Opioid Crisis Funding

InsightfulRX Analytics

Dec 6, 2023

Analytics Team

Ed Hui

Ed brings with him 25 years as an actuary. Ed will be the primary Subject Matter Expert on the opioid crisis, lending his expertise on the data while developing the data pipelines, models, and visualizations.

Curt Robords

Curt brings 15 years of experience as a data analyst and data engineer to our team. Curt will be working on the data architecture, models, and the final dashboard and mobile application.

Jared Sarabia

Jared has worked in the public sector developing data cleaning and data analytic tools. Jared will be involved in developing the data pipelines and architecture, models, dashboard/mobile app development, and general project management tasking.

William Tsai

William has worked in the technology startup industry for a number of years. Within this project, William will be involved in guiding strategic initiatives as the team builds out data pipelines / modeling mechanisms.

Dan Yara

Dan brings with him over 10 years of experience as a data analyst at a public fortune 100 firm. Within the project Dan will be responsible for assisting with the architecture and structure of the models.

Agenda

1. Executive Summary: Business Case & Objectives
2. Background on Data Sources
3. Model Development
4. Key Findings
5. Case Study
6. Dashboard Product Overview
7. Go-Forward Recommendations

Executive Summary

Why We Are Here

Summary Brief:

InsightfulRx Analytics has developed a greenfield predictive model for concentrating opioid fund disbursement to augment the U.S. Government's efforts in combating the present national health crisis of opioid proliferation.

Our work establishes the firm's position as a data-centric thought leader and key consulting partner to the federal government looking to increase their data analysis abilities and optimize their legacy data pipelines and data collection methodologies.

Over $50bn in Opioid Legal Settlement Funds will be Disbursed in 2023 to Combat Opioid Epidemic

Current Situation

Overview:

- Proliferation of prescription Opioids within the United States stemmed from FDA's approval of a highly addictive prescription drug, OxyCotin.
- Since then, Opioid consumption has skyrocketed, spreading across State lines.

Key Observations:

- Synthetic Opioids are a subset of the Opioid drug class, legally created for prescription use. The most commonly known synthetic opioids are Fentanyl & Morphine.
- Within this project, we dedicate our modeling resources to Synthetic Opioid, specifically Fentanyl, as the drug increase has been dramatic, responsible for over 70% of the 103K U.S. drug deaths in 2021.

PUBLIC HEALTH

Here's who controls the $50 billion opioid settlement funds in each state

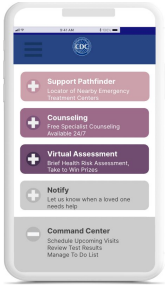
As more than $50 billion makes its way to state and local governments to compensate for the opioid epidemic, people with high hopes for the money are already fighting over a little-known bureaucratic arm of the process: state councils that wield immense power over how the cash is spent.

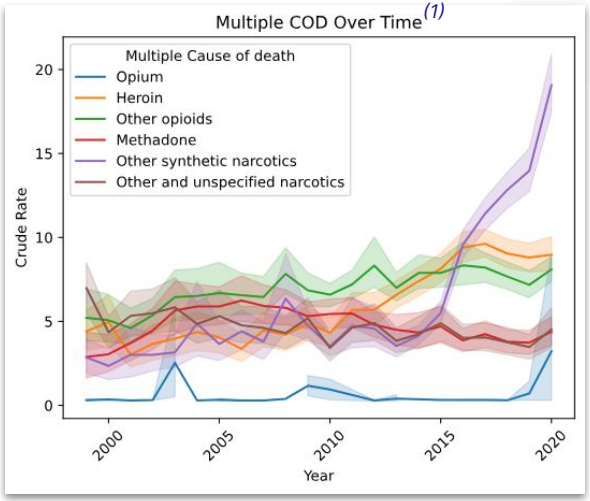
Economist, 2020, "The wider effects of America's opioid epidemic"

Business Case

Opportunity

- Dedicated federal funding to fund Opioid treatment centers & other mitigation efforts are inefficient and misallocated.
- Poor integration of data to drive current funding strategy, causing high risk U.S. States & Counties and demographic groups to miss out on invaluable resources.





Our Solution

- Develop data models based on existing Federal Healthcare & Economic Data and provide data-driven recommendations to the U.S. Government & commercial partners
- Our modeling methodologies come with integrated web dashboard & mobile applications to support federal fund management stakeholders and

InsightfulRx Analytics

(1) CoD refers to Cause of Death or Mortality Rate, sourced from U.S. Centers of Disease Control (CDC) data. This will be expanded on further in this deck.

