

# Rural jail responses to COVID-19 in Colorado

Albert Sun

4/30/2021

## Motivation

With overcrowding and characteristically unsanitary conditions, jails are uniquely positioned to spread infectious diseases such as COVID-19 in its quarters and to local communities. At the outbreak of the pandemic, jails across the country have made temporary (albeit significant) efforts to decrease their jail populations and slow down the intake of new inmates. Our project is interested in how Colorado’s data collection efforts can understand the state’s response to COVID-19 and what kind of disparities exist in the release of these jail inmates.

This study focuses on Colorado and its jail data collection. According to a study by the Prison Policy Institute, Colorado’s jail population decreased in greater rates than other state jail populations. [1] Our project is interested in how Colorado’s data collection efforts can understand the state’s response to COVID-19, how these statistics can motivate better public responses when facing future disasters, and what kind of disparities exist in the release of these jail inmates.

The Colorado Jail dataset was signed into existence by the bipartisan Colorado State House Bill 19-1297 in mid-2019, and the bill expanded previous and outdated efforts to catalogue Colorado jail data. The dataset is currently the most comprehensive, open-source published, and catalogued state-wide jail database in the United States. If proven useful in understanding the drivers of jail responses, the Colorado Jail Dataset can be a motivating example for other states to adopt a similar, modern system of data collection on their jails.

## Background

There is a strong political and organizational effort to make jail data surrounding COVID-19 transparent. For instance, the H.R.7983 (COVID-19 in Corrections Data Transparency Act) in the U.S. Congressional docket right now seeks to mandate public reports of data on “COVID-19 in Federal, State, and local correctional facilities” to be reported. [2]

Although the H.B. 19-1297 Colorado Jail Database does not include COVID-19 testing information, research using numbers from the HB 19-1297 database have already highlighted the dataset’s utility in understanding COVID-19 responses. Population data from existing HB 19-1297 data have already been used in aggregate counts for national jail population counts by studies done by the ACLU. [3] More data must be collected to fully document existing racial and ethnic disparities surrounding structural inequalities in the legal system. [4]

Other state-wide jail data collection efforts exist, but they are not nearly as comprehensive and public facing as Colorado’s efforts. Colorado’s data is published in Comma-separated values (.csv) format, allowing data researchers and the public to directly understand and run models on its data without needing to scrape it from a website.

This also encourages quick accessibility of the data. For California and Texas’ jail collection, research groups scraped and published the data periodically onto an unofficial data-sharing website called openICPSR [5][6]. Because Colorado’s jail data is immediately published, it saves data researchers time in not needing to scrape, and it allows them to verify their data by receiving it from an official, governmental source. This improves

reproducibility in data analyses, which is important for informed conversations about the statistical data surrounding Colorado jails. This allows the public to analyze the most recent data coming from their jails during a possible disaster like COVID-19 pandemic right in front of them.

Some stakeholders in Colorado and across the country, have attempted to reduce jail populations at the beginning of the outbreak. Colorado Governor Jared Polis signed an executive order relaxing the standards for early release in March in fear of the effects of excessive overcrowding in jails [7]. Jails who initially thought of expanding have halted their plans due to general decreases in jail populations [8]. Overall, Colorado has reported a net decrease in jail population during the COVID-19 crisis.

## Research Question

What factors are correlated with a county jail decreasing its jail population at the outbreak of and during the COVID-19 pandemic?

## Data Comments

Colorado has a uniform system for collecting jail data in the state, known as the HB19-1297 Jail Data Collection. [9] This dashboard is run by the Colorado Department of Corrections within the Division of Criminal Justice. HB19-1297 was a bipartisan bill primarily sponsored by State Representatives Michael Weissman (D) and Hugh McKean (R). [10][11] HB19-1297 was signed May of 2019. HB19-1297 expanded the information that jail keepers must submit to the Division of Criminal Justice, mandating that unless data is submitted quarterly, jail keepers could be subject to a \$500 fine. The public and published HB 19-1297 dataset represents 3 collection periods in 2020, split up into yearly quarters. Q1 was in January, Q2 was in April, Q3 was in July. A more-frequent periodic collection of data lends the dataset to good time series analysis, because we can better see the effects of policies or huge events, like COVID-19 over time.

