1 Objective:

Using Medicare Part D data, the possibility of opioid abuse is explored from outliers detection of healthcare providers claims.

2 Introduction:

Drug related deaths have steadily increased in the few decades. In 2016, approximately 63,693 Americans died from drug-overdosing. Noting that there may be more than one drug involved, of the deaths, 42,249 (66.4Due to their addictive nature, opioids are treated as controlled substances in the United States. Examples of opioids are heroin, oxycodone, hydrocodone, morphine, fentanyl, among others. Heroin is a schedule I drugs, deemed as having no medical use, unsafe to use, and illegal in all cases. Oxycodone, methadone, morphine are examples of schedule II opioids. Scheduling divides controls drugs into five categories used FDA and the DEA in order determine a drugs risk of abusing and level of medical acceptability. Lower number scheduling indicates higher potential of abuse and higher medical regulation. In the last few decades, there has been concern over the abuse of prescription opioids, in particular Oxycontin, an oxycodone hydrochloride salt approved in 1995 by the FDA as an anesthetic for those with moderate to severe pain. At the time Oxycodin was generally thought to be less additive due to the tablet's delayed onset of action. However, according to a 2005 study, it is the most heavily abused of the schedule II/III opioids.

3 Business Case and Value Proposition:

In 2017, 17% of individuals in the US were prescribed one or more opioid [2]. In 2016, it was estimated that 11.5 million or 4.3% of US population over the age of twelve misused illegal or prescription opioids [2]. Not only does the misuse of opioids cost lives but there is also a financial cost. Based on 2013 data, it was estimated that the economic burden for the misuse of opioids was $78.5 billion. The health care costs to private insurance companies was $14.0 thousand per patient.% It is therefore in the interest of these companies to track the misuse to prescription opioids.

4 Data Set Overview

The opioid data for this project comes from a 2013 Medicare Part D data set of the opioid prescription of health providers around the United States. This set have the total opioids and total number of drugs prescribe, the zip code, and

state of each health provider. It is important to note that since the data is of the prescriptions for patients on Medicare Part D, the data will reflect an older population. Individuals on Medicare Part D are 65 or older and it is expected they may require painkillers at a higher rate than the general public. From this data set, both specialty, and location can be correlated with the amount of opioids prescribed. Using data from the 2010 census of the population, the population of each zip code can be determined, and correlated with the levels of opioid prescription.

5 Data Acquisition

The data set comes from a 2013 Medicare Part D report on the opioid prescription of medical professionals. The data reports the health care provider's name, state, the zip code, the specialty, the number of opioid prescribed (new prescriptions and refills), the total of prescription, and the percentage of drugs prescribed that are opioid. This data set contains over 1 million different health care providers. It was downloaded using the Socrata module and the query limit set to 1.5 million so all data points included.

6 Data Wrangling and Cleaning

The first thing I did was change the name of the columns to something more descriptive. Each health care provider was given a unique identification number. This number was used as the index for the pandas dataframe. The name catalog was divided into two columns, a first and last name. In the cases where health care provider name listed were the health center rather than the name of the individual, the first name was empty. The first and last names were combined into a single column in order to remove empty columns. At several entries, the number and percentages of opioids prescribed were empty. These were assumed to be and filled with zeros. Two versions of this dataframe were created, one that contained health care providers that prescribed opioids and one that included all health care providers. The complete dataframe contains 1049326 health care providers, the dataframe that only includes health care providers that prescribed opioids has 496744 providers. For each unique specialty description, an integer identification number was given. These numbers will be used in order to categorize the health care providers. Note that several providers with similar specialty were given different number. For example, there are several different identification numbers for nurse. Several of these categories will most likely be combined in the future based on keywords (such as nurses and sur-

geons). There are currently 246 specialties found, only 169 of which prescribe opioids.The population of each zip code was reported in the 2010 census. The population of zip codes in the same city were summed over and then assigned to the health care provider using their zip code. After this, of the 1049326 health providers, 1836 (0.17%) had unknown area population. Some operated outside the US and therefore not included in the census. Some health care providers were in the zip codes missing from the census.

7 Outlier Detection:

Since detecting outliers is the purpose of this analysis, for the purpose of cleaning the data, outlier detection was limited to filtering the data points outside of the feasible range. The only concern was if the percentages of opioids prescriptions were more than 100% or below 0%. The maximum percentage of opioid prescription was 100% and the minimum was 0%.

8 Statistics

8.1 Opioid claims of a city population:

From Figure 1, we can see that the number of opioid claims increases with city population. This is what is expected; as the size of the population increases, the number of people who would be prescribed opioids also increases. But when we switch from total opioid claims to percentage of claims that are for opioids, the trends change. For smaller cities with lower population ranges, there is a huge range of percentage opioid claims, with values reaching as high as 100%. On closer inspection (see the scatterplot of populations ¡80,000) it is clear that an inverse relation between population size and percentage opioid claims exists. The maximum opioid prescribed percentages settles in the range 5-10%. As the number of claims increase with population, any health care provider who prescribed high levels of opioids is unlikely to have a large influence on the overall percentages; this most likely the reason why this inverse relation between population and opioid prescribed percentages occurs.