Our Objectives

- Evaluate Key Variables that drive mortality rate in current Opioid Epidemic
- Derive Recommendations to optimize U.S. Health Strategy against Opioid Epidemic
- Design & Implement Scalable Data Architecture & Pipeline

01	**Recommend solutions to mitigate the use and impact of Opioids throughout the US**	• Use analytics to build predictive models • Use factors such as socioeconomic, demographics, existing funding • Decrease Opioid related deaths and addictions
02	**Assist in future allocation of Federal Funds to US demographics in need**	• Use analytics to recommend new funding plans • Be sure to see that counties in need, get the help they require

Project Development Timeline

Completed Oct 4

Project Initial Scoping
Identify business case, consolidating datasets, scoping deliverables

Completed Oct 23

Deliverables Design
Create web & mobile dashboard mockup and outline visualizations

Completed Nov 1

Stakeholder Status Check
Align with C-Suite stakeholders on preliminary data insights, deliverable designs, and resource allocation

CURRENT Nov 21

Results Showcase
Submission of Final Report, Business Recommendations & Dashboard Walkthrough to Stakeholders

Data Analysis & EDA
Evaluate dataset and pre-processing data to confirm modeling methodology

Completed Oct 17

Model Validation
Test model performance and refine to optimize actionable insights

Completed Oct 28

Develop Dashboards
Deliver beta version of Dashboard for stakeholder review

Completed Nov 12

Further Consultation
With CEO Approval, our Project Team to embed into Eng teams to provide database/dashboard training & support

CURRENT Dec 6, 2023

Data & Model Deep Dive

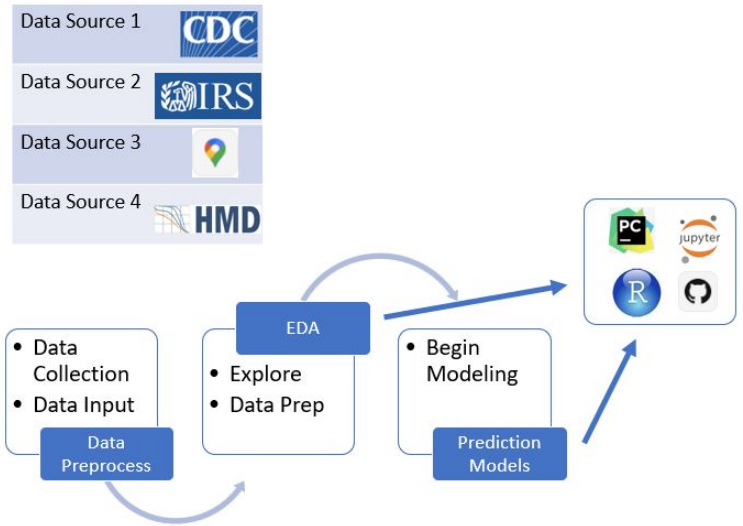
Data Sources and Background

1. Centers for Disease and Prevention (CDC) WONDER (Wide-Ranging Online Data for Epidemiological Research)
   a. Public free dataset
   b. 75 different variables such as race, gender, age

2. Internal Revenue Service (IRS) 990 Tax Forms (Charitable donations)
   a. Over 4,000 charitable donations from the Bill and Melinda Gates foundation