The HB 19-1297 data collection has the following variables: population count, gender, race, Hispanic, capacity, beds, deaths, bookings, releases. It splits up each of these counts based on different metrics, specifically: total number of inmates, sentenced, unsentenced and hold, unsentenced and no hold, unsentenced and no hold felonies, unsentenced and no hold misdemeanors, municipal charge, administrative segregation, competency evaluation, average daily population, average LOS felonies, Felony Releases, Average LOS Misdemeanors, Misdemeanor Releases, and Homeless. The data is mostly complete. Only 1 county jail out of all 53 county jails in Colorado didn't submit data for the last quarter. Given that the HB 19-1297 jail dataset collects a lot of data surrounding race and different measures, only 19.649% of the data is missing. However, overall, all jails have submitted overall population counts and race counts. Finally, we seek to impute demographic variables in a particular region (general population, political profile, etc.) from the MIT Election Lab Dataset, allowing us to investigate different regional factors in the likelihood of releases in jails. [13]

## Data Preparation

### Check Proportion Missing

Instead of leaving blank values in missing columns, the Colorado HB 19-1297 jail dataset uses the `not_available` column to annotate and comment on missingness. They add 0 to a datapoint that is missing. Thus, because we cannot use conventional functions like `is.na()` to detect missingness, we will take a look at the jail observations that contain missing data. Here are the 10 most common NA messages.

<code>not_available</code>	n
NA	1792
Eforce does not seprate this data	18
JMS does not calculate this	18
ESTIMATES	13
This is not tracked in our system	13

not_available	n
My JMS doesn't break unsentenced inmates by Gender or Race/Ethnicity.	12
Not able to capture information	11
JMS does not currently break down "sentenced" by gender/race/ethnicity	10
Not able to capture information	10
Population down due to COVID-19	10

Most of the data that is missing is that a jail's JMS (Jail Management System) might not break down types of sentences by gender, race, or ethnicity. When conducting data analysis on race and gender for some particular measures, it will be a good idea to remove these rows, or at least account for them.

Out of 2280 rows, there are 488 (2280-1792) rows with some sort of **not\_available** message.

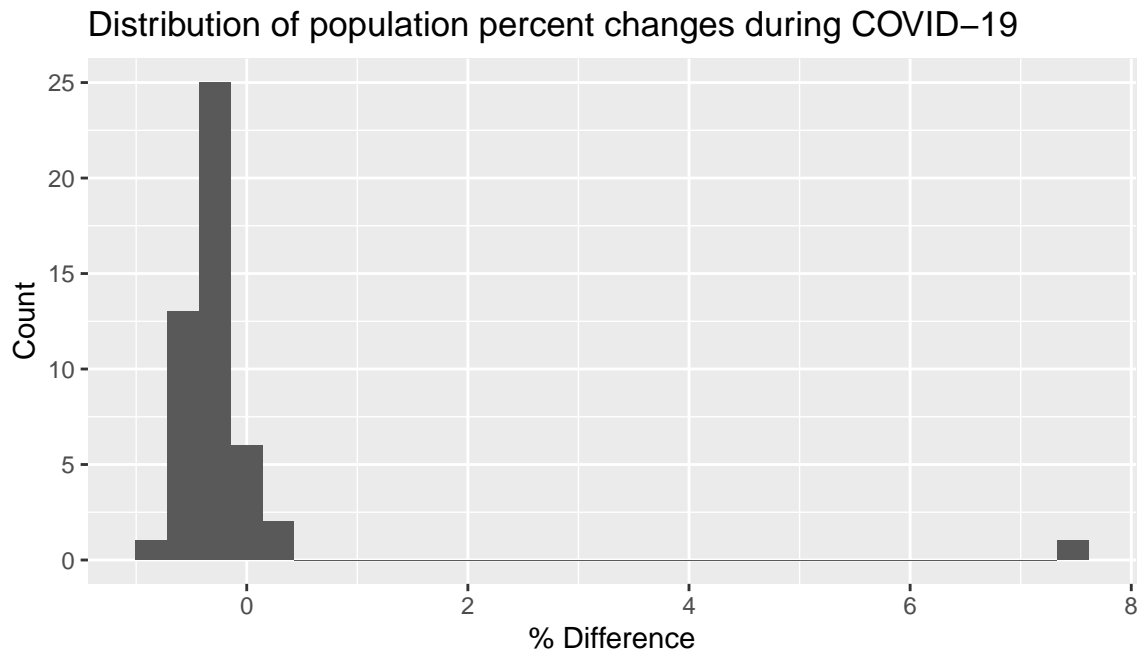
This means that 19.6491228% of the data has some sort of **not\_available** message to it, which is relatively low.

