**Data Dictionary**

Summary

The data is based on the COVID-19 case counts as reported by the COVID Prison Project (CPP), with additional information about the ICE detention facilities collected from a variety of other sources (see information in “Data Sources” and “Terms” below).

There are two data files in this repository. The one entitled “all\_cases.csv” includes the weekly case counts for all of the facilities included in the study. The one entitled “summary\_data.csv” is summary statistics based on the case counts in “all\_cases.csv” and was used for the main analysis in our publication. Each of the columns in these two documents is explained in the “Terms” section below. Also included in this repository is the R code that was used for determining the presence of outbreaks and statistical analysis.

Please contact Emily Woods at [emilywoods@stanford.edu](mailto:emilywoods@stanford.edu) if you have questions about the materials in this repository.

Data Sources

1. UNC Re-imagining Health and Justice Lab. COVID Prison Project. https://github.com/healthandjustice/covid-prison-project/tree/main/data. Published Aug 13, 2021. Accessed March 2023.
2. National Immigrant Justice Center. ICE Detention Facilities as of November 2017*.* https://immigrantjustice.org/ice-detention-facilities-november-2017. Accessed March 2023.
3. Transactional Records Clearinghouse. Detention Facilities Average Daily Population. Syracuse University, Syracuse, NY. https://trac.syr.edu/immigration/detentionstats/facilities.html. Accessed March 2023.
4. Wang, EA, Brinkley-Rubinstein, L, and Puglisi LB. The Case for Prioritizing COVID-19 Vaccines in Prisons and Jails. *The Appeal*. https://theappeal.org/the-lab/report/covid-19-vaccines-in-prisons-and-jails/. Published Feb 22, 2021. Accessed March 2023.
5. Centers for Disease Control and Prevention. COVID-19 Vaccinations in the United States. COVID Data Tracker. https://covid.cdc.gov/covid-data-tracker. Updated 2023. Accessed March 2023.
6. Woolley, J and Peters, G. Statistics: 2020. The American Presidency Project. UC Santa Barbara. https://www.presidency.ucsb.edu/statistics/elections/2020. Updated 2021. Accessed March 2023.

Terms

In all\_cases.csv

**date**: date of recorded case count in YYYY-MM-DD format

**facility**: name of ICE detention center

**state**: state in which facility is located (from the CPP data set1)

**operator\_class**: whether a facility is publicly (“Public”) or privately (“Private”) operated (from National Immigrant Justice Center2)

**type\_class**: whether a facility is dedicated to immigrants only (“dedicated”) or includes immigrants mixed with other incarcerated populations (“mixed”) (from National Immigrant Justice Center2)

**pop\_est**: estimated facility population on that date (fromTransactional Records Clearinghouse3)

**cummulative\_cases**: number of COVID-19 cases reported at that facility, counts are cumulative since the start of data reporting (calculated for one-week intervals based on data reported in CPP data set1)

**incidence**: number of new COVID-19 cases on that date (calculated as difference in cumulative case number compared to the previous week)

**normalized\_incidence**: incidence for that date divided by the total number of weeks of data available for that facility

**week3\_total**: total incidence over the three-week period ending in the indicated date

**outbreak**: TRUE/FALSE whether the week3\_total was ≥ 10% of the facility population (i.e. met our definition of an outbreak)

**unique\_outbreak**: TRUE if the “outbreak” column is TRUE and it is the first time it has switched to TRUE after being FALSE (i.e. only counts the first TRUE in a series of TRUEs)

**vax\_timing**: whether the state in which the facility is located had policies in place to vaccinate incarcerated populations in the first round (aka phase 1) of COVID-19 vaccinations (“early”) or later phases (“late”) (based on information reported by Wang, et al.4)

**community\_vax\_rate**: whether the state in which the facility is located had at least 60% of the adult population vaccinated against COVID-19 by the end of the study period (“high”) or less than 60% (“low”) (based on statistics from CDC5)

**political\_leaning**: political leaning of the state in which a facility is located based on which candidate won the 2020 Presidential election (“R” = Republican, i.e. state voted for Trump; “D” = Democrat, i.e. state voted for Biden) (based on election results6)

In summary\_data.csv

**facility**: name of ICE detention center

**start\_date**: first date in the COVID Prison Project1 data set when any data was reported for a facility

**end\_date**: last date in the COVID Prison Project1 data set when any data was reported for a facility

**weeks\_of\_data**: difference between “start\_date” and “end\_date” in weeks, indicates the total number of weeks during which any data was reported for a facility

**operator:** whether a facility is publicly (“Public”) or privately (“Private”) operated (from National Immigrant Justice Center2)

**type:** whether a facility is dedicated to immigrants only (“dedicated”) or includes immigrants mixed with other incarcerated populations (“mixed”) (from National Immigrant Justice Center2)

**state:** state in which facility is located (from the CPP data set1)

**avg\_pop**: mean population over the course of the study period (calculated from “pop\_est” in “all\_cases.csv”)

**max\_pop**: maximum population size over the course of the study period (determined from “pop\_est” in “all\_cases.csv”)

**min\_pop**: minimum population size over the course of the study period (determined from “pop\_est” in “all\_cases.csv”)

**pop\_change\_percent**: ((population at the start of the study period) - (population at the end of the study period))/(population at the start of the study period) \* 100

**vax\_timing**: whether the state in which the facility is located had policies in place to vaccinate incarcerated populations in the first round (aka phase 1) of COVID-19 vaccinations (“early”) or later phases (“late”) (based on information reported by Wang, et al.4)

**community\_vax\_rate:** whether the state in which the facility is located had at least 60% of the adult population vaccinated against COVID-19 by the end of the study period (“high”) or less than 60% (“low”) (based on statistics from CDC5)

**political\_leaning:** political leaning of the state in which a facility is located based on which candidate won the 2020 Presidential election (“R” = Republican, i.e. state voted for Trump; “D” = Democrat, i.e. state voted for Biden) (based on election results6)

**staff\_positive**: total number of staff members reported to have tested positive for COVID-19 during the study period (from CPP data set1)

**any\_outbreak**: TRUE/FALSE whether at least one outbreak was detected during the study period

**unique\_outbreaks**: number of unique outbreaks during the study period (see definition under all\_cases.csv)

**unique\_outbreaks\_perweeksdata**: value in “unique\_outbreaks” column divided by value in “weeks\_of\_data” column

**outbreak\_weeks**: number of weeks that met outbreak criteria (all weeks, not just unique instances)

**outbreak\_weeks\_percent**: (value in “outbreak\_weeks” column) / (value in “weeks\_of\_data” column)

**max\_incidence\_normalized: (**maximum weekly incidence during the study period) / (value in “weeks\_of\_data” column)

**max\_incidence\_percent:** (maximum weekly incidence during the study period) / (population size at the time that the maximum incidence occurred) \* 100