Utilizing Google’s Mobility Report to understand Infectious Disease Dynamics

Vincent Nguyen

**Authors**

* Vincent Nguyen

**Author affiliations**

1. College of Public Health, University of Georgia, Athens, GA, USA.

Corresponding author: vln27447@uga.edu

# 1. Summary/Abstract

*Write a summary of your project.*

# 2. Introduction

The emergence of the novel corona-virus SARS-CoV-2 caused one of the most significant global health crises in modern history. First identified in Wuhan, China, in late 2019, the virus’s transmissibility and globalism led to its rapid spread worldwide, to leading the World Health Organization to declare it a pandemic on March 11, 2020. Just two days later, on March 13th, 2020, the Trump Administration declared a nationwide emergency due to the COVID-19 pandemic, marking the biggest virus outbreak since \_\_\_. In response, public health agencies implemented measures to curb the virus’ spread including travel restrictions, social distancing, and lockdown procedures (CDC 2024)(1).

## 2.1 Description of data and data source

This project explores mobility dynamics, as measured by [Google’s Community Mobility Reports](https://www.google.com/covid19/mobility/), in an infectious disease context to understand and forecast disease spread. By integrating time-series infection data (provided by [Johns Hopkins University](https://github.com/CSSEGISandData/COVID-19)), the analysis aims to simulate disease outbreak scenarios and provide insights into optimal intervention strategies for mitigating future outbreaks. Additionally, the [County Health Ranking](https://www.countyhealthrankings.org/) is used to assess these dynamics within the context of a county’s health care access and performance.  
  
Google’s Community Mobility Reports are anonymized insights collected from Google products, such as maps, that detail population movement trends across various categories, such as retail stores, workplaces, residencies, and more.  
  
The COVID-19 Data Repository by the Center for Systems Science and Engineering at Johns Hopkins University is a comprehensive data set that tracks global COVID-19 cases, recoveries, and deaths. It was updated regularly and compiles data from various official sources to support research.

The County Health Ranking is a data set that provides a comprehensive overview of various health factors and outcomes across U.S. counties. Data collected includes metrics on healthcare access, insurance coverage, income levels, violent crime, air and water quality, transportation access, and more. [Not sure about the inclusion of this]

## 2.2 Questions/Hypotheses to be addressed

How have mobility dynamics influenced the spread of COVID-19 in the U.S. during 2020-2022?  
  
Can mobility data, in combination with COVID-19 case counts, be used to forecast future infection trends?

How do socio-economic and health factors influence the relationship between mobility patterns and COVID-19 infection rates across U.S. counties?

Investigate if counties with similar health rankings but different mobility trends experienced different outbreak severities, potentially revealing the effectiveness of local interventions

Examine ifthere was lag time between changes in mobility and subsequent changes in infection rates

To cite other work (important everywhere, but likely happens first in introduction), make sure your references are in the bibtex file specified in the YAML header above and have the right bibtex key. Then you can include like this:

# 3. Methods

*Describe your methods. That should describe the data, the cleaning processes, and the analysis approaches. You might want to provide a shorter description here and all the details in the supplement.*

## 3.1 Schematic of workflow

Sometimes you might want to show a schematic diagram/figure that was not created with code (if you can do it with code, do it). **?@fig-schematic** is an example of some - completely random/unrelated - schematic that was generated with Biorender. We store those figures in the assets folder.

## 3.2 Data acquisition

*COVID-19 case data was provided by the Center of Systems and Science and Engineering at JHU and mobility data was provided by Google’s Community Mobility Report.*

## 3.3 Data import and cleaning

Since the pandemic spanned across 3 years, data on mobility was collected from 2020 to 2022. The data on mobility is quite extensive and is stratified by county, resulting in the large amount of mobility. The mobility data measures changes in mobility patterns across different key areas, such as retail, workplace, residential, and more.

