REGRESSION ANALYSIS OF RESILIENCE AND COVID-19 IN IDAHO COUNTIES

by

Ishrat Zaman

A thesis

submitted in partial fulfillment of

the requirements for the degree of

Master of Science in Mathematics

Boise State University

August 2022

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

Ishrat Zaman

Thesis Title: Regression Analysis of Resilience & Covid-19 in Idaho Counties

Date of Final Oral Examination: 28 June 2022

The following individuals read and discussed the thesis submitted by student Ishrat Zaman, and they evaluated his presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Jodi Mead, Ph.D.	Chair, Supervisory Committee
Kyungduk Ko, Ph.D.	Member, Supervisory Committee
Brittany Brand, Ph.D.	Member, Supervisory Committee

The final reading approval of the thesis was granted by Jodi Mead, Ph.D., Chair of the Supervisory Committee. The thesis was approved by the Graduate College.

ABSTRACT

Global pandemic Coronavirus Disease 2019 (COVID-19) has serious harmful effects on our day-to-day lives. To overcome challenges such as this, critical preparedness, readiness, and response actions are required. This thesis uses estimates of community resilience available through the CRE Tool, published by the US Census Bureau, and COVID19 cases published by John Hopkins Coronavirus Research Center for Idaho counties. Simple linear regression analysis was performed to identify a correlation between COVID-19 cases and deaths in Idaho counties and measures of their resilience. Understanding this correlation could lead to better estimation and prediction of the effect of disasters in Idaho's counties.

We determined that there is a weak negative correlation exists between the number of COVID-19 cases and the percentage of people who fall into low-risk categories, a weak positive correlation between the number of cases and the percentage of people who fall into medium-risk categories. We also determined that there is a moderate positive correlation between the number of deaths and the percentage of people in a high-risk category. Analysis of the residuals requires further study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

COVID-19 is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The common symptoms of COVID-19 include fever, dry cough, difficulty in breathing, etc. While most cases result in mild to moderate symptoms, some progress to pneumonia, multi-organ failure, and even death. The COVID-19 pandemic makes up a global health shock, with a death toll of over 6.2 million and over 500 million people reported ill by 29 March 2022. The rapidly increasing number of COVID-19 cases represents an unprecedented public health challenge for federal, state, and local authorities. Given the need to control the spread of the virus, the evolving guidelines for front-line providers, institution and business closures, and social distancing have been highly disruptive to essential healthcare, economic activities, and social services, leading to a significant increase in individual and community stress.

As the number of cases of COVID-19 increases, so does the associated anxiety. Disaster situations and traumatic events overwhelm our ability to cope. While this long-term disease is still ongoing, and we still have no idea when this will end, it is possible to mitigate some consequences by measuring communities' vulnerability and resilience. Community resilience can play a significant role in coping with shocks. However, it is an ambiguous concept, hard to define and measure. We define it as a complex and dialogical process in which communities create, develop and engage their resources to cope with shocks and their consequent uncertainty [23].

1.1 Background

Coronavirus disease was first reported as an outbreak of respiratory illness cases in Wuhan City, Hubei Province, China, in late December 2019. Later, WHO named the novel coronavirus disease COVID-19 in February 2020. In addition, the United States reported its first confirmed case of the novel coronavirus in January 2020. This virus has many potential naturals, intermediate, and final hosts (Fig. 1.1); due to these characteristics, there is a great challenge in preventing and treating the virus infection. The high infectivity and transmissibility make this disease a pandemic. And within a few months, researchers, manufacturers, organizations, media, and governments worldwide collaborate to gather and evaluate the COVID-19 data to predict and slow down the pandemic at any cost. WHO published healthcare, technical, preparedness and response, and social guidelines and then advised countries on responding suitably in the situations. Then, the first U.S. Food and Drug Administration (FDA) approved vaccine was available in August 2021, which was a milestone against the pandemic. Preventive measures for COVID-19 also include testing, isolating, maintaining social distancing, washing hands frequently, and avoiding touching the mouth, nose, and face.

1.2 Problem Statement

In this thesis, we focus on the community resilience of Idaho counties regarding COVID19 confirmed cases and confirmed deaths. Our objective is to use linear regression to identify the relationship between the community resilience data and the COVID-19 case data for 44 Idaho counties. An understanding of these relationships has the potential to predict and forecast how communities will respond to an epidemic.

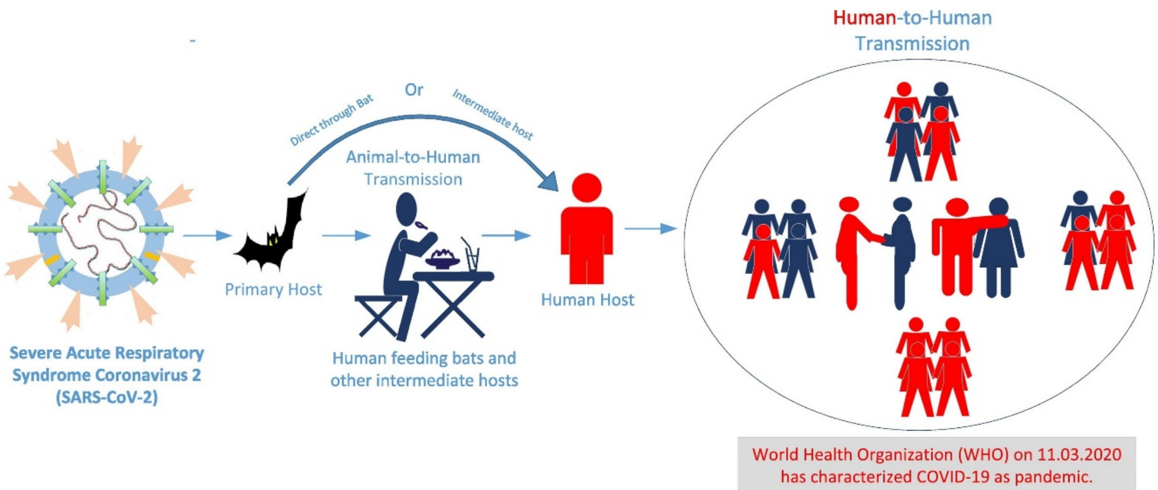


Figure 1.1 Transmission of COVID-19 [40]

Our research objectives focus on the following questions.

