

# Final Report

Brandyn Ruiz

## Website Home

### Introduction

When the pandemic first started in early March and the effects of the corona virus inhibiting a person's ability to breathe I first thought if our state's air pollution would also be a contributor to contracting the virus. With air pollution being so different in all of the 58 counties within California, I wanted to see if contracting confirmed cases of covid would also be on the rise if a person's breathing ability was already impacted by air pollution. My hypothesis that I want to further explore is whether there is an association between air quality and confirmed cases of covid amongst people within the counties of California.

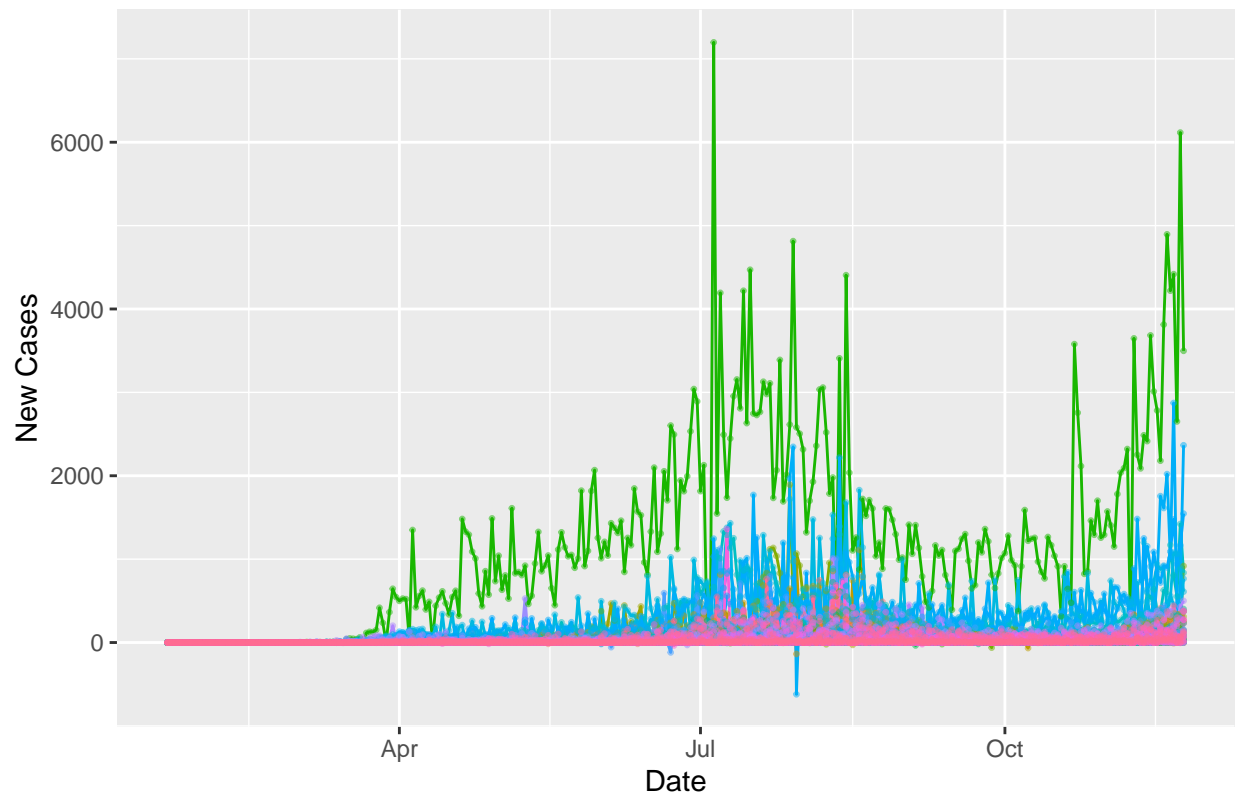
### Covid Cases within the States

```
download.file("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_data.csv")
covidGithub <- data.table::fread("time_series_covid19_confirmed_US.csv")
```

### Incidence Cases Across the Pandemic

```
# Time series visual of incidence cases from the start of the pandemic
ggplot(covidGithubmelt, aes(Date, new_cases, color = county))+
  geom_line()+
  geom_point(size = .5, alpha = 0.5)+
  labs(title = 'Incidences of Covid Cases in California by County', y = 'New Cases')+
  theme(legend.position = 'none')
```

