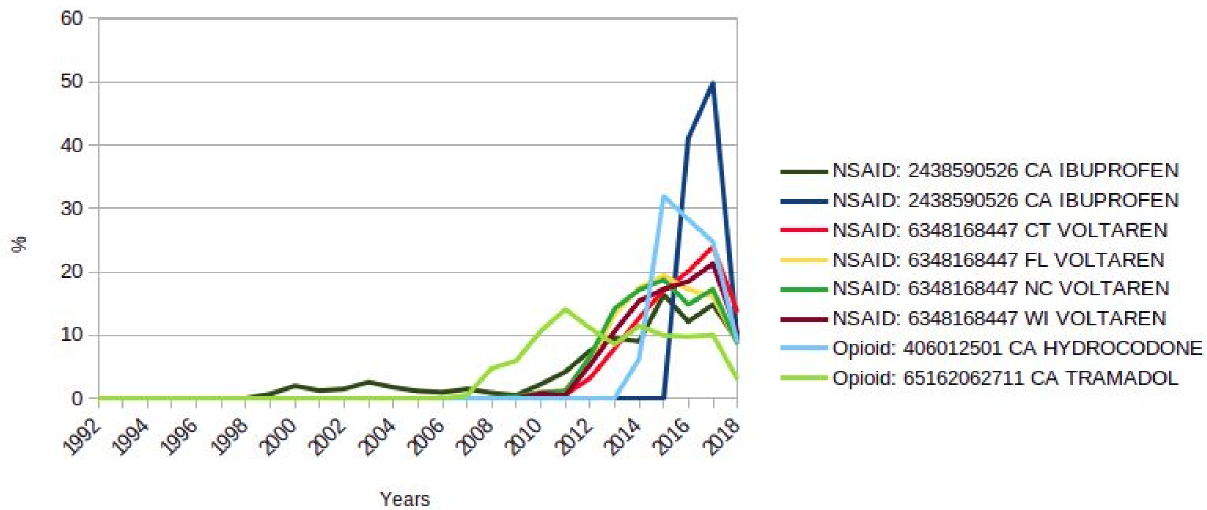
When we compute the correlation between top 10 opioids and nonopioids nationwide we can see that some high positive correlations are found. Therefore, we plotted the data below to observe the trends and visually conclude a common behaviour. This is probably one of the most representative analysis (as explained in previous section) and it contributes to the conclusion that our hypothesis of a non-opioid replacement by opioids if not valid as both types of drugs show a similar tendency to grow.



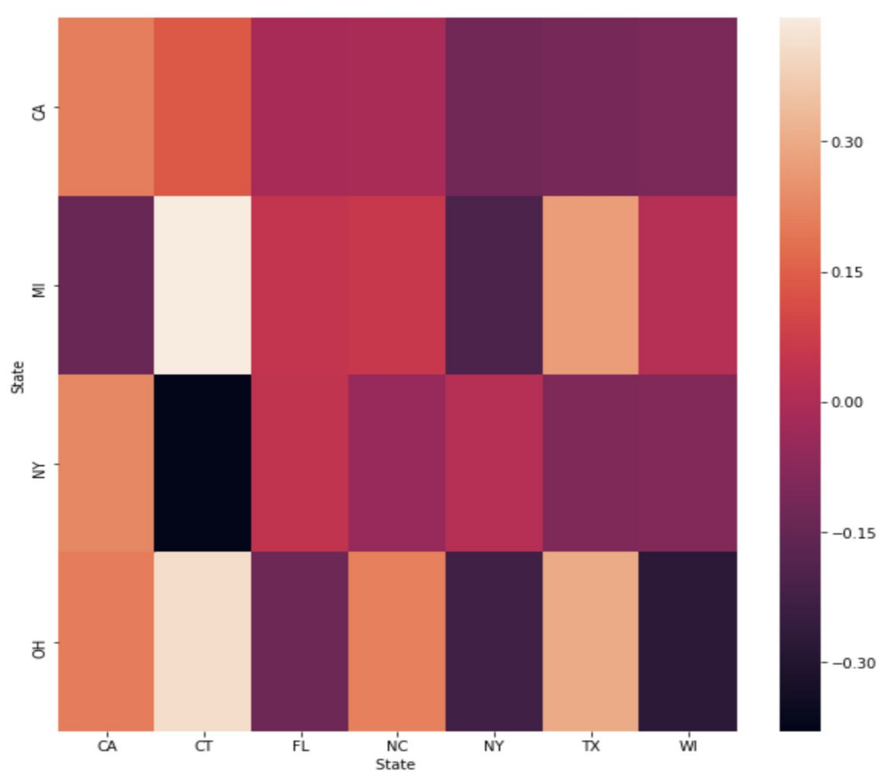
Correlation opioid vs non-opioid top 10 acum. statewise



In the graph above it can be seen that two opioids (Y-axis) are highly correlated with several NSAIDs (X-axis). To better understand these results we plot the trends of those drugs and we can see in the trend below that the reason of these high positive correlations are due to the rapidly increase of the number of prescription in the last decade. The results are similar to the analysis done nationwide with the top 10 drug of each group.