qtr	isNA	n
1	TRUE	170
2	TRUE	162
3	TRUE	156

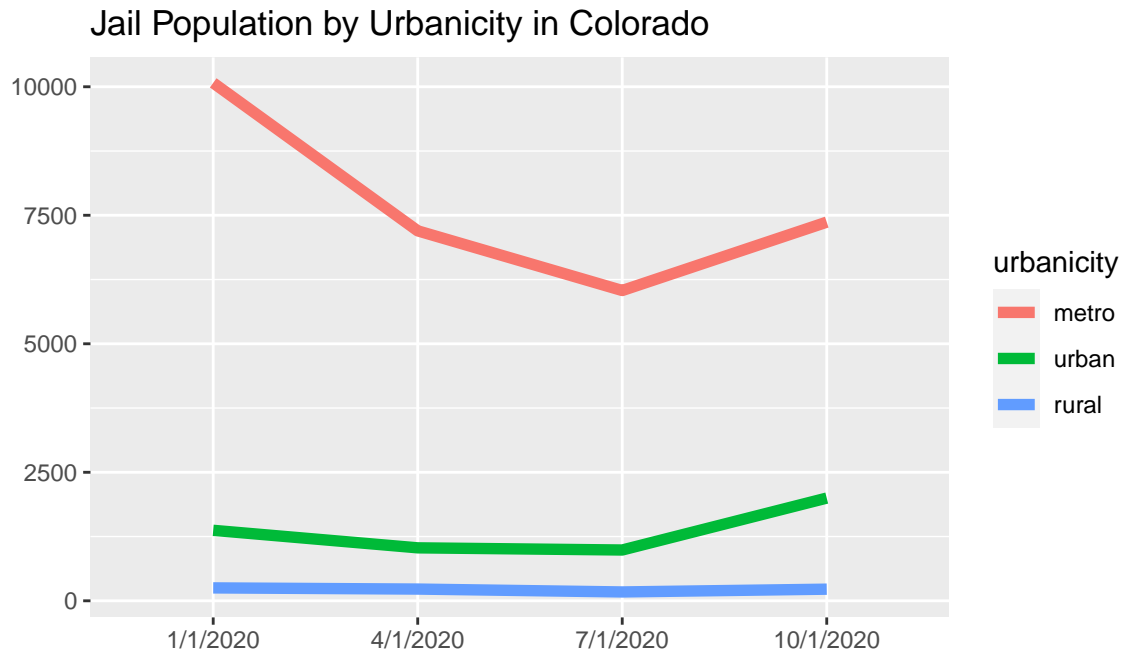
Missingness was generally reduced throughout the four quarters of jail data collection in 2020, possibly suggesting improvements in jail collection throughout this time period.

## Exploratory Data Analysis

The percent change in population numbers for individual jails in Colorado has a unimodal normal distribution. It has a slight right-skew with one significant outlier at +7.5% difference. This means that most jails decreased their jail population, with an average of -0.1642023% difference in jail population:



Jail population visibly decreases significantly in metropolitan jails, but doesn't visibly decrease significantly in rural jails during COVID-19.



## Method

After cleaning our data, we plan on using a few different types of regression models to fit our data and hopefully understand drivers of reduced jail intake/outtake rates during COVID-19. We find that the insights from regression models are highly accessible to the public. Regression analysis can be used to forecast, create time series models, and find the causal effect relationship between variables (given enough variables at hand).