1. How is resilience defined?
2. What is community disaster resilience, and how can it be incorporated in the COVID19 pandemic in particular?
3. How reliable is the proposed CRE tool, available through the U.S. Census Bureau, as a quantitative measure of how well counties in Idaho responded to the COVID-19 pandemic?

CHAPTER 2: COMMUNITY DISASTER RESILIENCE

In this chapter, we review the concepts of community disaster resilience to form a better understanding of community resilience for a recent worldwide disaster, the COVID-19 pandemic. In addition, this chapter provides a theoretical foundation for developing the conceptual framework and relevance for measuring resilience discussed in Chapter 3.

2.1 Resilience

A system is usually designed to behave in a certain way under normal circumstances. However, when disturbed from equilibrium by a disruptive event, the performance of the system will deviate from its design level. The resilience of the system is its ability to reduce both the magnitude and duration of the deviation as efficiently as possible to its usual targeted system performance levels. Fig. 2.1 shows how a system returns to its normal equilibrium position after the disturbance occurs [36].

The term “resilience” was initially used in physics and mathematics to describe the capacity of a material or system to return to equilibrium after a displacement. Today, resilience has emerged in many wide-ranging diverse disciplines, including hazards, ecology, psychology, sociology, geography, economics, urban planning, and public health [28, 27]. As a concept, resilience thinking can be found anywhere from self-help guides on coping with hardships to major international agendas on reducing impacts from climatic change and is defined in various ways depending on the discipline. However, when applied to people and their environments, “resilience” is fundamentally a metaphor.

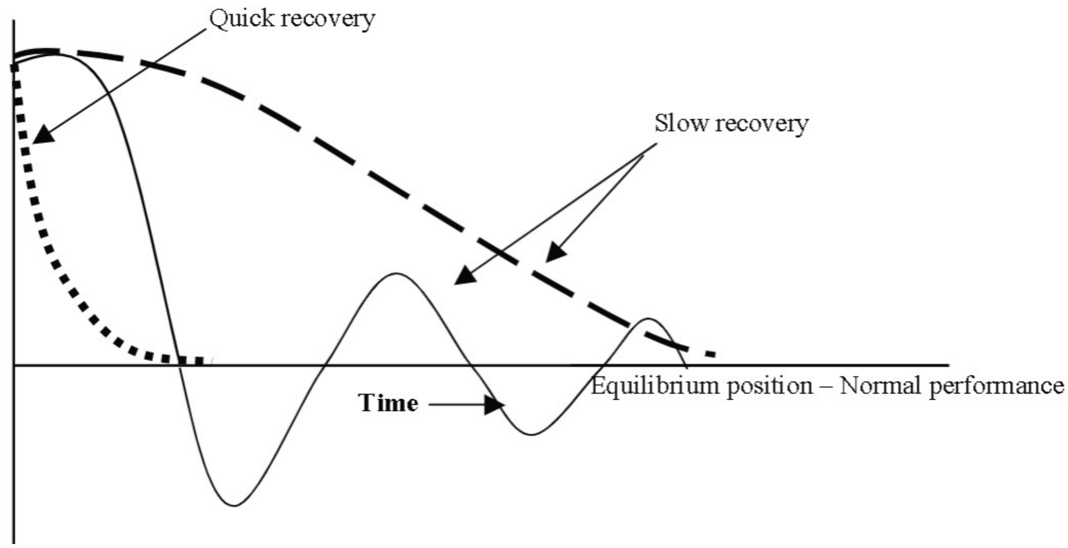


Figure 2.1 How a system returns to normal performance equilibrium position [36]

In recent years, the term ’resilience’ has gained increasing attention in the field of hazards and disasters. However, with its growing use in increasingly diverse areas, scholarly and policy prominence has come a fair amount of conceptual confusion and misapplication. So then, what does ‘resilience’ actually mean? Our primary focus is to identify key concepts and explores the relevance of resilience for disaster planning for communities. Resilience is a process linking a set of adaptive capacities to a positive trajectory of functioning and adaptation after a disturbance [28]. Government, industry, and charitable organizations have an increasing focus on programs intended to support community resilience to disasters. Most definitions of resilience share four common elements: context; disturbance; capacity; and reaction, shown in Fig 2.2 [17].

But has consensus been reached on what defines ’community resilience’ and its core characteristics?

