Springboard Data Science Career Track Plus: Capstone 02

"Mortality in Guard and Inmate Populations in US Correctional Facilities: a multifactor analysis of differential effects"

09.10.2021
—

Courtney Jackson

Springboard

Suite 1100, 22 Battery Street,

San Francisco, CA 94111

Overview

The data set for analysis is derived from a New York Times repository of data on COVID-19 cases in the United States Correctional System from 2020. Two additional datasets will be used for additional analysis, both from the Centers for Medicaid and Medicare Services (cms.gov). They include the MDCR PROVIDERS 4 report (Medicare Providers: Number of Medicare Certified Hospitals and Skilled Nursing Facilities, and Number of Beds, by State, Territories, Possessions and Other, Calendar Year 2019) and the MDCR PROVIDERS 5 report (Medicare Providers: Number of Medicare Certified Providers, by Type of Provider, by State, Territories, Possessions, and Other Areas, Calendar Year 2019). I will also include COVID tracking data from the COVID Tracking Project for data on different counties and states.

Goals

1. Analyze the rates of covid amongst inmates and guards in US correctional facilities and the relationship of these rates to one another. I will also look into the rates in the surrounding counties and states.

2. Analyze the effect that availability of medicare or medicaid availability may have on the populations local to but outside of these facilities. The purpose would be to discover if availability of publicly funded healthcare is predictive of higher rates of COVID and COVID-related mortality.

Specifications

I will first describe the summary statistics of covid cases and deaths among guards and inmates of each facility, each county, each state, and each geographical region. I will then use this information to find correlations between these variables and how densely populated the facilities are, and how close to their maximum capacity they are.

COVID-19 infection and death rates of a county, state, or geographical region may be predictive of the cases and death of the inmates and guards. I will also seek to find if there is a differential effect of this prediction on inmates as compared to guards, who live in these areas and interact with people outside of the prison. It may be that prison guard cases are predictive of inmate cases more so than the inverse, because guards may bring the disease inside the facility due to their contact with the surrounding communities.

Geographical regions will be established as the Northeast, Southwest, West, Southeast, and Midwest.

Milestones

I.   EDA

Exploratory Data Analysis will show whether there is a relationship between inmate cases and guard cases. Exploratory Data Analysis will also show whether there is a relationship between inmate and guard cases and the county and state their facilities are located in. It is possible that these relationships will exist in some facilities and localities and not in others. That will aid the determination of what should be included in the model, and whether some facilities may need to be excluded from the model so that we can examine these effects in facilities where a relationship between guard and inmate cases and between these cases and localities does exist. If the model fits, it could be deployed to better understand why these excluded facilities may not demonstrate the relationships that others do.

II.   Modelling

The modelling would first be conducted on the data from the New York Times to better understand predictive factors like maximum population capacity, and the ratio of the actual population to maximum capacity. I would then build a model to understand how inmate and guard cases may be predictive of each other, and integrate a prediction of max capacity and population ratios into a study of how they may have a differential effect on inmates and guards, respectively. After this I would develop a model to understand how localities and their mortality and case numbers, as well as the availability of Medicare and Medicaid facilities and providers may relate to these other relationships.