We plan on using linear regression and logistic regression to understand the variables (with special focus to race, gender, and other demographic variables) that may have motivated jail population rates to decrease.

## Model Selection

### Variables:

**liberal** (Factor): 1 if county majority voted for Clinton in 2016, 0 if county majority voted for Trump

**lesscollege\_whites\_pct** (Double): white population with an education of less than a bachelor's degree as a percentage of total population

**population** (Double): Total population of a county

**jail\_male\_pct** (Double): Percent male of a jail

**jail\_black\_pct** (Double): Percent Black of a jail

**jail\_hispanic\_pct** (Double): Percent Hispanic of a jail

**urbanicity** (Factor): metro - Counties in Metropolitan Areas. urban - Counties in Urban Areas. rural - Counties in Rural Areas.

Definitions of Metropolitan, Urban, and Rural are designated by the US Office of Management and Budget (OMB) delineation as of February 2013 **black\_pct** (Double): Black population as a percentage of total population

### Initial Model Fitting

First, put as many predictor variables as possible to create a full model:

Backwards selection with AIC as selection criterion to choose the model with the best AIC.

Based on backwards AIC selection, the two significant predictors for are the percent of males in a jail and whether a jail is in a rural, urban, or metropolitan area.

### Interaction Term

Now, let's use nested F-test to check if the addition of an interaction term between the two variables above is statistically significant:

res.df	rss	df	sumsq	statistic	p.value
44	42.201	NA	NA	NA	NA
42	19.223	2	22.978	25.102	0

Since F-statistic is high and p-value is close to 0, the interaction effect between jail\_male\_pct \* ruralurban\_cc is statistically significant.

## Results

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	12.347	1.334	9.254	0	9.654	15.039
jail_male_pct	-14.067	1.566	-8.984	0	-17.227	-10.907
urbanicityurban	-11.840	1.602	-7.392	0	-15.073	-8.608
urbanicitymetro	-13.398	3.158	-4.242	0	-19.771	-7.025
jail_male_pct:urbanicityurban	13.107	1.906	6.878	0	9.261	16.953
jail_male_pct:urbanicitymetro	14.841	3.749	3.959	0	7.276	22.407

### Urbanicity:

Rural jails have failed to slow to decreasing their jail population during COVID-19 in comparison to urban/metropolitan areas:

- A jail in an urban area is expected to decrease its population 12 percent more than a jail in a rural area, on average.
- A jail in an metropolitan area is expected to decrease its population 13 percent more than a jail in a rural area, on average.

### Male Population Percentage and its Interaction with Urbanicity:

Jails in rural and urban areas with a higher male population have a higher chance to decreasing their jail population. Jails with higher male populations in metropolitan areas have a higher chance of increasing their jail population. Specifically,

- For rural jails, for every one percent increase in male inmates, there is expected to be a 14 percent decrease in jail population between Jan to Sept 2020, on average.
- For urban jails, for every one percent increase in male inmates, there is expected to be a 1 percent decrease in jail population between Jan to Sept 2020, on average.
- For metropolitan jails, for every one percent increase in male inmates, there is expected to be a 1 percent increase in jail population between Jan to Sept 2020, on average.

## Discussion

Our finding that rural jails are more resilient to the COVID-19 pandemic fits with larger national trends. According to the Vera Institute of Justice, there's been a "jail boom" in rural incarceration rates across the country where rural county budgets are allocating more money toward building and expanding new jails instead of housing, transportation, and public health [14]. Further research should be done to see whether the same fiscal incentives that are driving the jail boom are primary causes to our findings about why rural jail populations were so resilient to COVID-19.

Furthermore, this type of analysis would not be possible in North Carolina, because North Carolina doesn't have a regularly updated jail database. Existing national sources of data such as the BJS Census of Jails or Annual Survey are either sparsely collected (every 7 years) or incomplete. Furthermore, the data collection frequency for national data collections by the Bureau of Justice Statistics is at most yearly; a statewide jail database in North Carolina structured like the HB19-1297 Dataset can collect data more frequently. Frequent data collection and information about inmates' race, gender, ethnicity unique to the HB19-1297 Dataset will allow us to fit a more complete regression model to understand the correlation between urbanicity and willingness to reduce population.