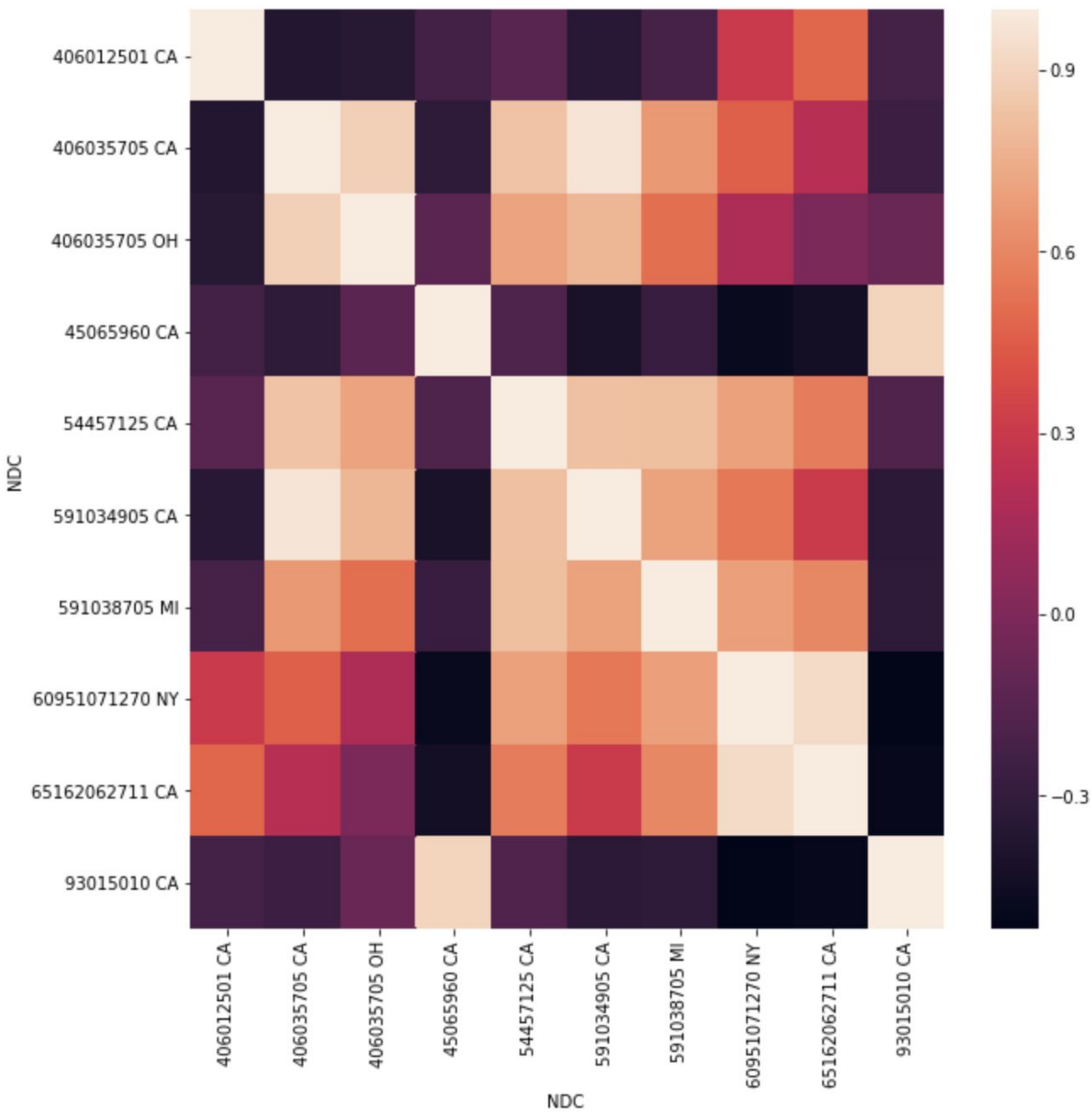


Correlation opioid vs non-opioid top K acum. statewise



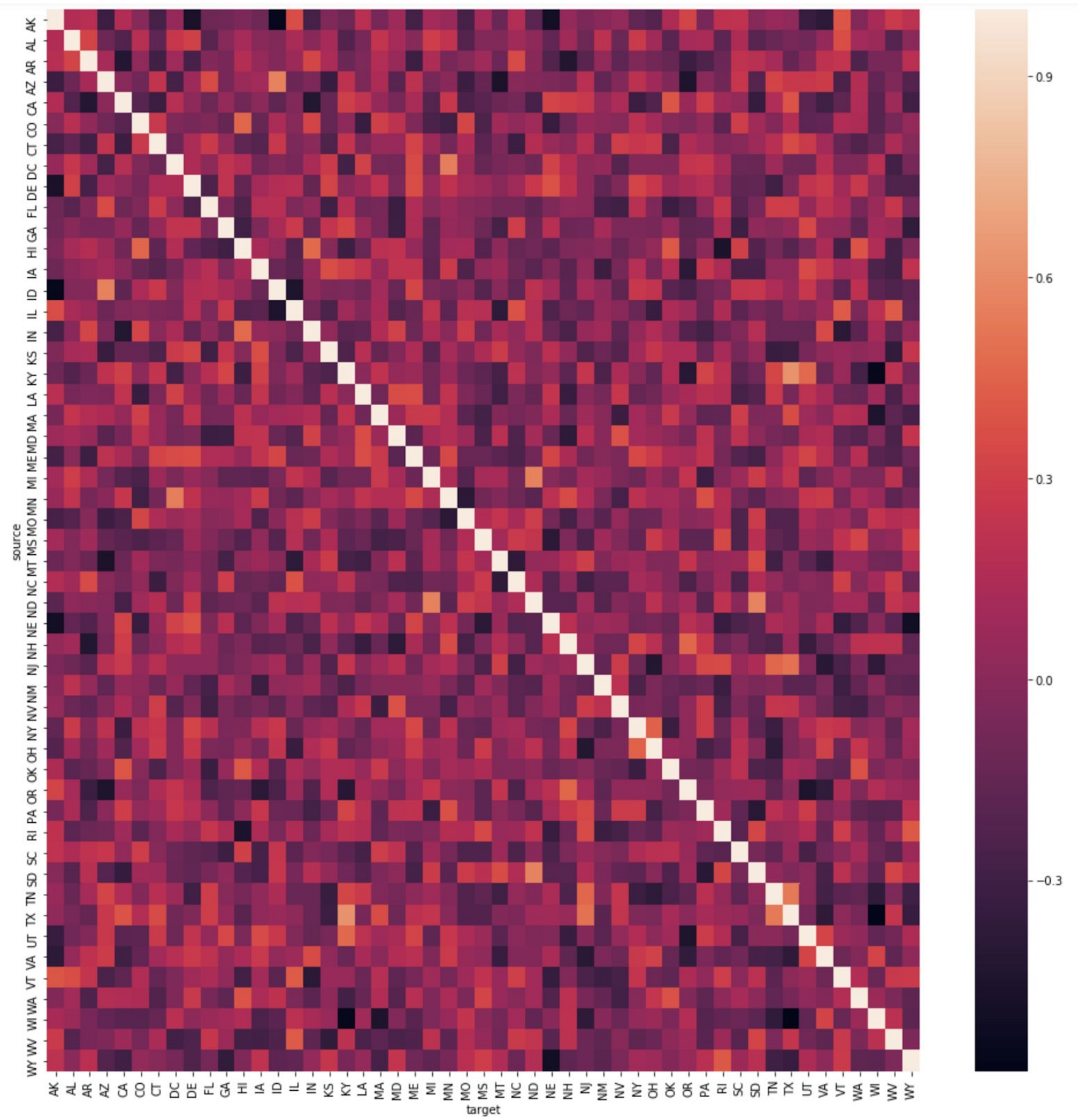
When accumulating the drugs per state per year the positive correlations decrease indicating that there might not be an average common trend between both types of drugs. This can be cause for example for state policies on certain medications which they vary among states.