The time series data has 3342 observations of 1154 variables. Each row represents a different city. The columns record cases as days progress and locational data like longitude and latitude.

library(here)

Warning: package 'here' was built under R version 4.3.3

here() starts at C:/Users/vince/OneDrive/Desktop/MADA-Project/project-MADA

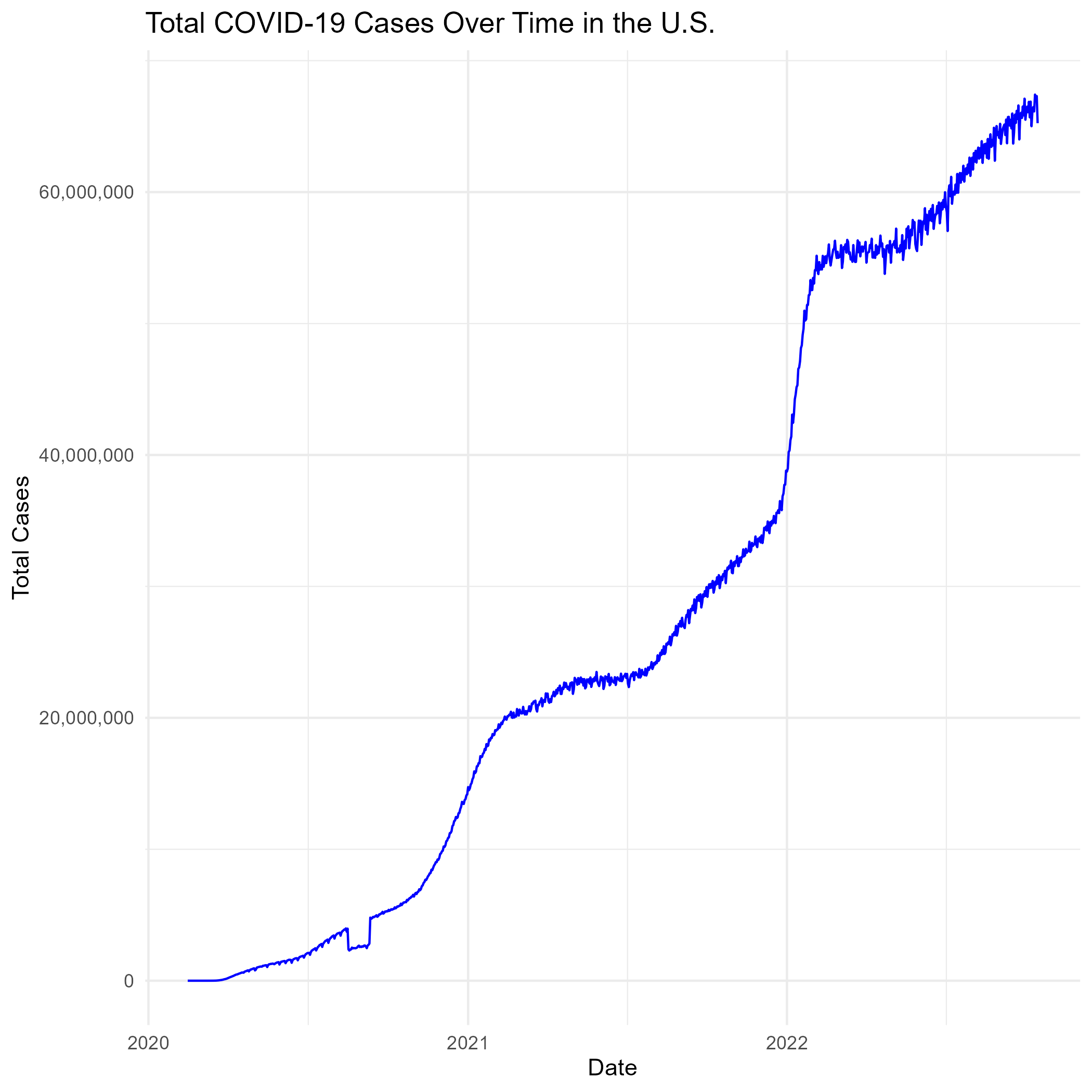
# importing data  
mobility\_data\_2020 <- read.csv(here("data", "raw-data", "2020\_US\_Region\_Mobility\_Report.csv"))  
mobility\_data\_2021 <- read.csv(here("data", "raw-data", "2021\_US\_Region\_Mobility\_Report.csv"))  
mobility\_data\_2022 <- read.csv(here("data", "raw-data", "2022\_US\_Region\_Mobility\_Report.csv"))  
  