8.2 Opioids prescribed by health providers:

Figure 2 shows two histograms; the left presents the number of opioid claims by prescribed by healthcare providers and the right presents percentage of drug claims that are opioids prescribed by healthcare providers. As shown, the majority of healthcare providers each provided ¡25 opioid prescriptions for their

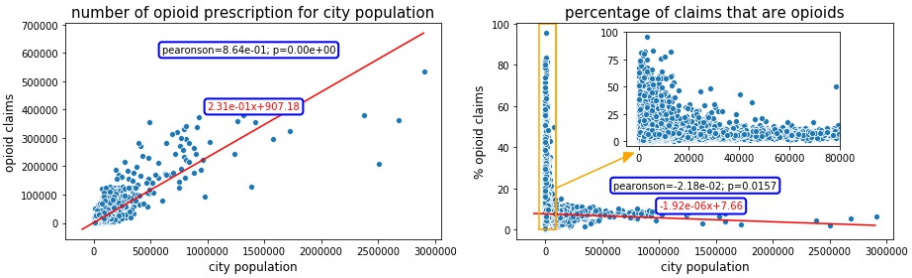

Figure 1: dfsaf

patients in 2013; which consisted of about 0%-15% of drug claims.

8.3 Opioid prescribed by clinicians:

Opioid prescribed by clinicians Figure 4 examines the amount of opioid prescriptions of the ten top prescribers by specialty. Medicare Part D patients are prescribed opioids more frequently by family practice and internal medicine physicians than any specialist, 21.6 million and 18.9 million in total, respectively. The third most frequent prescribers of opioids, nurse practitioners, are responsible for far fewer opioid claims, 6.4 million. Interesting, only about 5.5% and 4.6% of the total number of prescriptions by family practice and internal medicine physicians respectively were opioids. Family practice and internal medicine physicians treat a broad range of illnesses and disorders and are generally the first clinicians that patients consult for any complaint. This can explain why these two groups are responsible for far more opioid claims but opioids only include a small portion of their prescriptions. Orthopedic surgery, physical medicine and rehabilitation, anesthesiology, and interventional pain management are the most frequent prescribers of opioids; with more than 40% of their prescriptions falling into that category. This would be expected because two treat patients in or recovering from pain (physical medicine and rehabilitation and interventional pain management) and two perform surgeries (orthopedic surgery and anesthesiology) for which opioids are commonly given for post-operative pain.

8.4 Range of opioid prescription by clinicians

Figure 5 examines the range of percentages of opioid prescribed by the specialists in Figure 4. While family practice, internal medicine, and general practice physicians have narrow spreads of their percentages (the whiskers in their boxes range from 0% to 14%), they have outliers that reach 100%. However, their outliers are only about 6% of the total number of healthcare providers in those

categories. Healthcare providers are considered outliers of their specialty if their percentage of their opioid claims are outside the first and third quartile by 1.5 the interquartile ranges (difference of the first to third quartile). Physical medicine and rehabilitation's boxplot has a whiskers span of 0% to 100% while orthopedic surgery, physician assistant, and anesthesiology's whiskers span most of this range. All of the top ten prescribers of opioids, as outliers or in the normal range of the plot, span 0 to 100%. Despite the wide range of values, only general practice had more than 10% of healthcare providers as outliers, at 11%. Nurse practitioner were slightly below at 9%. While it makes sense that specialties that deal with pain relief or surgery would have a wide range, that family practice and internal medicine healthcare providers have a significant number of outliers suggests that they are too lenient with opioid prescriptions

8.5 Correlation Between Population and Number Opioid Claims by Healthcare Provider

In Figure 1, the correlation with number of opioid claims and percentage opioid claims of a healthcare provider to city population was examined. Both were weakly correlated with population but were statistically significant (p ¡ 0.05). Surprisingly the total number of claims negatively correlated. Because the correlation is weak, the ability of population to detect outliers opioid claims by a healthcare provider is suspect.

8.6 Correlation Between Population and Number Opioid Claims by Healthcare Provider

In Figure 2, the correlation between lowest month average temperature of 2013 with both the the number opioid claim and the precent opioid claim of the healthcare provider is examined. Lowest month average temperature is weakly positively correlated with number of opioid claims and weakly negative with percent opioid claims. Both correlations were statistically significant (p ¡ 0.05) but similar to population because it is weak, the predictability of temperature in detecting opioid claim outliers is suspect. Similar plot using mean and highest average temperature of 2013 were also produced with similar results.