## Incidences of Covid Cases in California by County

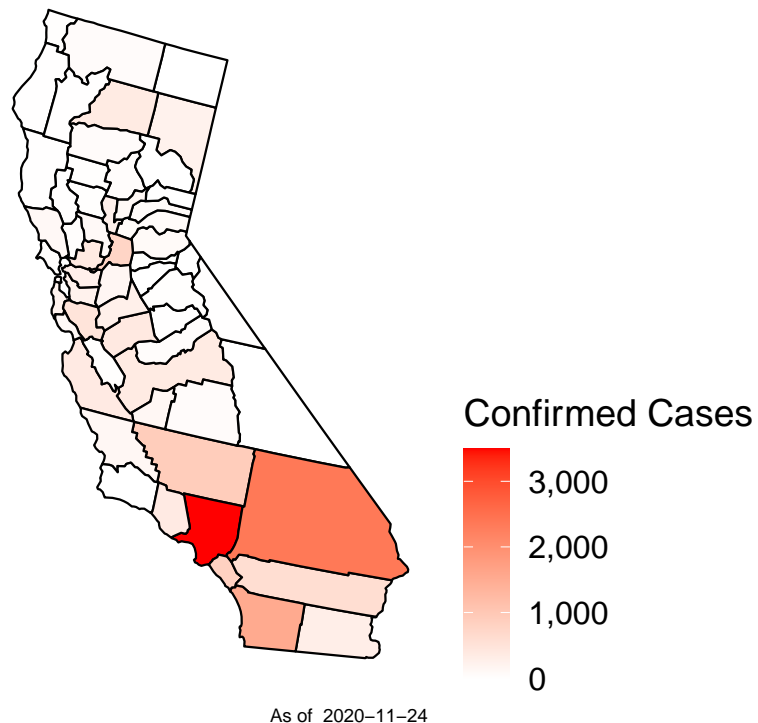


## Incidence Cases

```
plot_usmap(data = mapdf, values='new_cases', include = 'CA')+
  scale_fill_continuous(low = 'white', high = 'red', name = 'Confirmed Cases', label = scales::comma)+
  labs(title = 'Incidence Cases by California County',
        subtitle = 'Source: Johns Hopkins University Center for Systems Science and Engineering (JHU CSS)',
        caption = paste('As of ', as.Date(max(mapdf$Date))))+
  theme(plot.title = element_text(hjust = 0.5, vjust = 0.25, size = 18), legend.position = "right",
        legend.title = element_text(size = 14), legend.text = element_text(size = 12))
```

# Incidence Cases by California County

Source: Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE)