time\_series\_covid <- read.csv(here("data", "raw-data", "time\_series\_covid19\_confirmed\_US.csv"))  
  
chr\_data\_2024 <- read.csv(here("data", "raw-data", "chr\_trends\_csv\_2024.csv"))

## 3.4 Statistical analysis

# 4. Results

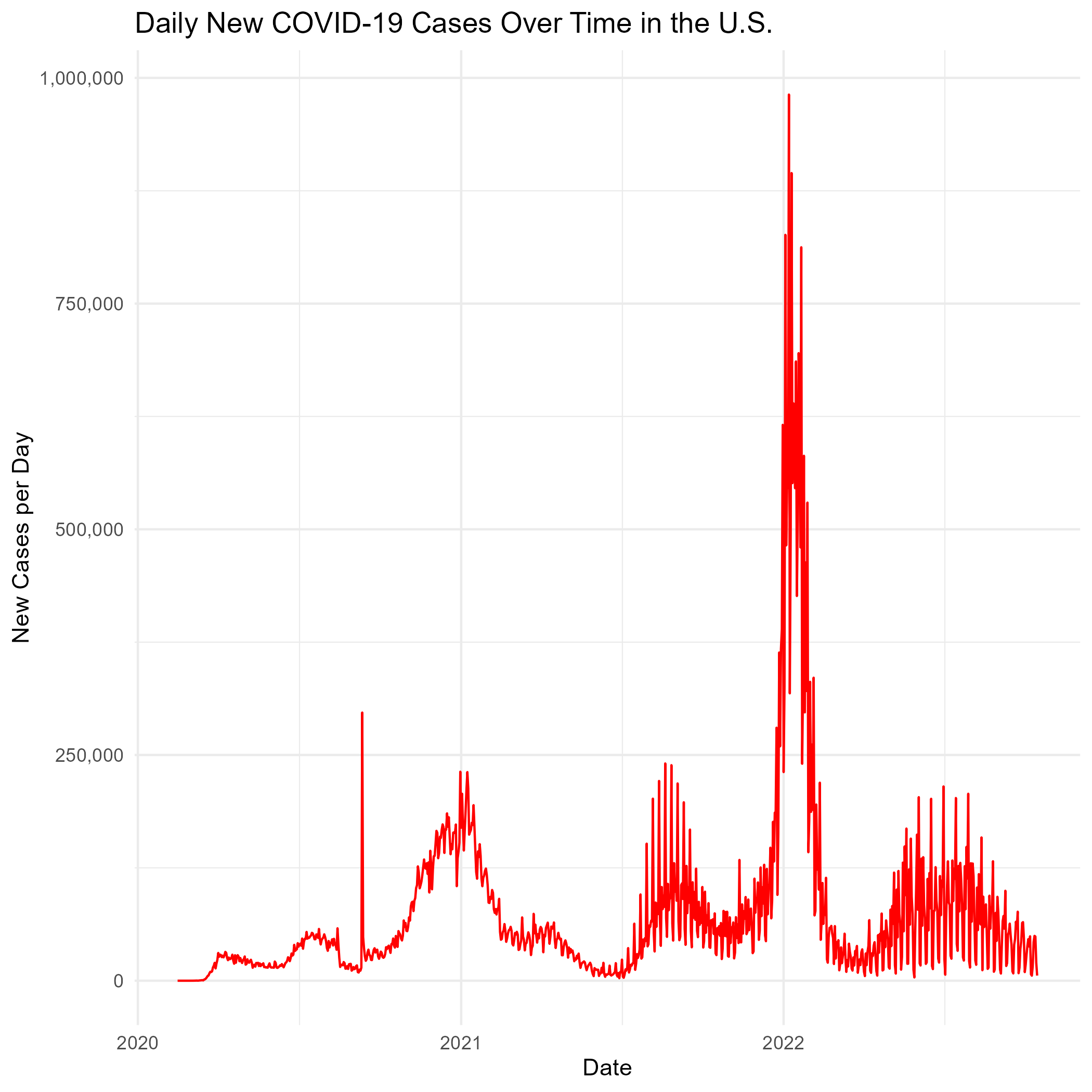
## 4.1 Exploratory/Descriptive analysis

To gain a clearer perspective on the overall scale of the pandemic, we visualized the total number of COVID-19 cases over time. This plot provides a overview of how case counts evolved.The graph reveals a large spike in cases in the beginning of 2022. This does correspond with figures further.