3. Society of Actuaries FIP Codes
   a. Socioeconomic stratified county IDs

4. Human Mortality Database
   a. Worldwide mortality rates by countries

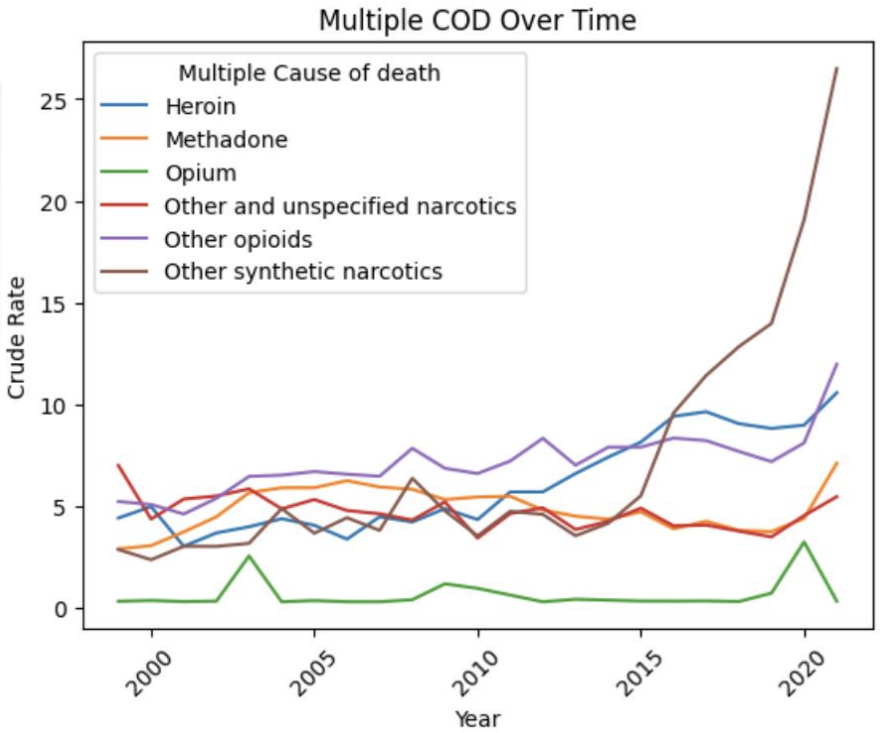


Key Trend: Opioid Deaths

Our Exploratory Data Analysis yielded interesting results:

- Rapid increase in Opioid related deaths in recent years
- Fentanyl related deaths categorized as " Other Synthetic Narcotics"
- Demonstrates the complexity of the crisis as it is not slowing down

Fentanyl Deaths have increased 25x in the past 10 years



General Modeling Objectives

1 Data Smoothing

Provide recommendations of funding allocations conducted on equitable, reasonable, & optimized basis

Our Approach

Pursue Data Smoothing within Model to reduce extreme outliers and retain predictive model efficacy to support funding allocation suggestions

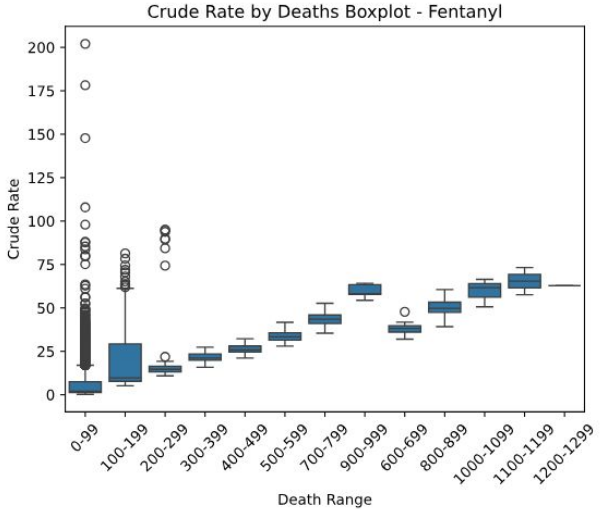
2 Feature Attribution

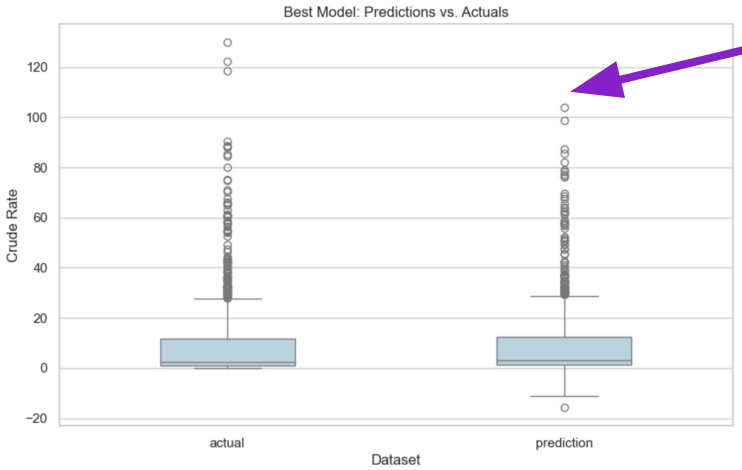
Identifying which input features (i.e. race, gender, age) impact a model & related recommendations for future funding