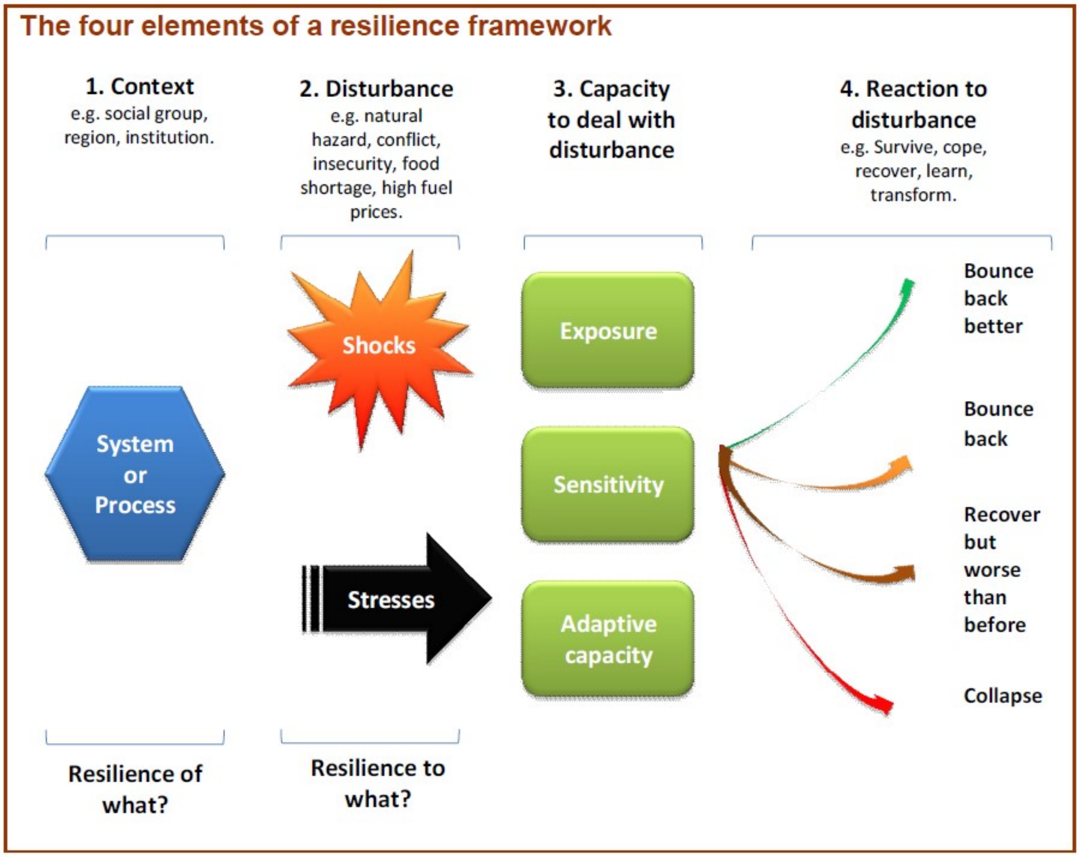


Figure 2.2 Four elements of resilience defined by the Department for International Development [17]

2.2 Community resilience

Communities are complex entities, and the challenges they face in this 21st century are getting complicated. For example, human-caused and natural disasters are more frequent and cost more modern lives [22]. In addition, factors like climate change, globalization, and increased urbanization can bring disaster-related risks to more significant numbers of people. Addressing these threats, we need to be well prepared for upcoming unknown disasters by gathering all the known experiences and taking proper actions to secure our human society.

Community resilience can play a vital role in coping with shocks. However, the concept of "community resilience" raises the same concerns as the concept of resilience per se but is further complicated by variation in the meaning of community. Typically, a community is an entity that has geographic boundaries and shared fate. Communities are composed of built, natural, social, and economic environments that complexly influence one another. Past writings on community resilience have described everything from grass-roots groups and neighborhoods to complex amalgams of formal institutions and sectors in larger geopolitical units. This description is not inappropriate, as resilience can be understood and addressed at different levels of analysis. However, discussions of community resilience often note that the "whole is more than the sum of its parts," meaning that a collection of individual resilience does not guarantee a resilient community [28].

The U.S. Census Bureau defines: "Community resilience is a measure of the capacity of individuals and households to absorb, endure, and recover from the health, social, and economic impacts of a disaster such as a hurricane or pandemic. When disasters occur, recovery depends on the community's ability to withstand the effects of the event"[5]. To define community resilience, Paton said it is important to identify the personal and community characteristics and processes that promote a capability to "bounce back" and effectively use physical and economic resources to aid recovery following exposure to hazard activity. Resilience should be conceptualized and managed in a contingent rather than a prescriptive manner. Researchers, planners, and emergency managers must acknowledge heterogeneity in community characteristics and perceptual processes and develop models that accommodate contingent relationships between hazard effects and community, cultural, geographical, and temporal factors within resilience models[35].

Community resilience is a process linking a set of networked adaptive capacities to a positive trajectory of functioning and adaptation in constituent populations after a disturbance. Community resilience describes the capability (or process) of a community adapting and functioning in the face of disturbance [28]. Communities and individuals harness local resources and expertise to help themselves in an emergency in a way that complements the response of the emergency services [6].

From a literature review of Patel and colleagues [34], three general types of definition were found:

1. **process** i.e. an ongoing process of change and adaptation
2. **absence of adverse effect** i.e. an ability to maintain stable functioning and
3. **range of attributes** i.e. a broad collection of response-related abilities

More recent studies tended to adopt the first type of definition 1. Community resilience is defined as “a reflection of people’s shared and unique capacities to manage and adaptively respond to the extraordinary demands on resources and the losses associated with disasters” [35] [28][12]. Furthermore, community resilience is defined as “a capability (or process) of a community adapting and functioning in the face of disturbance” [8].

The **absence of adverse effect** definition 2 uses the desired outcome of ’maintaining stable functioning’ as their basis. Contrasting to the first type of definition, Gibson [18] stated that “...resilience is not a process, it is not a management system standard, nor is it a consulting product. Instead, resilience is a demonstrable outcome of an organization's capability to cope with uncertainty and change in an often volatile environment. Resilience is thus a product of an organization's capabilities interacting with its environment.

This notion of community resilience as an outcome was adapted by others who noted the importance of explicitly identifying and strengthening abilities in a community, creating definition 3. An example of these definitions can be found in which community resilience is defined as "communities and individuals harnessing local resources and expertise to help themselves in an emergency, in a way that complements the response of the emergency services" [6]. This report suggests that primarily, community resilience has to do with having a responsive and collective action of local support to help the community after/during an incident. "A community's capacities, skills, and knowledge that allows the community to participate fully in the recovery from disasters" [28] is called community resilience. Additionally, definitions exist that blend one or more of these general definition types. In a recent review, Ostad Taghizadeh and colleagues [33] produced a definition that blended definition 2 and 3, now used by the United Nations Office for Disaster Risk Reduction (UNDDR): the "ability of a system, community, or society exposed to hazards to resist, absorb, accommodate to and recover from the effects of a hazard in a timely and efficient manner including through the preservation and restoration of its essential basic structures and functions."

