

Question

What industries show the highest COVID-19 outbreak incidents in Colorado, what areas are these industries in, and what are the primary ethnicities of the labor force for these industries? Is the work environment contributing to the spread of COVID-19?

The primary data sets that I will be focusing on come from the Colorado Department of Public Health and Environment. The datasets are updated weekly, so the latest data for the project will be coming from 7/15/2020. I will not limit myself to this data, but it will be the backbone of my analysis.

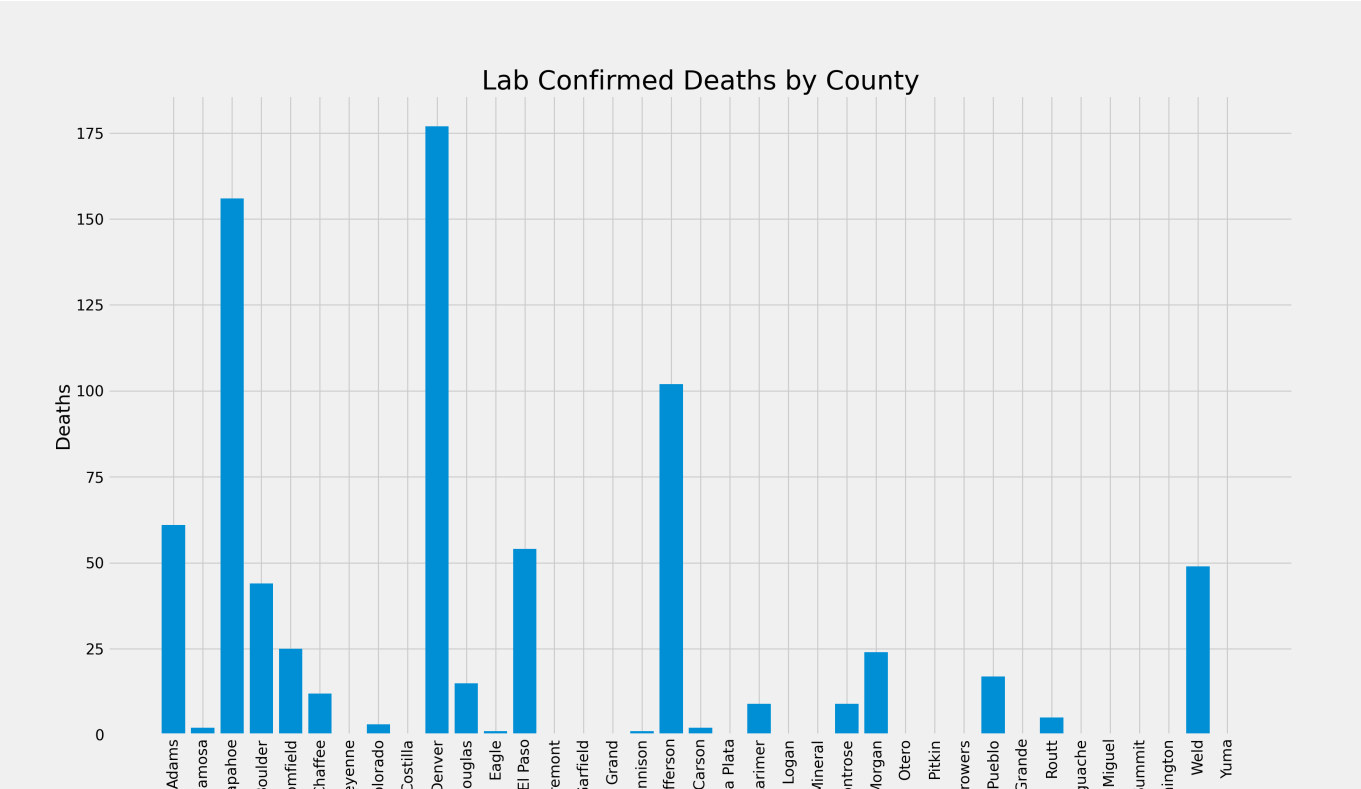
I have performed some preliminary analysis on the dataset and the results are highlighted below.