Correlation opioid vs opioid top K statewide

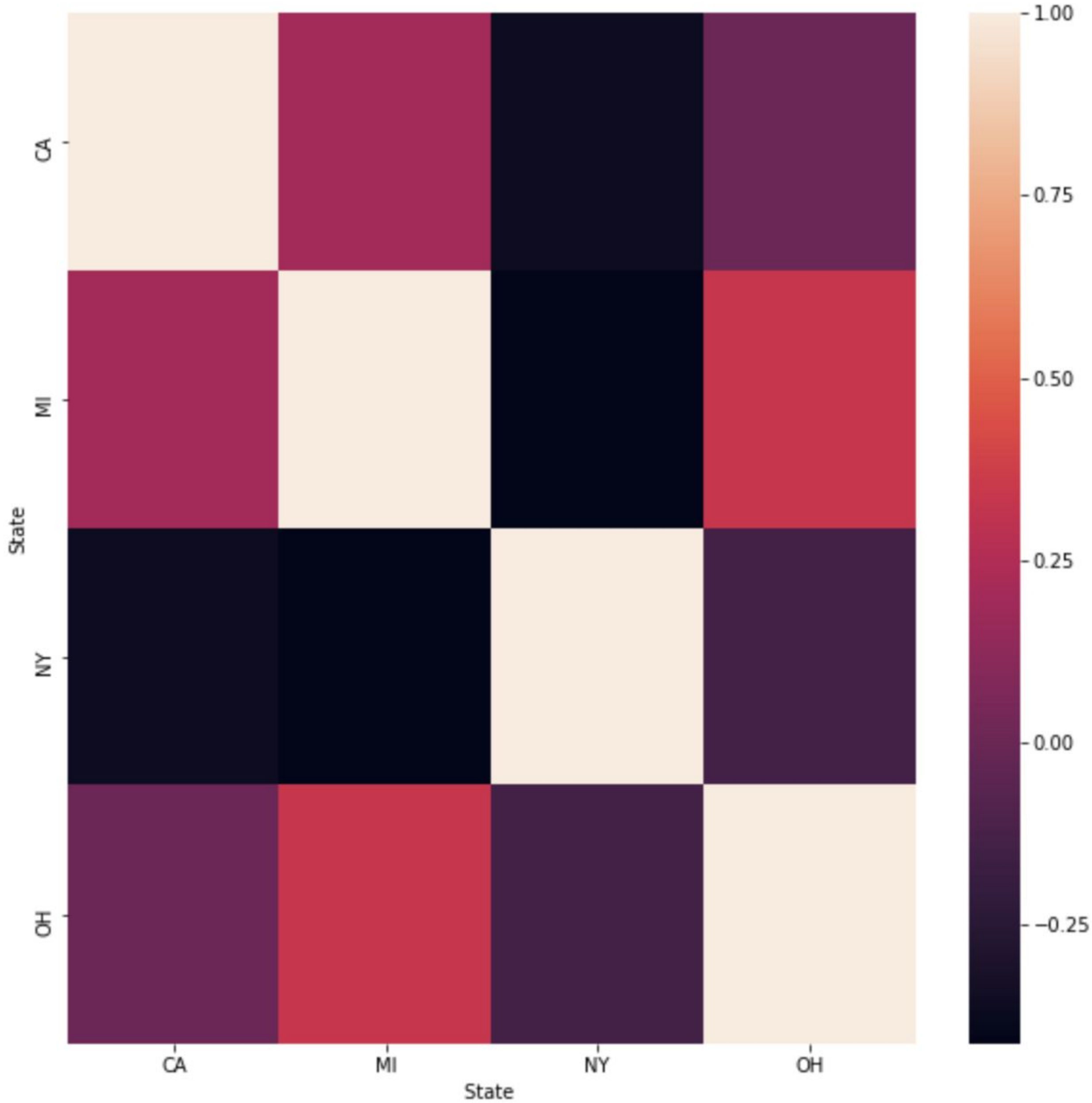


The above correlations are skewed as in the top 10 opioid 70% are from California. Therefore, we can conclude than within the state of California there is a high positive correlation between the most sold opioids, but not nationwide.