Therefore, community resilience was an amorphous concept that different research groups understood and applied differently to different communities. In essence, community resilience can either be seen as an ongoing process of adaptation, the simple absence of adverse effects, the presence of a range of positive attributes, or a mixture of all three [34]. Unfortunately, we currently have no consensus on what exactly a culture or community should look like to be more resilient. Until we resolve this fundamental question, attempts to measure or enhance resilience will remain discordant and inefficient, while the academic literature will continue to be confused by papers assessing different concepts but using the same terminology. However, now that we have grown some concept of community resilience, we will address the question of what community resilience means in the context of disaster?

2.3 Community resilience in a disaster

The word disaster is used in diverse ways, primarily to refer to any sudden, unexpected or extraordinary misfortune, regardless of the number of people, region, country, or the entire world. In the disaster context, resilience is often treated as the simple inverse of fragility [36].

At present, definitions of community disaster resilience tend to either focus on specific aspects of the concept that may lead to overlooking some elements or tend towards all encompassing definitions that may be too complex to apply at the local level. It may be more appropriate to consider community resilience as a term for the range of elements that may be important for a community facing or recovering from a disaster. Resilience is now a key element of the United Nations International Strategy for Disaster Reduction(UNISDR), defining it as 'the ability of a system, community or society exposed to hazards to resist, absorb, accommodate to and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions...is determined by the degree to which the community has the necessary resources and is capable of organizing itself both before and during times of need' [27]. In conceptual terms, vulnerability and disaster resilience are closely related. Some researchers see vulnerability as the opposite of disaster resilience, while others view vulnerability as a risk factor and disaster resilience as the capacity to respond [24]. While labeling an individual or group of individuals as "vulnerable" seems to discourage peoples' efforts in dealing with disasters, the concept of community disaster resilience appears to be more proactive. It encourages collective efforts in a community to deal with disasters.

Community disaster resilience is a broader concept that encompasses a large part of the risk spectrum [42]. It emphasizes the community's capacities and how to strengthen them, and it places less emphasis on the factors which make the community vulnerable. Some researchers see resilience as "a multi-dimensional attribute that in its different forms contributes in various but equally important ways to disaster recovery" [11].

The literature review indicates that conceptual and methodological problems still exist concerning community disaster resilience that needs to be addressed. The concept of disaster resilience seems to be central to understanding the complex interactions within and across communities and how communities respond and function during disasters. Researchers agree that disaster resilience is the capacity or ability of people, a group of people, a community, or a society to continue functioning in the face of a disaster, the ability of a system to absorb, resist or deflect disaster impact and when impacted to relatively quickly recover and learn or adapt to future risks [25]. In general, we can define that community disaster resilience describes a community's intrinsic capacity to resist and recover from a disaster or disturbance, capacity to adapt to crises.

Developing community resilience helps disaster planners and community members to build traditional preparedness while promoting robust community systems and addressing the many factors contributing to health. In addition, a resilience approach adds features like building social communication and improving the community systems' everyday health and overall wellness.

2.4 Community resilience in the context of COVID-19

Worldwide, basic knowledge of COVID-19 across populations is now expected – including knowledge about COVID-19 symptoms. Available global data suggests that 64% of survey participants could correctly describe COVID-19 signs and symptoms [38]. Risk perception is a crucial driver of behaviors, and there is growing evidence that people's risk perception of COVID-19 infection is declining. People are becoming complacent; thus, risk perceptions are lowering. People feel less confident in what they can do to control the virus. People do recognize COVID-19 as a serious disease; however, they often feel COVID-19 is more of a threat to others: their friends and family, their community and country, than to themselves [30]. Understanding the transmission of this highly contagious disease, shown in Fig 1.1, is vital in the sense of control and preventing the spread of the virus, as well as reducing community transmission and being supportive of the affected people of our society.

People's behavior and their willingness to follow public health and social measures remain the most powerful weapons to stop spreading the virus. However, human behavior is complex. Therefore, it is crucial for risk management to understand people's changing perceptions and attitudes and the barriers and enablers influencing their ability and motivation to adopt and sustain positive health behaviors. Multiple efforts are made to collect, analyze, and use socio-behavioral evidence in response to this pandemic. Consequently, there is an unprecedented need to elevate community resilience to mitigate the impact of pandemics. People have enough knowledge about COVID-19 and the necessary preventive measures. However, as the situation continues, ’pandemic fatigue’ is occurring. This fatigue is likely to lead to a decrease in people’s motivation to follow recommended preventive behaviors and create a number of detrimental emotions, experiences, and perceptions [32]. However, community-led approaches are championed widely, resulting in increased trust and social cohesion and ultimately a reduction in the negative impacts of COVID-19.

Pandemic fatigue can be influenced by a variety of factors depending on the context. These factors include a decrease in risk perceptions related to the disease; an increase in the socio-economic and psychological impact of the crisis and restrictions; the urge for self-control and self-determination in a constantly changing and restricting environment; and the feeling of getting used to the situation. Self-efficacy is a vital driver of behavior change towards stronger community participation. In countries where people feel less confident in their ability to protect themselves, people are also less likely to practice preventive measures [20]· Social perceptions have consequences. They can hamper efforts to stop or slow the spread of COVID-19 and mitigate its impacts. Community engagement and participation have played a critical role in successful disease control and elimination campaigns in many countries [2], including outbreaks that occurred previous to COVID-19 [19], [4]. That is, people-centered and community-led approaches are championed widely, resulting in increased trust and social cohesion and ultimately a reduction in the negative impacts of COVID-19.

According to Giamberardino et al., 2020 [14], it took about 18 months for an infectious disease to spread throughout the world in the nineteenth century. It took less than 36 hours in recent years, which is shorter than the incubation period of most diseases. Also, nearly 400 million people go to another country or region every year, which undoubtedly accelerates the spread of the virus. The global outbreak of COVID-19 is a non-linear and dynamic process.