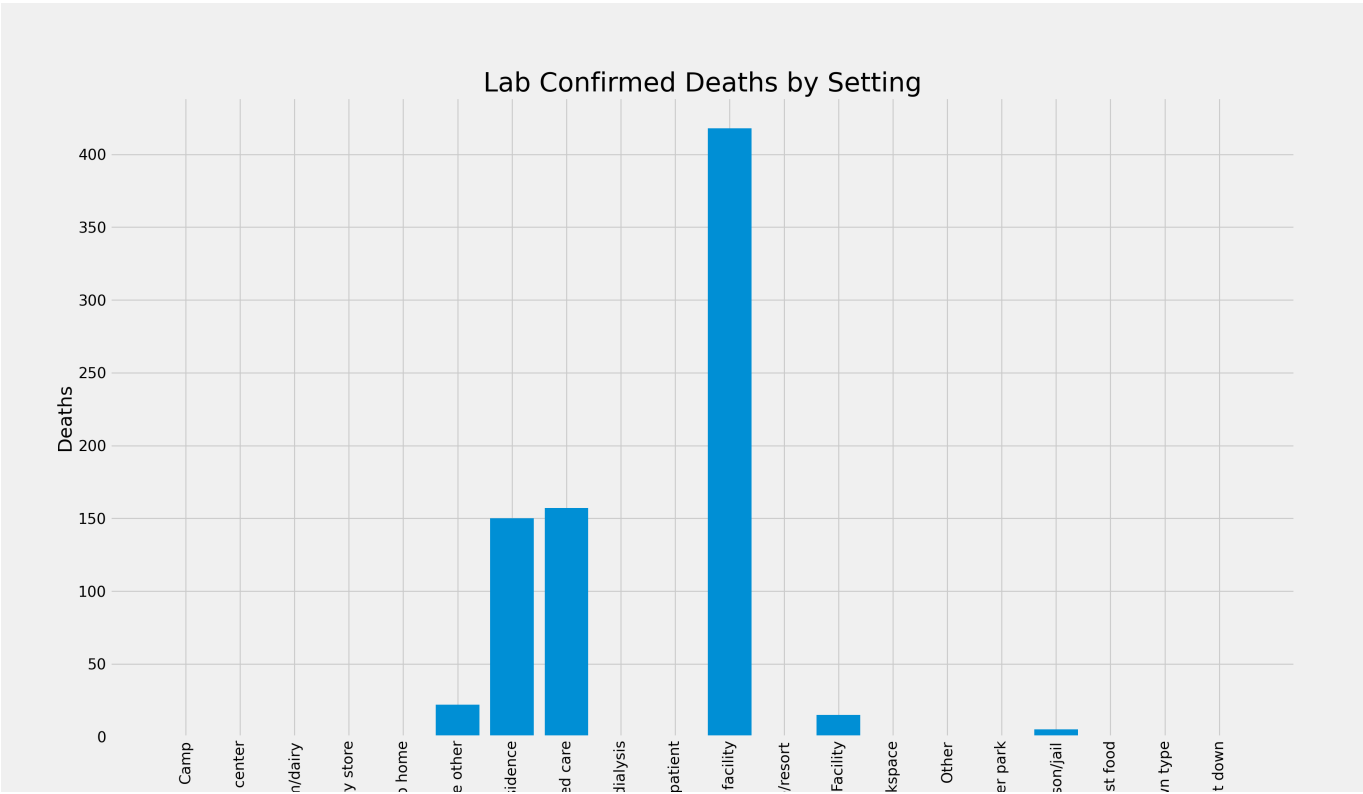
Initial EDA

The industry categories for the data, labeled '**Setting type**' and '**If setting type is other, specify**', total 19 and 110 respectively. These will need to be consolidated to approximately 14 categories, which will be defined by Sandra.

There are 37 Counties in the data set, the other 37 Counties in Colorado have not reported industrial outbreaks.

Below are are two graphs I have generated comparing COVID-19 lab proven deaths in the industries and counties.





4

The other columns that can be investigated from the initial dataset are:

Setting name

Investigation status

Date Outbreak Resolved

Setting type

If setting type is other, specify

Colorado county (exposure location)

Date illnesses were determined to be an outbreak

Number of residents positive for COVID-19 (lab confirmed)

Number of residents with probable COVID-19 (NOT lab confirmed)

Number of COVID-19 deaths (NOT lab confirmed/probable)

Number of staff who are positive for COVID-19 (lab confirmed)

Number of staff with probable COVID-19 (NOT lab confirmed)

Number of COVID-19 staff deaths (lab confirmed/confirmed)

Number of COVID-19 staff deaths (NOT lab confirmed/probable)

Number of attendees who are positive for COVID-19 (lab confirmed)

Number of attendees with probable COVID-19 (NOT lab confirmed)

Number of COVID-19 attendee deaths (lab confirmed/confirmed)

Number of COVID-19 attendee deaths (NOT lab confirmed/probable)

MVP

The main objective of this project is to determine if workers in certain industries are more vulnerable to catching COVID-19, what the primary ethnicity employed by those industries is, and what factors would contribute to the transmissibility of COVID-19 in those higher risk industries.

MVP++

A secondary goal is to take a deeper dive into these industries. I would like to take a look at wage/salary levels, exposure rates (with the public), risk aversion, etc. that may contribute to the spread of COVID-19 within the population that is employed by these industries.