Correlation opioid vs opioid acum. statewise

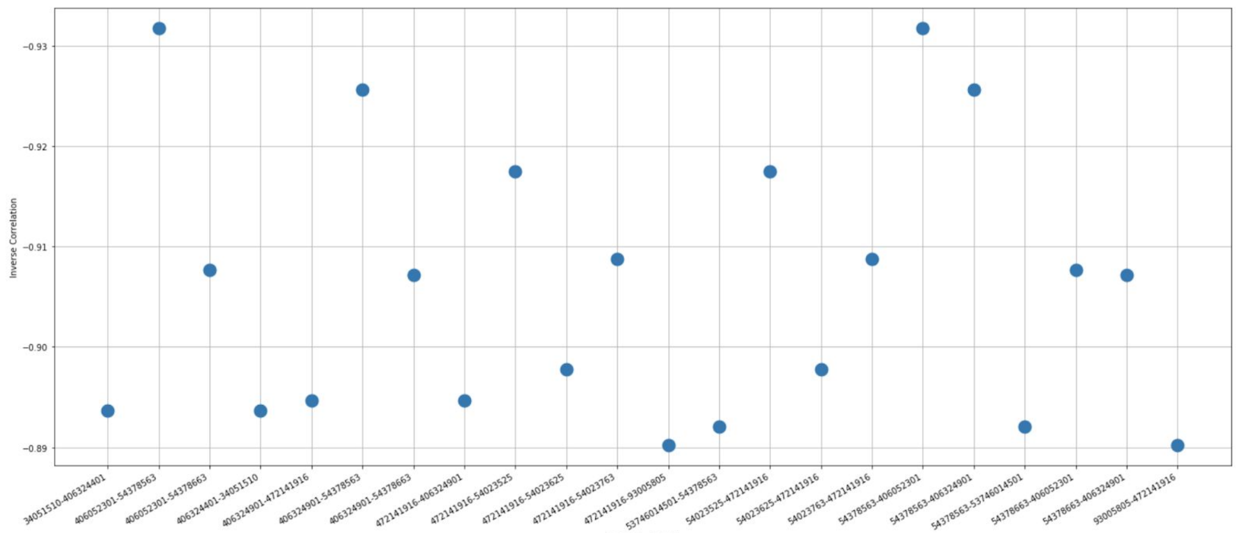


Correlation opioid vs opioid top K acum. statewise



As the previous accumulated analysis (opioid vs non-opioid), we cannot find any correlation, which might be due to drug regulations across states.

High Inverse correlations of opioid vs opioid Nationwide



Due to the extremely large amounts of data, and the fact that many drugs have zero values for many years which creates false correlations, we decided to represent this with the highest negative correlations. There was a small amount of inverse correlations, about 300, compared to around 30,000 high positive correlations. Even if a large portion of the positive correlations are false, this ratio would still suggest that most opioids are following similar trends.

Summary of correlations

correlations	|corr| avg	+ corr	- corr	|corr|>0.7
opioid acum state <--> nonopioid acum state	0.16	14	0	0.05%
opioid topK acum state <--> nonopioid topK acum state	0.16	0	0	
opioid acum national <--> nonopioid acum national	0.58	0	0	
opioid topK national <--> nonopioid topK national	0.4	16	0	16.00%
opioid topK acum national <--> nonopioid topK acum national	0.045	0	0	
opioid topK state <--> nonopioid topK state	0.32	7	0	7.00%
opioid national <--> opioid national	0.23	94142	149	10.00%

opioid acum state <--> opioid acum state	0.16	0	0	
opioid topK state <--> opioid topK state	0.5	19	0	19.00%

Regression Analysis[14] [15]

We fit a linear regression model to the total reimbursement values of opioids and nonopioids. Opioids is on the left, and nonopioids is on the right. We can see from the P-value that the linear regression models are safe to use. They also both have R^2 values that are greater than .5, which means that over 50% of the data can be explained by the linear regression model, which helps us justify using the model. The linear regression model also supports a constant variance, and the residual plot shows a random distribution which allows us to conclude that the linear regression model shows us that both opioids and nonopioids have had a positive correlation in reimbursement value since 2000.

OLS Regression Results

Dep. Variable:	y	**R-squared:**	0.940
Model:	OLS	**Adj. R-squared:**	0.933
Method:	Least Squares	**F-statistic:**	140.3
Date:	Tue, 27 Nov 2018	**Prob (F-statistic):**	8.60e-07
Time:	08:03:02	**Log-Likelihood:**	-142.24
No. Observations:	11	**AIC:**	288.5
Df Residuals:	9	**BIC:**	289.3
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[95.0% Conf. Int.]
const	-2.505e+08	2.12e+07	-11.822	0.000	-2.98e+08 -2.03e+08
x1	1.247e+05	1.05e+04	11.845	0.000	1.01e+05 1.49e+05

Omnibus:	0.718	**Durbin-Watson:**	0.569
Prob(Omnibus):	0.698	**Jarque-Bera (JB):**	0.571
Skew:	0.465	**Prob(JB):**	0.752
Kurtosis:	2.382	**Cond. No.**	1.28e+06