The COVID-19 pandemic is a disaster that combines a biological threat with various vulnerabilities, such as physical, social, and economic. Various attempts have been made in the last years to study the spread of epidemic diseases, and the main problem is to understand and describe the modalities of the geographic spread among nations with different community characteristics and health statuses and the effects of migration phenomena. Governments have taken health, social and economic measures to address the emergency and reduce the impact of the crisis on the most vulnerable. Most of the countries in the region have made notable efforts, considering their reduced fiscal space.

2.5 Summary

The concept of community resilience is widely used in the academic and policy literature, yet the meanings of the term differ. Nevertheless, these core elements have been consistently suggested as constituting community resilience as it applies to disasters: local knowledge, community networks and relationships, communication, health, governance and leadership, resources, economic investment, preparedness, and mental outlook. Further exploration of these individual elements may lead to a greater understanding of community resilience and how it can be measured and enhanced for the worldwide pandemic Covid-19. In the meantime, the attempt to define the concept of community resilience would be unhelpful if it obscures the main point- being resilient as a community when a disaster occurs.

The definitions and various concepts reviewed in this chapter provide a better understanding of resilience, community resilience and how they should be conceptualized and applied in the research for pandemic Covid-19. Only by understanding the character of a community adequately, i.e., their communication and collaboration skills, knowledge, needs, and gaps in understanding about COVID-19, can we achieve a community-driven response that will reduce the spread of the virus and create a disaster-resilient community.

CHAPTER 3: COMMUNITY RESILIENCE DATA

This chapter discusses the importance of creating a tool to measure community disaster resilience problems associated with creating an effective tool and highlights the applications of the tools in specific fields. Then it introduces a newly developed experimental estimate, Community Resilience Estimates (CRE) tool, developed by the U.S. Census Bureau. To understand this tool, we reviewed concepts of resilience, community resilience, community resilience for disaster, and COVID-19 from the literature to identify key elements that can be used to validate the tool.

Public health emergencies in the United States have been complex, frequent, and increasingly costly in the past decade. Emergencies are not always predictable, and adequate resources are not always available to prepare in advance when a new threat emerges [9]. Communities in Idaho rely on self-governance and state fund allocation to combat natural disasters, diseases, and everyday life. Improved readiness can mitigate the impact of disasters on at-risk populations and the economic burden on individuals, households, and governments. However, often it can be difficult to fully understand which areas are most at risk for these unexpected events. Therefore, it is vital to find a metric to identify at-risk populations and adequately allocate resources to those communities. Identifying the generic principles, i.e., risk factors, that support resilience can facilitate the development of models capable of use with diverse communities and hazards as well as provide emergency managers with a framework within which they can develop suitable strategies tailored to the specific context (e.g., a mix of hazard and community characteristics). Resilience variables must have predictive validity independent of the community or hazard under investigation to be useful for emergency planning.

The U.S. Census Bureau has completed an analysis that classifies counties based on their community resilience. The Bureau is the nation's leading provider of quality data with advanced statistical capabilities about its people and economy and is uniquely positioned to provide the most accurate and timely measures for an individually focused community resilience indicator. It uses detailed demographic and economic data about individuals to build these estimates, with lower sampling error, compared to other institutions. In addition, the Bureau can adapt the estimates as needed to incorporate the latest and most relevant data. As a result, the Bureau produces estimates with the most granularity, highest statistical quality, and broadest coverage while still protecting privacy.

The U.S. Census Bureau has created a tool to help measure the degree of a community's resilience in the face of disasters and other emergencies in June 2020. This newly developed experimental estimate, Community Resilience Estimates (CRE), is a resilience measure that identifies a community's ability to endure, respond and recover from the impact of disasters. This tool can be used for any purpose where specific risk factors are helpful at low levels of geography, i.e., by county. The CRE tool estimates community resilience to disasters using small area estimation (SAE) techniques to combine data from several sources and produce high-quality estimates. These techniques are flexible and can easily be modified for a broad range of uses (hurricanes, tornadoes, floods, economic recovery, etc.). Resilience to a disaster is partly determined by the vulnerabilities within a community. The Bureau designed population estimates based on individual and household-level risk factors to measure these vulnerabilities and construct the community resilience estimates.

Community resilience does not necessarily improve when they are able to cope with disasters and their aftermath alone; rather, it improves when public health systems strengthen protective factors such as social networks that aid people and communities to manage, adapt, and ultimately recover well from disasters. Indeed, a good measure of resilience implies that communities' day-to-day health and wellbeing can help reduce the negative impacts of disasters and being a member of multiple social networks or groups can affect health and wellbeing, particularly during times of change. Therefore, strengthening community resilience in the months and years will require a whole system approach working with different sectors. The updated version of the CRE tool (updated on August 2021) is produced using the information on individuals and households from the 2019 American Community Survey (ACS) and the Census Bureau's Population Estimates Program (PEP) to identify the population most at risk of the Coronavirus pandemic.

The ACS is a nationally representative survey with data on the characteristics of the U.S. population. The sample is selected from all counties and county-equivalents and has a sample size of about 3.5 million housing units each year. It is the premier source for detailed population and housing information about the U.S. and communities within it. The estimates analyze the individual, and household level restricted ACS microdata to determine the number of individual risk factors. The PEP produces and publishes estimates of the population living at a given time within a geographic entity in the U.S.

and Puerto Rico. The estimates use population data from the PEP by tract, age group, race and ethnicity, and sex. Once the weighted estimates are tabulated, small area modeling techniques are utilized to create the CRE tool.