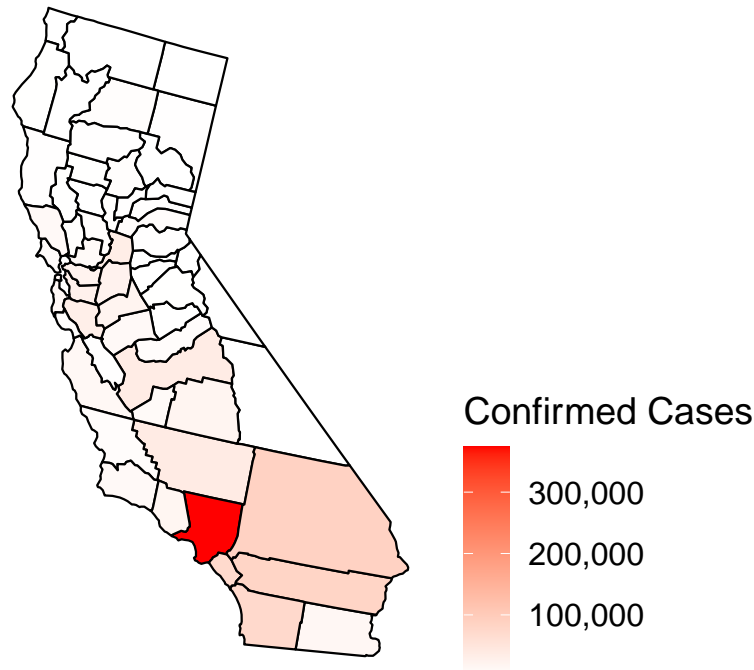


## Total Confirmed Cases

```
plot_usmap(data = mapdf, values='Confirmed', include = 'CA')+  
  scale_fill_continuous(low = 'white', high = 'red', name = 'Confirmed Cases', label = scales::comma)+  
  labs(title = 'Total Cases in California by County',  
        subtitle = 'Source: Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE)',  
        caption = paste('As of ', as.Date(max(mapdf$Date))))+  
  theme(plot.title = element_text(hjust = 0.5, vjust = 0.25, size = 18), legend.position = "right",  
        legend.title = element_text(size = 14), legend.text = element_text(size = 12))
```

# Total Cases in California by County

Source: Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE)



As of 2020-11-24

## Exploration of Covid Cases within California

Figure 1: Incidence

Figure 2: Prevalence

## US Census data of California

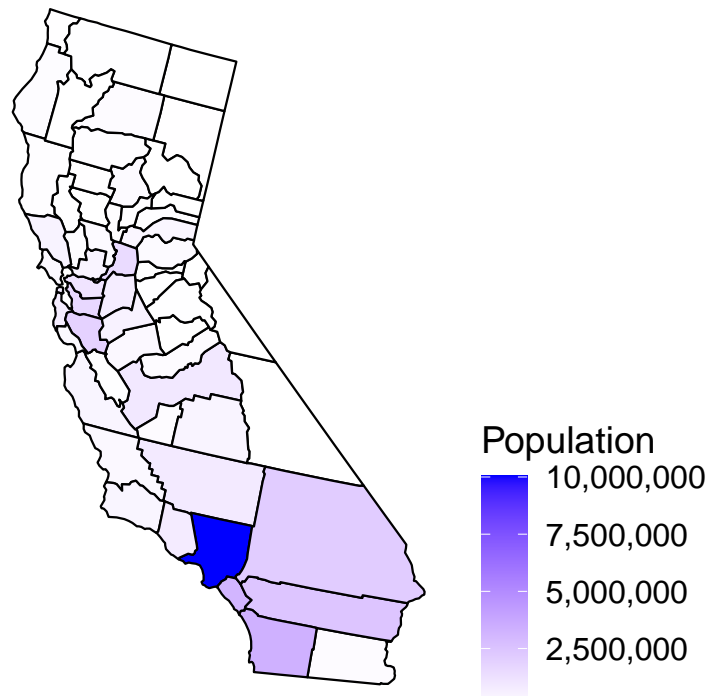
```
CAcensus <- read_csv("data/cc-est2019-alldata-06.csv")
```

## County Populations Within California

```
plot_usmap(data = Popmapdf, values='TOT_POP', include = 'CA')+  
  scale_fill_continuous(low = 'white', high = 'blue', name = 'Population', label = scales::comma)+  
  labs(title = '2019 County Population Estimates', subtitle = 'Source: US Census')+  
  theme(plot.title = element_text(hjust = 0.5, vjust = 0.25, size = 18), legend.position = "right",  
        legend.title = element_text(size = 14), legend.text = element_text(size = 12))
```

# 2019 County Population Estimates

Source: US Census



## Running Total of Confirmed Cases by Population of Counties

```
# Visual of how county population influences covid cases
# cvd_pop%>%
#   plot_ly(x = ~TOT_POP, y = ~Confirmed,
#           type = 'scatter', mode = 'markers', color = ~county,
#           size = ~TOT_POP, sizes = c(5, 70), marker = list(sizemode='diameter', opacity=0.5),
#           hoverinfo = 'text',
#           text = ~paste( paste(county, ":", sep=""), paste(" Cases per 100k: ", per100k, sep=""),
#                       paste(' Population: ', TOT_POP, sep=""), sep = "<br>"),
#           width = 800)%>%
#   layout(title = "Covid Cases vs Population of Each County",
#           yaxis = list(title = "Cases per 100k"), xaxis = list(title = "Population"))

# ggplot(cvd_pop, aes(TOT_POP, Confirmed, color = county, size = TOT_POP))+
#   geom_point()

# Could mention plot from website
```

## Environmental Protection Agency (EPA) Air Quality Index

```
# Daily AQI for every county in California
csvAQI_data <- read_csv("data/ad_viz_plotval_data.csv")
```