In terms of next steps, the statistical model employed in this report, regression modeling, only examines correlation, not causation. As such, in order to better understand why rural jail populations are more resilient, other research techniques should be employed. Qualitative researchers can interview rural jail stakeholders to better understand the response to COVID-19, and whether rural jail incentives came into play. Qualitative researchers and statisticians can use the existing dataset to employ popular causal inference techniques, such as the Structural Causal Model [15]. Additionally, researchers should use this technique to examine whether rural jail incarceration rates were resilient in other states as well during COVID-19.

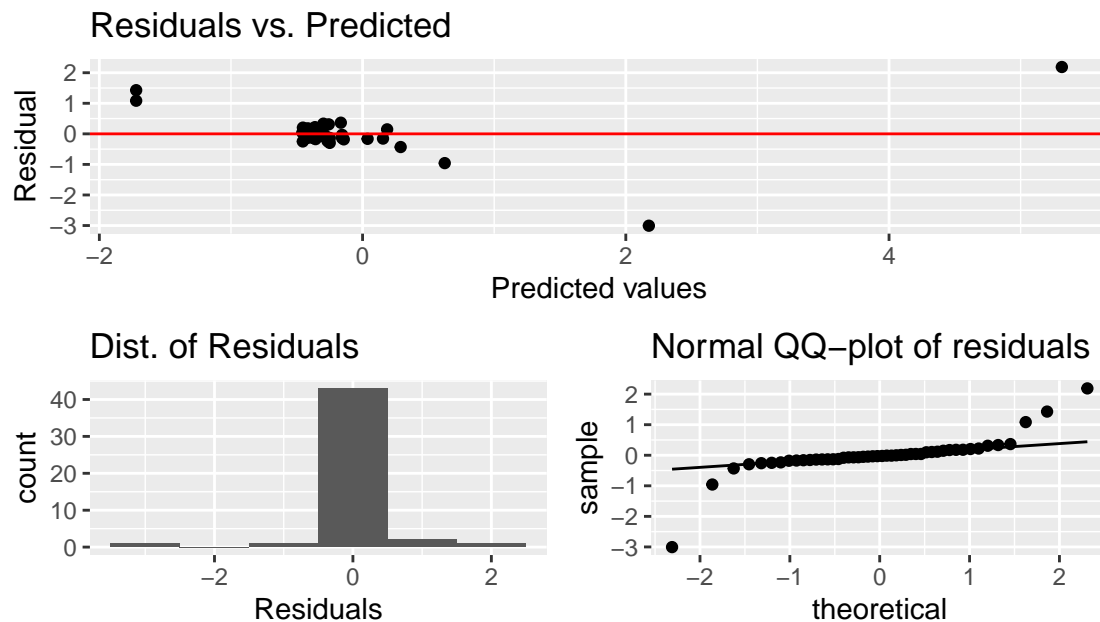
## References

- [1] Ollove, Michael. How COVID-19 in Jails and Prisons Threatens Nearby Communities. <https://pew.org/3eMrIM1>. Accessed 27 Jan. 2021.
- [2] Pressley, Ayanna. Text - H.R.7983 - 116th Congress (2019-2020): COVID-19 in Corrections Data Transparency Act. 7 Aug. 2020, <https://www.congress.gov/bill/116th-congress/house-bill/7983/text>.
- [3] "COVID-19 Model Finds 100,000 More Deaths Than Current Estimates, Due to Failures to Reduce Jails." ACLU. 1 Aug. 2020, [https://www.aclu.org/sites/default/files/field\\_document/aclu\\_covid19-jail-report\\_2020-8\\_1.pdf](https://www.aclu.org/sites/default/files/field_document/aclu_covid19-jail-report_2020-8_1.pdf)
- [4] Haskell-Hoehl, Micah. "Policymakers Must Lift the Veil on COVID-19 in Jails and Prisons." Vera Institute of Justice. 2 Sept. 2020, <https://www.vera.org/blog/covid-19-1/policymakers-must-lift-the-veil-on-covid-19-in-jails-and-prisons>
- [5] Kaplan, Jacob. California Jail Profile Survey 1995-2020. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2020-06-29. <https://doi.org/10.3886/E104560V6>
- [6] Kaplan, Jacob. Texas Commission on Jail Standards Data 1992-2017. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2018-07-14. <https://doi.org/10.3886/E104643V1>
- [7] Polis, Jard. D 2020 016 Colorado Executive Order. State of Colorado Governor. 25 Mar. 2020, <https://drive.google.com/file/d/1I9dyGYreaekjXq0ME0fDTcTEvQCtMr4/view>
- [8] Ashby, Charles. COVID forces changes to jail, making expansion unneeded. The Daily Sentinel. 6 Sep. 2020, [https://www.gjsentinel.com/news/western\\_colorado/covid-forces-changes-to-jail-making-expansion-unneeded/article\\_33242d3a-eed4-11ea-b079-af5d33237325.html](https://www.gjsentinel.com/news/western_colorado/covid-forces-changes-to-jail-making-expansion-unneeded/article_33242d3a-eed4-11ea-b079-af5d33237325.html)