OLS Regression Results

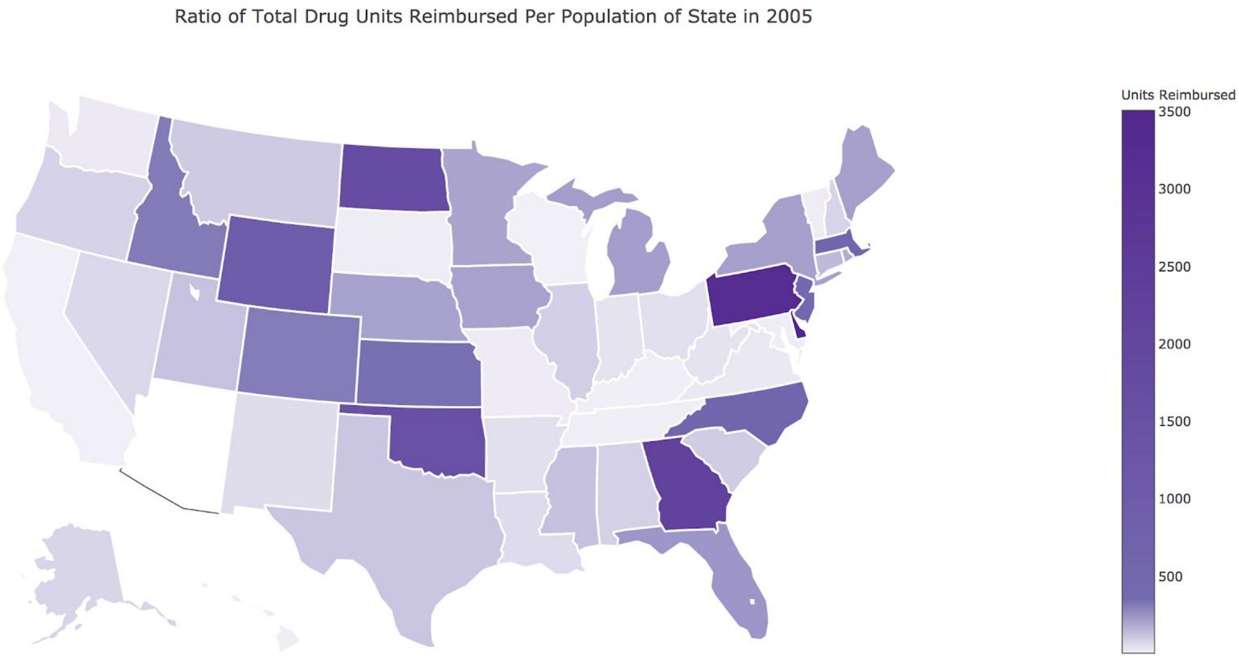
Dep. Variable:	y	**R-squared:**	0.728
Model:	OLS	**Adj. R-squared:**	0.713
Method:	Least Squares	**F-statistic:**	50.76
Date:	Tue, 27 Nov 2018	**Prob (F-statistic):**	8.96e-07
Time:	08:03:02	**Log-Likelihood:**	-286.28
No. Observations:	21	**AIC:**	576.6
Df Residuals:	19	**BIC:**	578.6
Df Model:	1		
Covariance Type:	nonrobust		

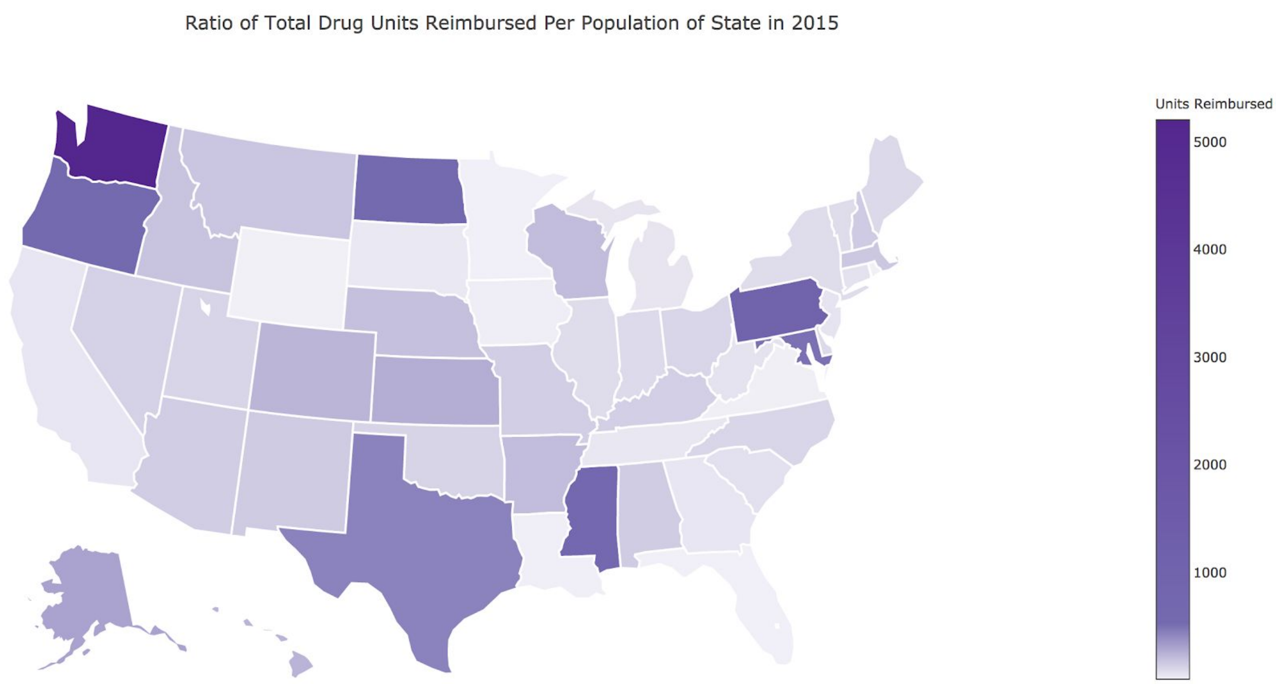
	coef	std err	t	P>|t|	[95.0% Conf. Int.]
const	-1.088e+08	1.53e+07	-7.105	0.000	-1.41e+08 -7.68e+07
x1	5.438e+04	7633.109	7.124	0.000	3.84e+04 7.04e+04