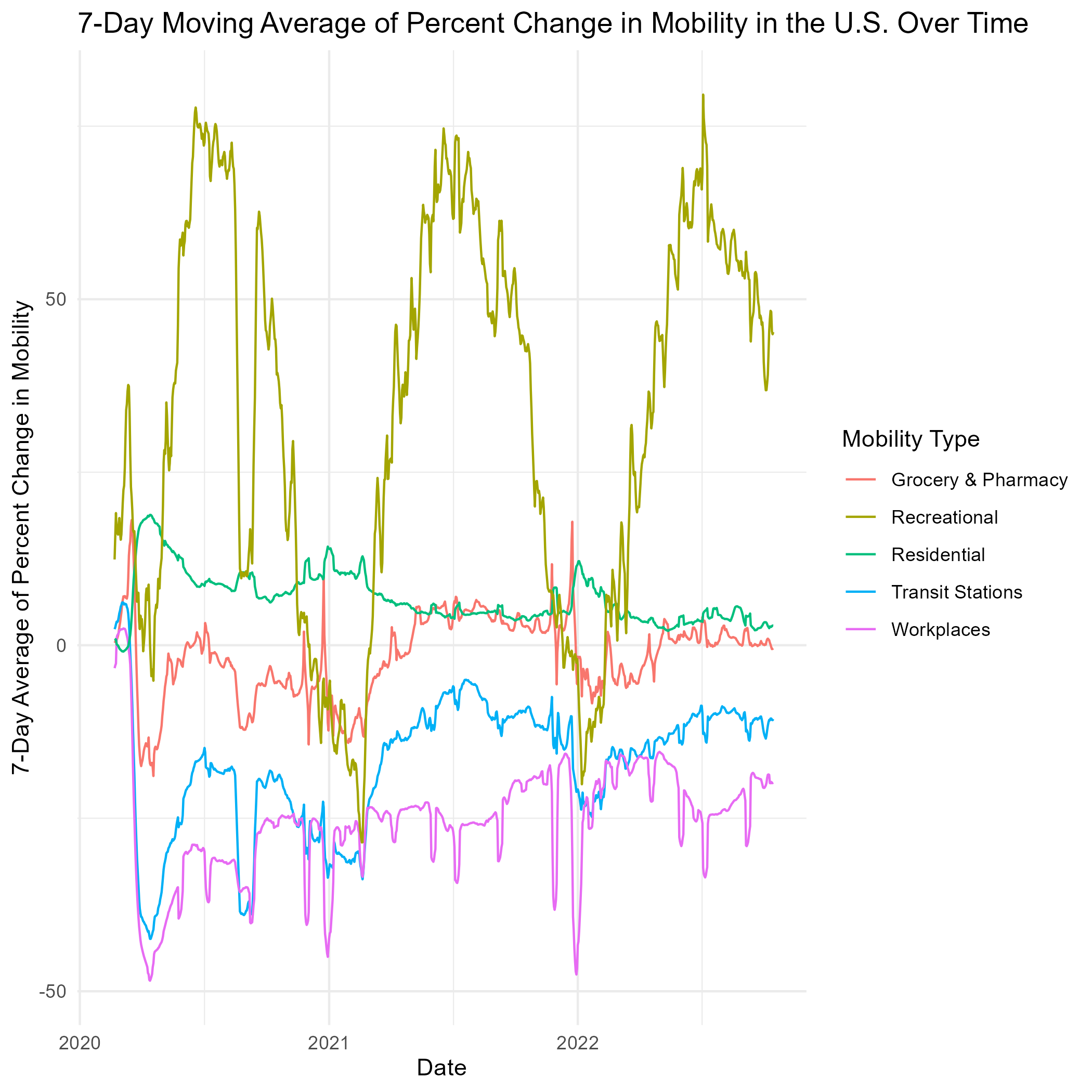
Total COVID-19 Cases overtime

To gain a clearer perspective on the overall scale of the pandemic, we visualized new caes of COVID-19 over time. This plot provides a overview of how case counts