Our Approach

Evaluate true relative risk relationships for identified variables, deploying statistical analysis to determine predictability

Model Objectives: Data Smoothing





Fewer Extreme Outliers

Key Details

- Reduces extreme outliers, reducing data volatility & bias (i.e. large/small allocation to counties)
- Makes the model more explainable and allows for further analysis to be conducted
- Establishes accountability for federal stakeholders to justify specific funding allocation amounts and strategy for specific counties by attributing back to this data model & related outputs

Model Objectives: Feature Attribution

Factors	Coef Importance
Gender	High
Race	High
Age Group	Medium
Socioeconomic Group	Medium
Population	Low

Key Details

- Evaluating true relative risk relationships for identified variables, using crude fentanyl death as our baseline while deploying various statistical measures to evaluate fit & predictability.
- As highlighted, various factors lead to higher consistency of fentanyl-caused mortality occurring.

Modeling Development & Progression

1

- Data Collection
- Clean Data
- Exploratory Data Analysis

2

- Define baseline prediction
- Build regression
- Interpret Regression
- Apply business context to model

3

- Build additional models
- Test models
- Fine tune parameters
- Settle on best model
- Extract feature importance
- Compare to baseline
- Finalize best model
- Publish best and most fine tuned model for production

Model Compare & Contrast

Regression Model

- Identify key variables
- Rule out unimportant variables
- Ability to compare features

Tree Based Model

- Simple layout of specific variables that are important
- Easily explain and trace through the results of the model
- Strong results with small amount of data
- Best metrics such as RMSE

Acknowledgement of Modeling Limitations

- Only have data that is publicly available
- External factors that are not measured
- Data pull limits on public and government websites

Findings & Recommendations

Key Findings

Model Outputs identified specific Demographic Factors that impact Opioid-related Mortality Rate

High Impact Factors

- Males are Most Affected Gender
- Native American & African American Ethnicity More At Risk
- Age is a Mild Factor in Opioid-related Cause of Death

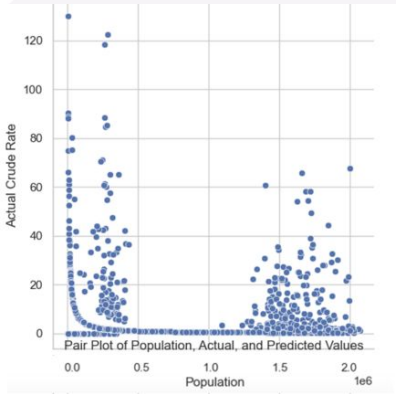
Low Impact Factors

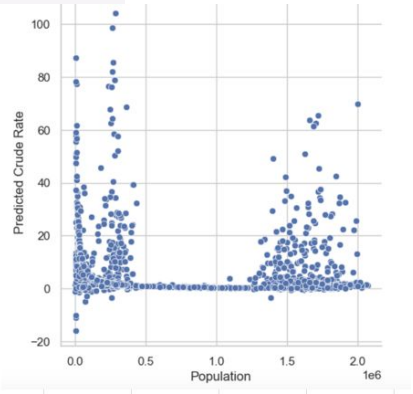
- Socioeconomic Quartiles are Largely Uncorrelated
- County Population Size is not Critical

Case Study: LA County

Thought Process and Goals

- Use best performing model on LA County
  - XGBoost model
- Apply model to LA County
  - Validation that the model works
- Government can use to make predictions on counties with no data
- Self-sustaining model that can be used for future years
- Model to be deployed on other counties as team continues to iterate scalability