Resilience is measured using risk factors. Risk factors are determined by examining the following ten demographic, socioeconomic, and housing characteristics in the ACS. Risk factors are binary components that add up to 10 possible risks. For household-level variables, if the household meets the criteria for the risk flag, every individual in the household receives that risk flag. Risk Factors (R.F.) for Households (H.H.) and Individuals (I) are:

- R.F. 1: Income-to-Poverty Ratio (IPR) < 130 percent (H.H.)
- R.F. 2: Single or zero caregiver household - only one or no individuals living in the household age 18-64 (H.H.)
- R.F. 3: Unit-level crowding defined as > 0.75 persons per room (HH)
- R.F. 4: Communication barrier defined as either -Limited English speaking households (H.H.) or; no one in the household over the age of 16 with a high school diploma (H.H.)
- R.F. 5: No one in the household is employed full-time, year-round. The flag is not applied if all household residents are aged 65 years or older (H.H.)
- R.F. 6: Disability posing constraint to significant life activity - Persons who report having any one of the six disability types (I): hearing difficulty, vision difficulty, cognitive difficulty, ambulatory difficulty, self-care difficulty, and independent living difficulty
- R.F. 7: No health insurance coverage (I)

- R.F. 8: Being aged 65 years or older (I)
- R.F. 9: Households without a vehicle (H.H.)
- R.F. 10: Households without broadband Internet access (H.H.)

Note that risk factor four is not double flagged. For example, if a household is linguistically isolated and no one over the age of 16 has attained a high school diploma or more education, those in that household are only flagged once. A ”Limited English speaking household” is one in which no member 14 years old and over (1) speaks only English at home or (2) speaks a language other than English at home and speaks English ”Very well.”

The result is an index that produces aggregate-level (tract, county, and state) small area estimates: the CRE. The CRE provides an estimation of the total number of people living in a community by the number of risk factors. The estimates are categorized into three groups: 0 risk factors (Low risk), 1-2 risk factors (Medium risk), and three or more risk factors (High risk). Individuals with three or more risk factors – from health and income to age and living conditions – are considered high risk. Likewise, communities are at high risk if at least 30% of their population has three or more risk factors.

Communities with more risk factors are considered less resilient to disasters and, therefore, important to identify. By finding these counties, lawmakers can better allocate resources and provide targeted help to the most needy. And this is especially important in the recent pandemic of COVID-19. From the beginning of the pandemic, the adverse effects of COVID-19 have been strongly affecting individuals and households. This tool maps the risk assessment of local populations down to the neighborhood level and allows national and community leaders to respond more efficiently to emergencies. In addition, stakeholders can use CRE and other tools to help combat the current crisis and plan for future health and weather-related disasters.

CHAPTER 4: HEALTH EFFECT

This chapter looks at the global conditions due to the COVID-19 pandemic, the affected and death numbers of different countries due to other mortality methods, discuss the new variants, vaccines, U.S. situation, and then focuses on Idaho number. The U.S. health and risk factors levels are compared to the number of COVID-19 cases and deaths.

4.1 Global health effect

The COVID-19 pandemic makes up a global health shock, with an official death toll of over 5.6 million and over 350 million people reported ill by January 23, 2022 [3]. Once the pandemic took hold, the world witnessed the devastating collapse of health systems when the first wave of coronavirus attacked. There was so much chaos from the beginning of this novel coronavirus, starting from denying the new virus attack, ignoring the symptoms, finding an effective testing procedure, developing and choosing safe and efficient vaccines from dozens of vaccines through clinical trials, how the virus spread and fighting with new variants, which group of people gets affected most, whether masks, lockdown, isolation and social distancing work, which vaccine to take, etc. The ultimate impact of the pandemic directs people to uncertainty and ignorance, perplexing thinking, and reflexive panic; on top of everything, unemployment, insufficient medical care, and the misery of losing loved ones led worldwide people to protest against lockdowns, vaccines, and government [45], [10], [7].

A question central to the COVID-19 pandemic is why the COVID-19 mortality rate varies so greatly across countries. One reason is that COVID-19 testing methods differ from country to country. Another reason discovered recently is that the mortality methods vary from one research approach to another. For example, the Lancet, a peer-reviewed journal, measured the COVID-19 death number by excess mortality and found that 18.2 million people may have died globally from COVID-19, three times the official total, by December 2021 [43]. Another excess-mortality database maintained by The Economist also estimates global excess mortality & puts the figure above 20.3 million COVID-19 death (with 95% confidence interval) by March 27, 2022.

The numbers vary so broadly that accounting for them entirely changes the picture of the experience of individual nations but the whole world, rearranging everything about our gathered knowledge [1]. Based on the crude count of official death reports, North America and Europe have death counts almost eight times as high as Asia and 12 times as high as Africa. South America's death toll is ten times as high as Asia and 15 times as high as Africa. The excess-mortality data tells a different story. There is still a clear continent-by-continent pattern, but the gaps between them are much smaller, making the experiences of other parts of the world much less distinct and telling a universal story about the devastation wrought by this once-in-a-century contagion. With this view, Oceania, Europe, and North America were among the best at preventing deaths among the old, and they were better at protecting their elderly than Africa and South Asia. By almost any metric, Oceania does better, and the estimate of excess deaths among the elderly in New Zealand is zero, according to The Economist. In the country-by-country data, the world's worst pandemic, according to the data, has been in Bulgaria, followed by Serbia, North Macedonia, and Russia, then Lithuania, Bosnia, Belarus, Georgia, Romania, and Sudan [21], [43].

The coronavirus has mutated through time and causes different variants. The delta variant is considered one of the most contagious, causes severe illness, and the omicron is more infectious [16]. Fortunately, vaccination effectively prevents serious illness, hospitalization, and death from COVID-19. In the beginning, fewer people were vaccinated, which meant many people were vulnerable [37]. Other factors have impacted whether COVID-19 cases are increasing or declining in particular locations. These factors include the effectiveness of vaccines over time, human behavior, infection prevention policies, mutation of the virus, and the number of vulnerable people. We are now in this overall stable situation after many trial and error procedures of our government and health guidelines [41].