- [9] HB19-1297 Jail Data Dashboard. [https://cdpsdocs.state.co.us/ors/Data/Data\\_Instruments/HB1297/Dashboard/HB19-1297v2.html](https://cdpsdocs.state.co.us/ors/Data/Data_Instruments/HB1297/Dashboard/HB19-1297v2.html)
- [10] HB19-1297 Jail Capacity Data Collection Bill Text. <https://leg.colorado.gov/bills/hb19-1297>

- [11] “Who’s in Colorado’s County Jails? Data-Collection Legislation Would Provide Answers. | Colorado Freedom of Information Coalition.” Colorado Freedom of Information Coalition, 26 Apr. 2019, [coloradofaic.org/whos-in-colorados-county-jails-data-collection-legislation-would-provide-answers/](https://coloradofaic.org/whos-in-colorados-county-jails-data-collection-legislation-would-provide-answers/). Accessed 27 Jan. 2021.
- [12] “Data | MIT Election Lab.” Mit.edu, 14 Jan. 2021, [electionlab.mit.edu/data](https://electionlab.mit.edu/data). Accessed 27 Jan. 2021.
- [13] “BJS Data Collection.” Bureau of Justice Statistics. <https://www.bjs.gov/index.cfm?ty=dcdetail&iid=255>. Accessed April 1, 2021.
- [14] Heiss, Jasmine. “Why reimagining safety looks different in rural America.” Washington Post, April 29, 2021. Accessed April 30, 2021. [https://www.washingtonpost.com/opinions/2021/04/29/reimagine-safety-cities-rural-communities/?utm\\_campaign=wp\\_opinions&utm\\_source=twitter&utm\\_medium=social](https://www.washingtonpost.com/opinions/2021/04/29/reimagine-safety-cities-rural-communities/?utm_campaign=wp_opinions&utm_source=twitter&utm_medium=social)
- [15] Pearl, Judea. “Causal inference in statistics: An overview”, September 2009. UCLA Computer Science Department. [https://ftp.cs.ucla.edu/pub/stat\\_ser/r350.pdf](https://ftp.cs.ucla.edu/pub/stat_ser/r350.pdf)

## Appendix

### 4 Conditions for Linear Regression:



To ensure the accuracy of the model, we looked at four different conditions for linear regression:

1. **Linearity:** Partially Satisfied - Although there is no distinguishable pattern in the residuals vs. predicted plot, the points are generally scattered. There seem to be some high leverage points that are worth investigating. However, leeway will be given given the small number of observations. Overall, the residuals are randomly scattered.
2. **Normality:** Satisfied - the points generally fall along a straight diagonal line on the normal quantile plot.
3. **Constant Variance:** Satisfied - the vertical spread of the residuals remains relatively constant across the plot.
4. **Independence:** Satisfied - The error for one county does not tell us anything about the error for another county. We also put urbanicity of the counties as one of the variables, which avoids some of the problems of spatial collinearity/correlation. However, in order to improve this model, we can add a “region” variable to account for the locations of the counties in the state of Colorado.