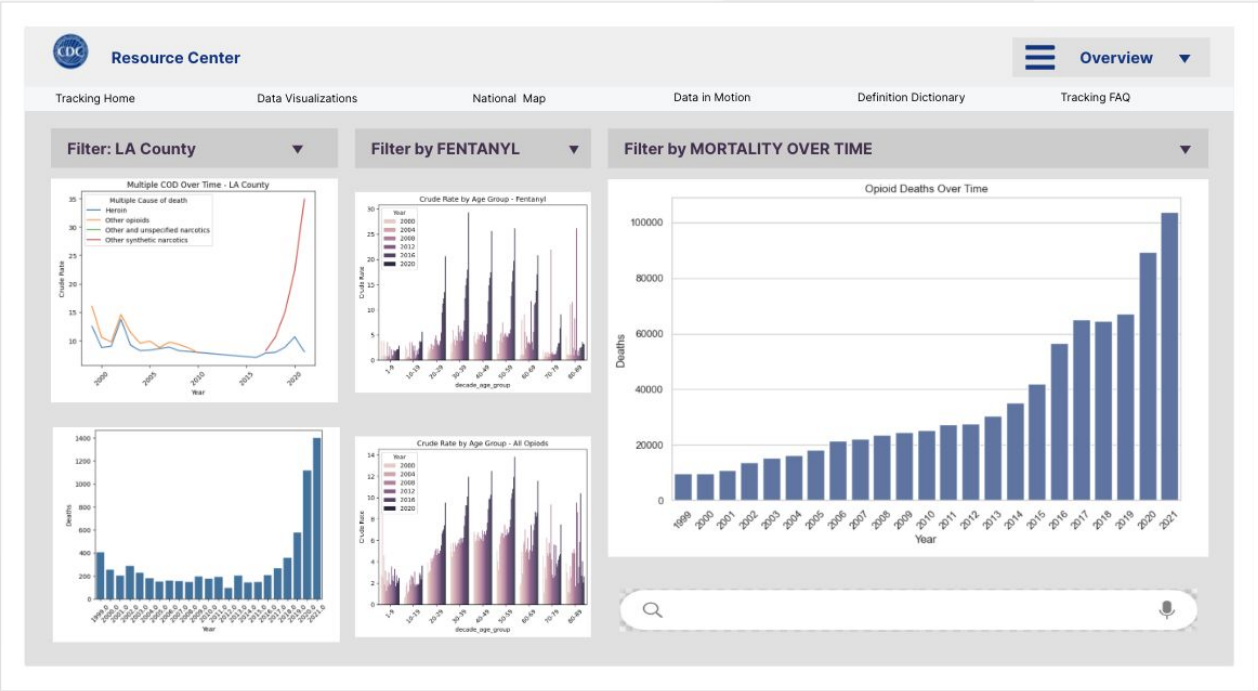




Results

- RMSE was inline with our base model (16.4 v 14.4)
  - Shows traction and ability to have impactful result

Deliverables: Web Dashboard

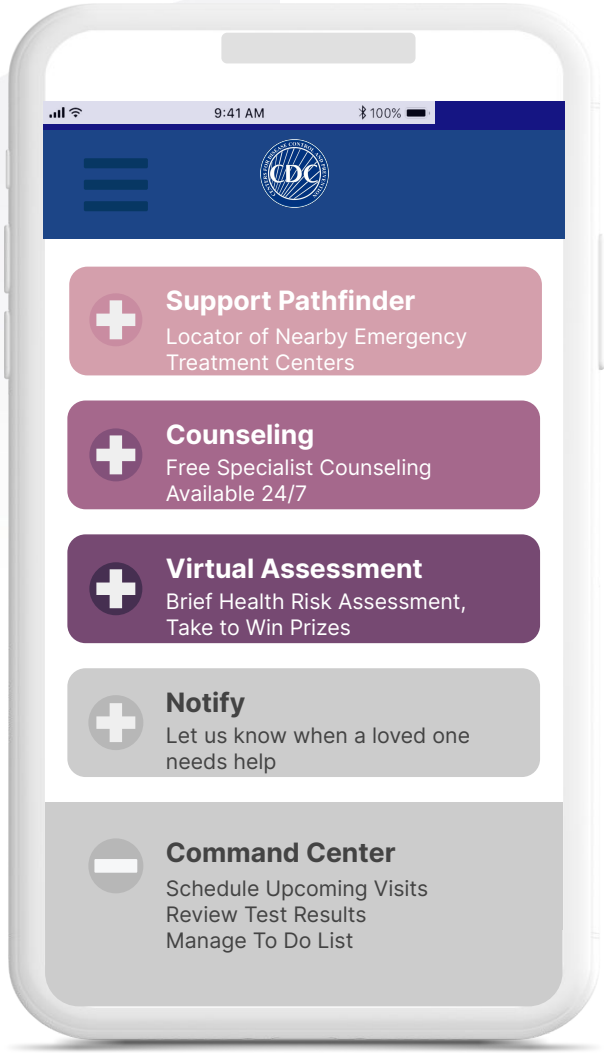


Dashboard Feature Overview

- **Audience:** Federal Government
- **Objective:** Consolidate U.S. County data from federal resources (i.e. CDC, IRS) to streamline data management & decision-making
- **Use Case:** Help stakeholder teams evaluate resource allocation efficacy and support distribution of new funding resources into high-risk U.S. Counties
- **Additional Features:** Data Filtering, Data alerts