and infection evolved.The graph reveals a large spike in cases in the beginning of 2022 which corresponds with the total case count graph.



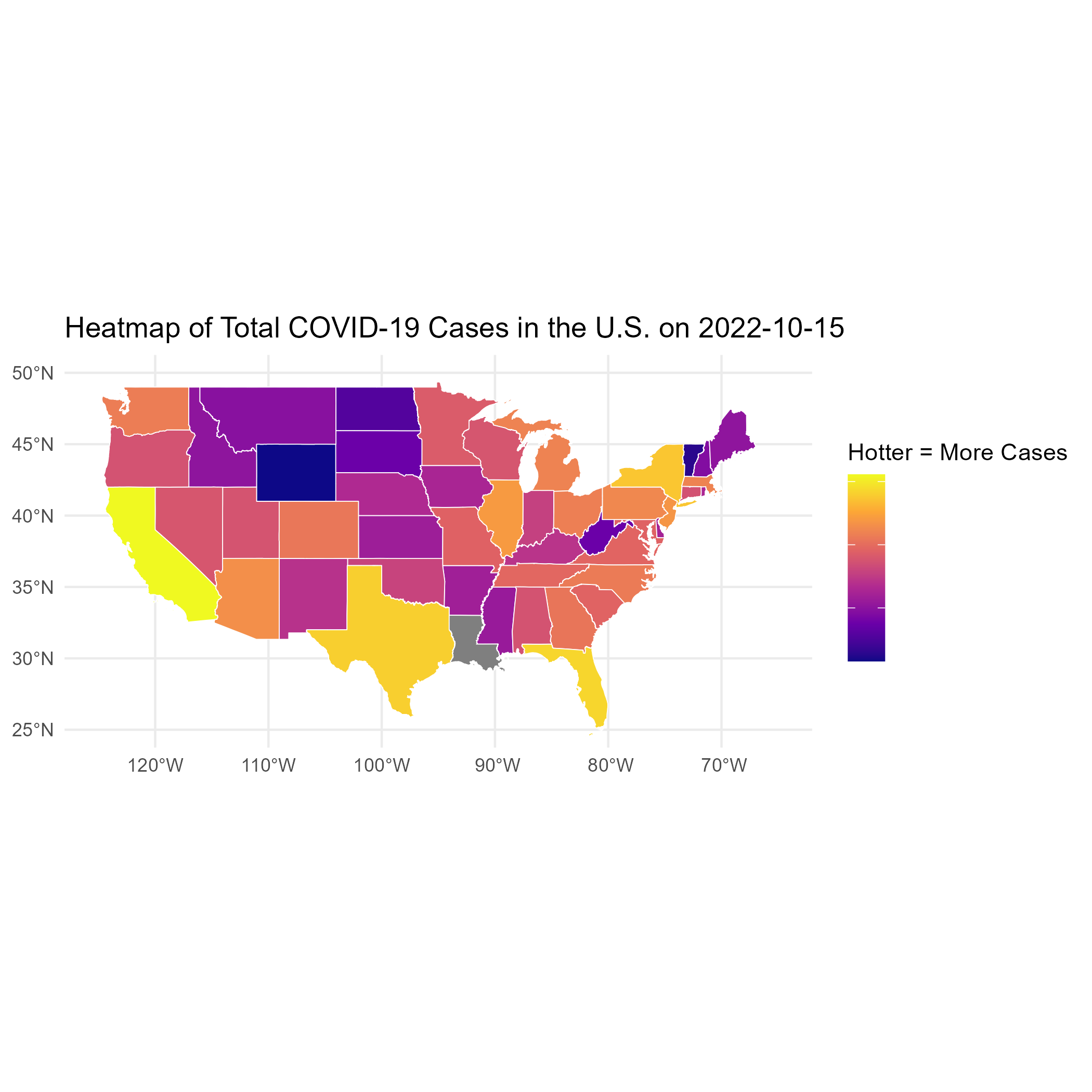
Daily New COVID-19 Cases overtime in the U.S.