4.2 U.S. health and community resilience

Health includes physical, behavioral, social, and well-being, which is a big part of overall resilience [44]. In many ways, health is a crucial foundation of resilience because almost everything we do to prepare for disaster and protect infrastructure is ultimately in the interest of preserving human health and welfare. In this pandemic, vaccines are serving the primary purpose, which is to prevent severe illness, hospitalizations, and death, critically reducing the load on the overburdened healthcare system, and overall improving resilience [13]. Since the pandemic began, U.S. states and territories have used different approaches to reporting data about COVID-19 cases, deaths, tests, and vaccines. The lack of uniformity has complicated efforts to track COVID-19 in near real-time. As of March 2022, 81.6 million COVID-19 cases and 1 million deaths have been reported, and more than 45 million COVID-19 vaccine doses have been administered in the U.S. [15], translating to 62% of the total population being fully vaccinated. The vast majority of individuals have access to a free vaccine. However, many have chosen not to accept one.

Table 4.1 Death rates for US states

US States with highest COVID-19 death rate	COVID-19 death rate per 100k population
Mississipi Arizona Alabama	416.5 396.8 393.2
US states with lowest COVID-19 death rate	COVID-19 death rate per 100k population
Hawai Vermont Utah	97 98.9 146.9
COVID-19 death rate in Idaho	274.9
*till March 29, 2022	

Table 4.2 Case rates for US states

US States with highest COVID-19 case rate	COVID-19 case rate per 100k population
Rhode Island Alaska North Dakota	34054 32590 31550
US states with lowest COVID-19 case rate	COVID-19 case rate per 100k population
Oregon Maryland Hawaii	16823 1686 17183
Total COVID-19 case rate in Idaho	24903
*till March 29, 2022	

The COVID-19 death rate is the number of fatalities from the disease per 100,000 people. Table 4.1, presents states with the highest and lowest death rates. Furthermore, Table 4.2 presents states with Rhode Island, Alaska, and North Dakota have the highest and Oregon, Maryland, Hawaii have the lowest COVID-19 case rates. However, rates are not the best measure when comparing data. According to the World Health Organization, ”any attempt to capture a single measure of fatality in a population will fail to account for the underlying heterogeneities between different risk groups, and the important bias that occurs due to their different distributions within and between populations” [31]. For instance, data shows COVID-19 death rate is higher among Latinos, Blacks, and Indigenous Americans than among non-Hispanic whites, shown in Figure 4.1. Also, rural counties are far more likely to have greater proportions of high-risk populations. Thirty percent of all rural counties are high-risk compared to 14% of all urban counties. Some of the factors associated with more high-risk communities include low income, especially in rural communities, a greater proportion of single mothers, a majority Black and Hispanic population, and a significant proportion of residents 65 and older are at considerable risk for infection and developing severe illness.

Counties with at least 30% of their population with three or more risk factors are considered high risk. The simple graph in Figure 4.2 shows the high risk counties of the U.S., provided by the U.S. Census Bureau. According to the Bureau, the CRE data indicates that Florida, Nevada, Texas, New York, New Mexico, and Arizona have a significant percentage of high-risk populations; around 30% of their population is highly vulnerable. For instance, Florida is the most at-risk state; 31% of the population is at high risk. Followed by Nevada, New York, and Texas have the highest percentage of high risk people. However, Texas has a significant number of high risk counties; for example, Real County, Texas, has 47% of its population with high risk factors, making its population vulnerable to pandemics. On the other hand, more than 60% of people in California and Hawaii have 1 or 2 risk factors, which means most people are at medium risk. Vermont, New Hemisphere, and Maine have more than 30% of people having no or low risks, pointing that they are less vulnerable to disaster. Countywise, Salt Lake County, Utah, has 66% zero or low risk people, having the most community resilience in the entire nation.