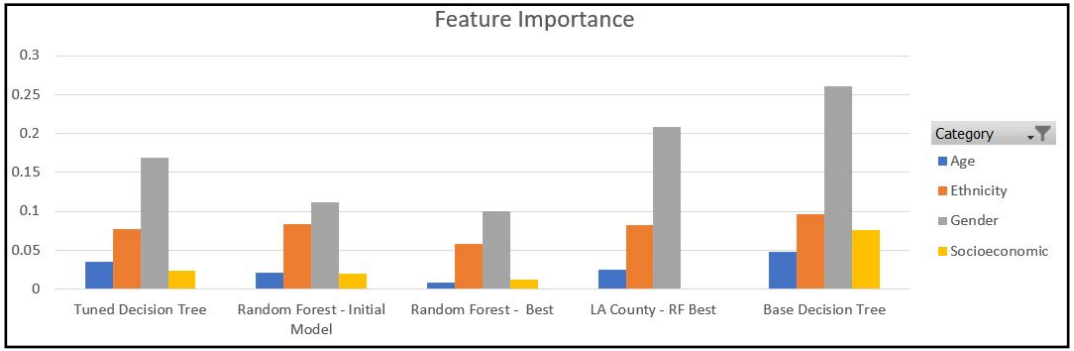
Deliverables: Mobile App

- **Audience:** General U.S Population
- **Objectives:**
  - Streamlining support resources on the county-level for at-risk citizenry while collecting live data for federal resources & databases
- **Key Features:** Live geo-mapping of nearby treatment centers, virtual health assessments & counseling, mobile data alerts, and additional links to self-help resources.
- **Use Cases:**
  - Inform users of impact of Opioids in their community and prevalent risk factors per individual.
  - Help locate nearby treatment centers & health guidelines.



Go-Forward Recommendations

- **Next Steps:**
  - Strategy and Model Scope
    - Expand data set to include more metrics such as crime and education
    - Drive central data warehousing to be used as inputs to the models
    - Create partnerships with drug stores such as CVS and Duane Reade
      - Obtain additional data
- **Dashboard Areas of Growth:**
  - **Web:** Focus on attributing real-time data (up to 2023 data) to guide federal stakeholders on fund disbursement as legal settlements continue
  - **Mobile:** Maintaining a high degree of interactive UI to receive user feedback and integrate Tableau Mobile to integrate web to mobile app pipeline



Archive

Citations

Drug Enforcement Administration. "Facts About Fentanyl." Accessed January 10, 2023. https://www.dea.gov/resources/facts-about-fentanyl.

Internal Revenue Service. “Form 990 Series Downloads.” Accessed October 1, 2023. https://www.irs.gov/charities-non-profits/form-990-series-downloads.

The Economist. 2020.. “The Wider Effects of America’s Opioid Epidemic.” January 16, 2020. https://www.economist.com/business/2020/01/16/the-wider-effects-of-americas-opioid-epidemic

The New York Times. "Fentanyl’s Toll: One Family’s Murder." Podcast audio, September 22, 2023. Accessed October 29, 2023. https://www.nytimes.com/2023/09/22/podcasts/the-daily/fentanyl-murder.html?showTranscript=1.

Data Source SWOT Analysis

Data Source	Strength	Limitations
Socioeconomic and CDC WONDER Dataset	Low level granular detail such as gender, race, age ranges, & socioeconomic data	Records can be excluded when death counts are too low based on aggregation unit
	Specific causes of death maintained by health agency	Data throttle limit of 6mm rows per query
IRS 990 forms data description (charities)	Wide range of data, given legal audit requirements for charitable orgs	County, Zip Code, and FIPS Connections can be inconsistent or out-of-date
	Disbursed funds are denoted down to the grantee level, with available relevant addresses	2020 (Covid) caused a backlog in missing tax records starting that year
	Readable PDFs accompany machine-readable XML files	Grantee usage of funds from charities are not listed, reducing clarity on usage towards opioid mitigation.
Society of Actuaries FIP Codes	County FIP ID Codes	
Human Mortality Database	Worldwide mortality rates by country from 1950s	Only Age, Gender, Year covariates available