## Air Quality Within California

```
# Averaged AQI for each county
AQI%>%
  plot_ly(x = ~COUNTY, y = ~MeanAQI, type = 'box', color = ~COUNTY, width = 750)%>%
  layout(title = 'Air Quality by California County', yaxis = list(title = 'AQI Value'),
         xaxis = list(title = 'County'), showlegend = FALSE)
```

## New Page

### Combined Visuals

```
# Stacked visuals of both AQI over covid cases during the pandemic

# Air Quality over course of the pandemic
line <- AQI%>%
  filter(COUNTY == 'Los Angeles')%>%
  plot_ly(x = ~Date, y = ~MeanAQI, type = 'scatter', mode = 'line',
          hoverinfo = 'text',
          text = ~paste( paste(COUNTY, ":", sep=""), paste(' Date: ', Date, sep = ''),
                        paste(' Avg AQI: ', MeanAQI, sep=""), sep = "<br>"), width = 800)%>%
  layout(title = 'Los Angeles County', yaxis = list(title = 'AQI Value'),
         xaxis = list(title = 'Date'))

# Incidences of covid cases over time
line2 <- covidGithubmelt%>%
  filter(county == 'Los Angeles')%>%
  plot_ly(x = ~Date, y = ~new_cases, type = 'scatter', mode = 'line',
          hoverinfo = 'text',
          text = ~paste( paste(county, ':', sep = ''), paste(' Date: ', Date, sep = ''),
                        paste(' New Cases: ', new_cases, sep = ''), sep = "<br>"), width = 800)

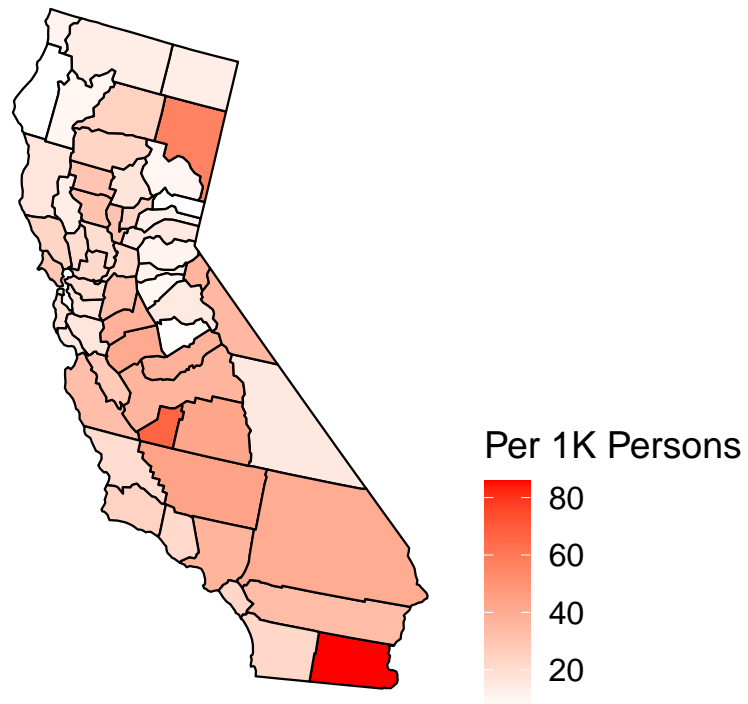
time_series <- subplot(line, line2, nrows = 2, shareX = TRUE)

time_series
```

### Attack Rate Normalized

```
# Incidence Rate of Covid Cases Normalized by County Population
plot_usmap(data = map_merge, values='ratePer1k', include = 'CA')+
  # scale_fill_distiller(type = 'qual', palette = 'YlOrRd')+
  scale_fill_continuous(low = 'white', high = 'red', name = 'Per 1K Persons')+
  labs(title = 'Attack Rate of Covid Cases',
       caption = paste('As of ', as.Date(max(mapdf$Date))))+
  theme(plot.title = element_text(hjust = 0.5, vjust = 0.25, size = 18), legend.position = "right",
        legend.title = element_text(size = 14), legend.text = element_text(size = 12))
```

## Attack Rate of Covid Cases



As of 2020-11-24