P8105 Final Project Proposal

Analyses in Drug Overdose and Death Counts Across States

Team members:

Jinghan Liu (jl6048); Paula Wu (pw2551); Yuxuan Chen (yc4018); Yuan Meng (ym2866); Yijia Chen (yc3883)

Motivation:

"Drug Overdose Deaths Remain High" - the headline posted by CDC rings the bell. Ever since 1999, nearly 841,000 people have died from drug overdoses. 70,630 drug overdose deaths occurred in the United States, in 2019 alone [1,2]. Deaths resulting from drug overdose increase drastically in every state, especially in those who recently passed bills to legalize/decriminalized recreational cannabis or marijuana in recent years. The abuse of and addiction to drugs is a genuine national crisis that influences public health as well as social and financial welfare. Though CDC has posted the US drug overdose death rate and rate change maps from 2014-2019 along with the rate change graphs from 2014-2017, there hasn't been any update for the past 2-3 years. Thus, our group wants to start from here and do more investigation upon the drug overdose death rate across states in the US. Fortunately, there are available drug-overdose-related data online to propel our motivation even further.

Intended Final Product

Our final products will include an organized report, a web page, a screencast, and corresponding Github repo. The report will fully describe our process and results that cover data collection and cleaning, EDA (e.g. heatmap), data mining, modeling, and result discussions. The webpage and screencast can give an overview of the project scope, data, approaches, visualizations, and other results. The content in our final product aims to help identify factors associated with drug overdose mortality rate. Based on identified factors, we could identify groups of people that are more vulnerable when facing drug overdose.We will include heat maps to indicate rates or changes over the past few years.

Anticipated Data Source

Main dataset: Provisional counts for drug overdose deaths within the 50 and the District of Columbia 2015-2021: https://data.cdc.gov/NCHS/VSRR-Provisional-Drug-Overdose-Death-Counts/xkb8-kh2a Other helpful sources: https://worldpopulationreview.com/state-rankings/crime-rate-by-state; https://www.cdc.gov/nchs/nvss/vsrr/drug-overdose-data.htm; https://www.drugabuse.gov/drug-topics/opioids/opioid-summaries-by-state; https://www.statista.com/

Planned Analyses

How different types of drugs affect the death rate? Drug usage in different states? What identities make people more likely to be trapped into drug overdose in the U.S? We will probably use time-series analysis and wants to find the most significant factors that contribute to a higher drug overdose rate / death rate over times.

Visualizations

Created a map that allows users to view the fatality rate of overdose in various states of the US. Mortality (rate) caused by each type of drug Types of drug overdose by region Fatality rate in different years with different types of drug Cause of death across different drug

Coding challenges

Multi-dimensional data processing Data tidying Data mining Combination of different datasets

Planned Timeline