Omnibus:	0.435	**Durbin-Watson:**	0.329
Prob(Omnibus):	0.805	**Jarque-Bera (JB):**	0.553
Skew:	0.136	**Prob(JB):**	0.758
Kurtosis:	2.253	**Cond. No.**	6.65e+05

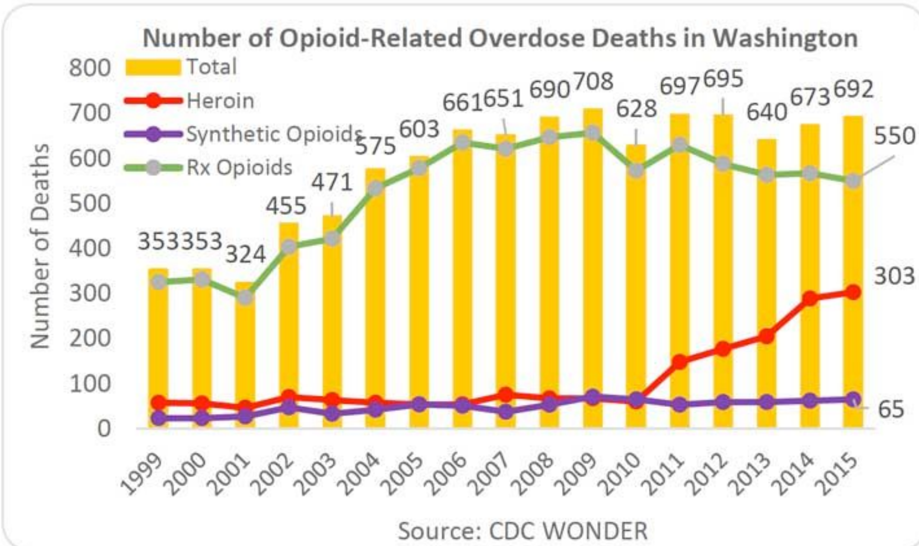
[14] "Simple and Multiple Linear Regression in ...." 8 may.. 2017, https://towardsdatascience.com/simple-and-multiple-linear-regression-in-python-c928425168f9. Se consultó el 28 nov.. 2018.
[15] "How to Calculate Prevalence Rates Per Thousand | Sciencing." 13 mar.. 2018, https://sciencing.com/calculate-prevalence-rates-per-thousand-7533277.html. Se consultó el 28 nov.. 2018.

Statewide Analysis:





We can see that some states had significantly different values in Drug Units Reimbursed Per Population in 2015 than in 2005. It's hard to say whether these are actual trends or just different values due to variation in factors such as population. Though we cannot make any statistical conclusions, we can compare data to other factors to find interesting possible trends that can be explored in future studies. One such example is when comparing the state of Washington to prescription overdoses, we can see a very obvious similarity to the increase in prescription reimbursement seen on the state choropleth map.



"Washington Overdose Summary ...." 2 Feb.. 2018, https://www.drugabuse.gov/drugs-abuse/opioids/opioid-summaries-by-state/washington-opioid-summary. Se consultó el 28 nov.. 2018.

II. Evaluation of Approach

Different types of normalization could have been used, for example once the top K were selected normalize so that the total accumulation off all drugs and all years was 100%.

We performed the Top K program using the Python Pandas Library in less than a quarter of code and much quicker. The dataset we used was split conveniently by year. Each yearly dataset was small enough that it could be conveniently read in and cleaned up using the Pandas library. The modeling of the statewide total drug reimbursement values was actually done by reading in the data through Pandas and performing simple aggregate functions. The Python code to do this is much cleaner and simpler than using Hadoop. This goes to show the lesson that there are plenty of ways of getting around using large-scale frameworks such as Hadoop. In this case we just needed to partition the data by year. Now this was only so convenient because the database was taken from a website that allowed us to choose different filters on the dataset, in which splitting by year was an option.