To gain a deeper understanding of how mobiltiy patterns shifted throughout the pandemic, weekly changes in mobility were visualized. This plot highlights fluctuations in movement across various sectors. Mobility was increased towards residential and recreational areas which warrants further exploration. This can help offer insight into how behavior towards COVID-19 restriction interventions evolved over time.

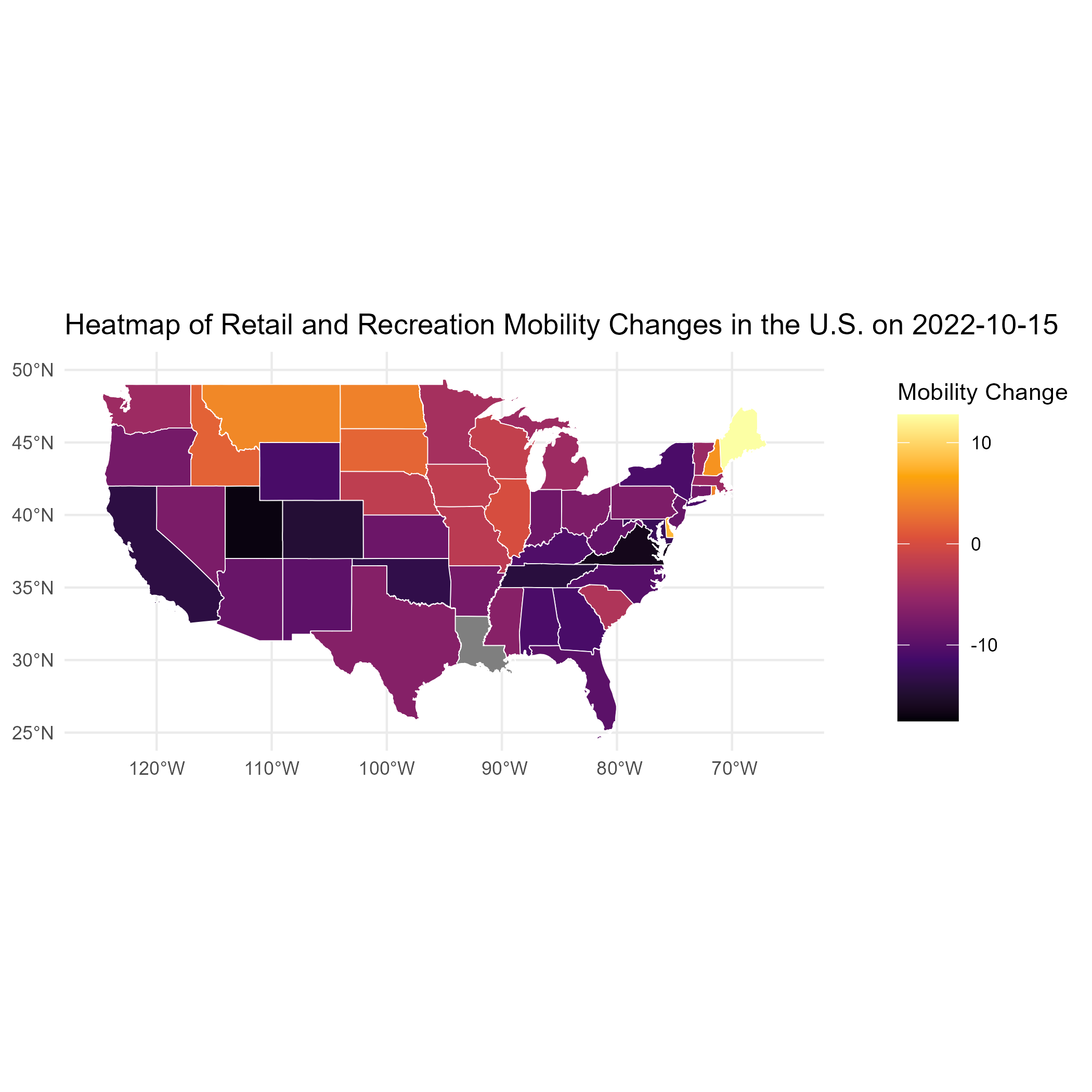


Percent Changes in Mobility per week in the U.S.

As hypothesized, an rise in COVID-19 case count is expected to lead to a decline in mobility to non-residential locations, as individuals alter their behavior in resposne to increasing risks. To explore this idea, total case count and changes in mobility were visualzied across the United States.



Total COVID-19 Cases on 10/15/2022 mapped to the U.S.



Average of mobility changes on 10/15/2022 mapped to the U.S.

Analysis of the mapped data reveals a trend: states with higher case counts generally experienced more significant decreases in mobility. This suggests a correlation between rising COVID-19 case count and a greater reduction in movement to non-residential locations.

## 4.2 Basic statistical analysis

*To get some further insight into your data, if reasonable you could compute simple statistics (e.g. simple models with 1 predictor) to look for associations between your outcome(s) and each individual predictor variable. Though note that unless you pre-specified the outcome and main exposure, any “p<0.05 means statistical significance” interpretation is not valid.*

## 4.3 Full analysis

*Use one or several suitable statistical/machine learning methods to analyze your data and to produce meaningful figures, tables, etc. This might again be code that is best placed in one or several separate R scripts that need to be well documented. You want the code to produce figures and data ready for display as tables, and save those. Then you load them here.*

# 5. Discussion

## 5.1 Summary and Interpretation

*Summarize what you did, what you found and what it means.*

## 5.2 Strengths and Limitations

*Discuss what you perceive as strengths and limitations of your analysis.*

## 5.3 Conclusions

*What are the main take-home messages?*

*Include citations in your Rmd file using bibtex, the list of references will automatically be placed at the end*

This paper (**leek2015?**) discusses types of analyses.

These papers (**mckay2020?**,**mckay2020a?**) are good examples of papers published using a fully reproducible setup similar to the one shown in this template.

Note that this cited reference will show up at the end of the document, the reference formatting is determined by the CSL file specified in the YAML header. Many more style files for almost any journal [are available](https://www.zotero.org/styles). You also specify the location of your bibtex reference file in the YAML. You can call your reference file anything you like.

# 6. References

1. for Disease Control C, Prevention. COVID-19 timeline. 2024;(<https://www.cdc.gov/museum/timeline/covid19.html>)