III. Conclusion

We have found that our hypothesis that opioids are increasing in a dramatic fashion because they are being used as a replacement for other painkillers has been rejected. The data has shown that non-opioid units have seen an increase in the same time period. This indicates that both types of drugs have been heavily rising since 2010. For opioid units there has been small drop in units reimbursed by 2015 and a mild decrease in units reimbursed since an opioid epidemic was declared.

The analysis of this data has caused us to formulate another hypothesis that fits the data better. Our second hypothesis is that there is not an opioid epidemic, but an overall prescribed drug epidemic. We suggest further research to fully accept our second hypothesis.

Project Contributions

A. Andre Hochmuth
   1. Data Modeling/Visualizations
      a) Trend Graphs
      b) Choropleth Map
   2. Statistical Analysis
   3. Python Statewide Top K Program
   4. Final Report Write Up

B. Joaquín Cuomo
   1. Database analysis and background research
   2. Hadoop Pre-Processing/Filtering Data
   3. Hadoop Top K and Normalization Program
   4. Correlational Analysis in Hadoop
   5. Github Wiki Explanations of Programs
   6. Trends and table in Heatmap section
   7. Final Report Write Up
      a) Problem formulation
      b) Algorithms
      c) Correlation

C. Mitchell Rouault
   1. Correlational Analysis Data Modeling
      a) HeatMaps
      b) Inverse Correlation Plot
   2. Final Report Write Up

D. Josh Munoz
   1. GitHub Repository Creation
   2. Final Report Write Up

Bibliography

"Opioid Data Analysis and Resources | Drug Overdose | CDC Injury ...." 9 feb.. 2017, https://www.cdc.gov/drugoverdose/data/analysis.html. Consulted on 27 nov.. 2018.

"Overdose Death Rates - National Institute ...." https://www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rates. Consulted on 27 nov.. 2018.

"The Opioid Crisis | The White House." https://www.whitehouse.gov/opioids/. Consulted on 27 nov.. 2018.

"Quantifying the Epidemic of Prescription Opioid Overdose Deaths ...." 1 abr.. 2018, https://ajph.aphapublications.org/doi/10.2105/AJPH.2017.304265. Consulted on 27 nov.. 2018.

"4: Opiates binding to opiate receptors in the nucleus accumbens ...." https://www.drugabuse.gov/publications/teaching-packets/neurobiology-drug-addiction/section-iii-action-heroin-morphine/4-opiates-binding-to-opiate-rece. Consulted on 27 nov.. 2018.

"WHO | Information sheet on opioid overdose - World Health Organization." https://www.who.int/substance_abuse/information-sheet/en/. Consulted on 27 nov.. 2018.

"Data Resources | Drug Overdose | CDC Injury Center." 19 oct.. 2018, https://www.cdc.gov/drugoverdose/resources/data.html. Consulted on 27 nov.. 2018.

"Other drugs for pain control | Coping with cancer | Cancer Research UK." https://www.cancerresearchuk.org/about-cancer/coping/physically/cancer-and-pain-control/treating-pain/painkillers/types-of-painkillers/other-drugs. Consulted on 27 nov.. 2018.

"Pain Relievers: MedlinePlus." 29 ago.. 2018, https://medlineplus.gov/painrelievers.html. Consulted on 27 nov.. 2018.

"Drug Safety and Availability > FDA Drug Safety Communication: FDA ...." 26 feb.. 2018, https://www.fda.gov/Drugs/DrugSafety/ucm451800.htm. Consulted on 27 nov.. 2018.

"National Drug Code Directory - FDA." 4 feb.. 2016, https://www.accessdata.fda.gov/scripts/cder/ndc/. Consulted on 27 nov.. 2018.

"Simple and Multiple Linear Regression in ...." 8 may.. 2017, https://towardsdatascience.com/simple-and-multiple-linear-regression-in-python-c928425168f9. Se consultó el 28 nov.. 2018.

"How to Calculate Prevalence Rates Per Thousand | Sciencing." 13 mar.. 2018, https://sciencing.com/calculate-prevalence-rates-per-thousand-7533277.html. Se consultó el 28 nov.. 2018.

"Washington Overdose Summary ...." 2 Feb.. 2018, https://www.drugabuse.gov/drugs-abuse/opioids/opioid-summaries-by-state/washington-opioid-summary. Se consultó el 28 nov.. 2018.

"hadoop-map-reduce-patterns/CartesianCommentComparison.java at …." https://github.com/geftimov/hadoop-map-reduce-patterns/blob/master/src/main/java/com/eftimoff/mapreduce/joins/cartesianproduct/CartesianCommentComparison.java. Consulted on 27 nov.. 2018.