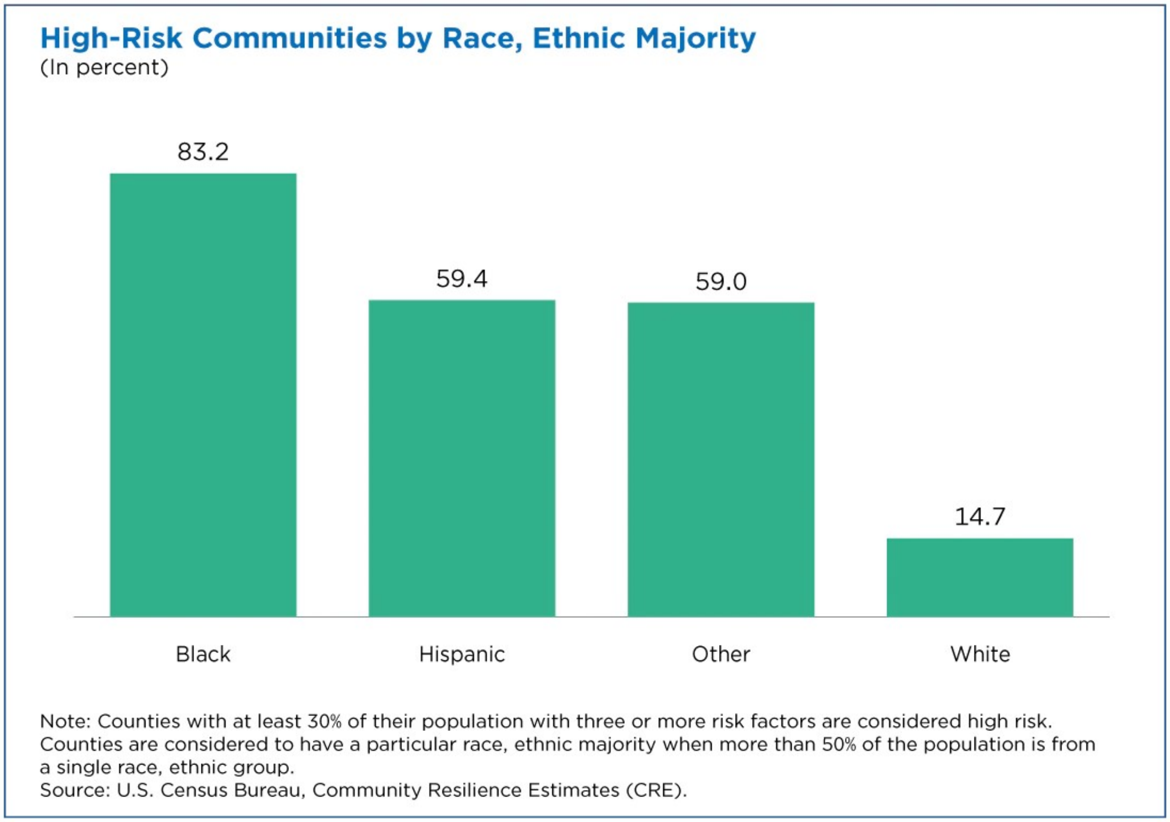


Figure 4.1 High risk communities of US by race, ethnicity majority

4.3 COVID-19 and Idaho resilience data

The first case relating to the COVID-19 pandemic in Idaho was confirmed on March 13, 2020, when a woman from Ada County tested positive, and Idaho's first COVID-19 deaths were on March 26 [39]. As of March 29, 2022, there have been 443,792 confirmed cases and 4,870 deaths within Idaho, while 930,380 people have been fully vaccinated (not including booster doses). To compare with other states, on average, counties in Idaho have 23% of its residents with high risk, 50% with medium risk, and 27% with low risk, see Figure 4.3.

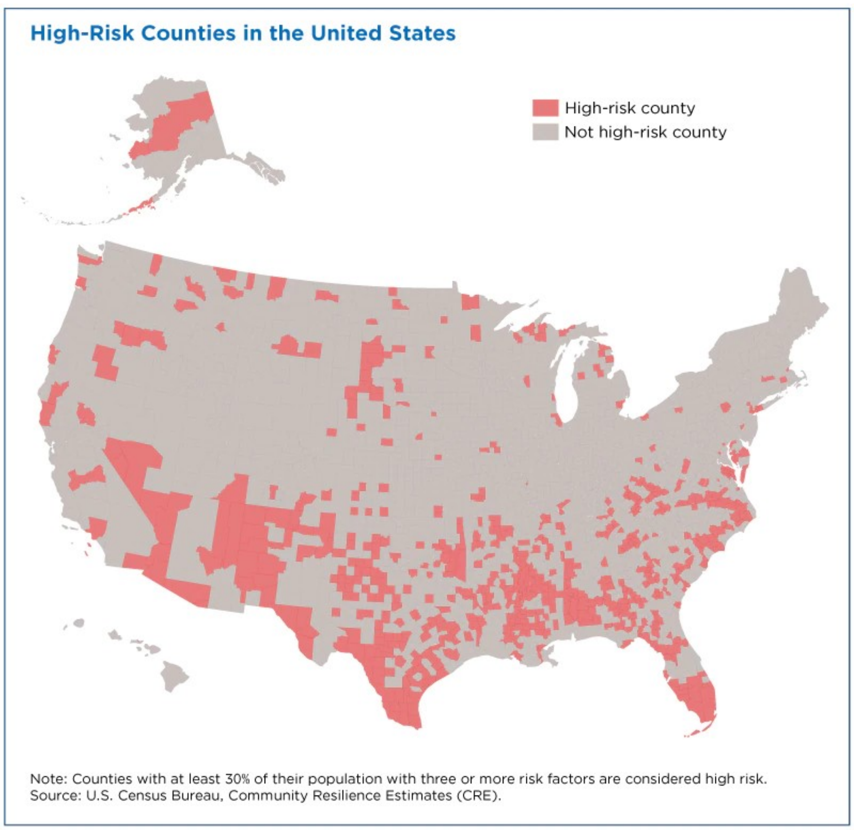


Figure 4.2: Community resilience of US by Census Bureau

Figure 4.3 shows the community resilience estimates for the county-level state profile for Idaho, sorted by the greatest percentage of residents in the county with three or more risk factors to least. According to the CRE tool, Clearwater is a high risk county with 33% of its residents at high risk, following Washington and Idaho counties having around 30% high risk people. In contrast, Teton County has 37% low risk people, being the most resilient in Idaho. Franklin, Jefferson, and Butte counties also keep their risk low. Madison, Jerome, Caribou, and Canyon counties have around 50% to 60% of medium risk populations.

Looking at the percent of death cases in Idaho counties, Clark (43% at medium risk) and Camas (54% at medium risk) counties have zero COVID-19 death cases, and Lewis (44.15% at medium risk) and Shoshone (45% at medium risk) counties have the highest percent of COVID-19 death cases, even though most people are at medium risk.

Table 4.3 Confirmed case rates for Idaho counties

Idaho counties with highest COVID-19 confirmed case rate	COVID-19 confirmed case rate per 100k population
Madison County Lewis County Clearwater County	269.25 236.87 229.9
Idaho counties with lowest COVID-19 confirmed case rate	COVID-19 confirmed case rate per 100k population
Camas County Custer County Clark County	99.37 104.4 110.3
*till March 29, 2022	

Table 4.4 Idaho counties sorted by high risk

Greatest percentage of high risk counties	percentage of residents with high risk
Clearwater County Washington County Idaho County	32.56% 30.26% 29.68%
Least percentage of high risk counties	
Latah County Jefferson County Caribou County	15% 16.63% 16.86%

4.4 Summary

The worldwide COVID-19 pandemic illuminated the fact that communities respond differently to disasters. For example, greater engagement in health-promoting behaviors may promote resilience in the face of infectious diseases like COVID-19 and prevent chronic diseases, including respiratory disease and diabetes. All countries can improve their readiness as we continue building a culture of preparedness. Collaborative efforts of the government and community are crucial to the success of the response to and recovery from public health emergencies and the resilience of a country. Because community resilience research is an emerging field, the indicators and indices of disaster resilience are sufficient with a wide range of tools that claim to measure disaster resilience, and our concern is to verify the CRE tool in this case.

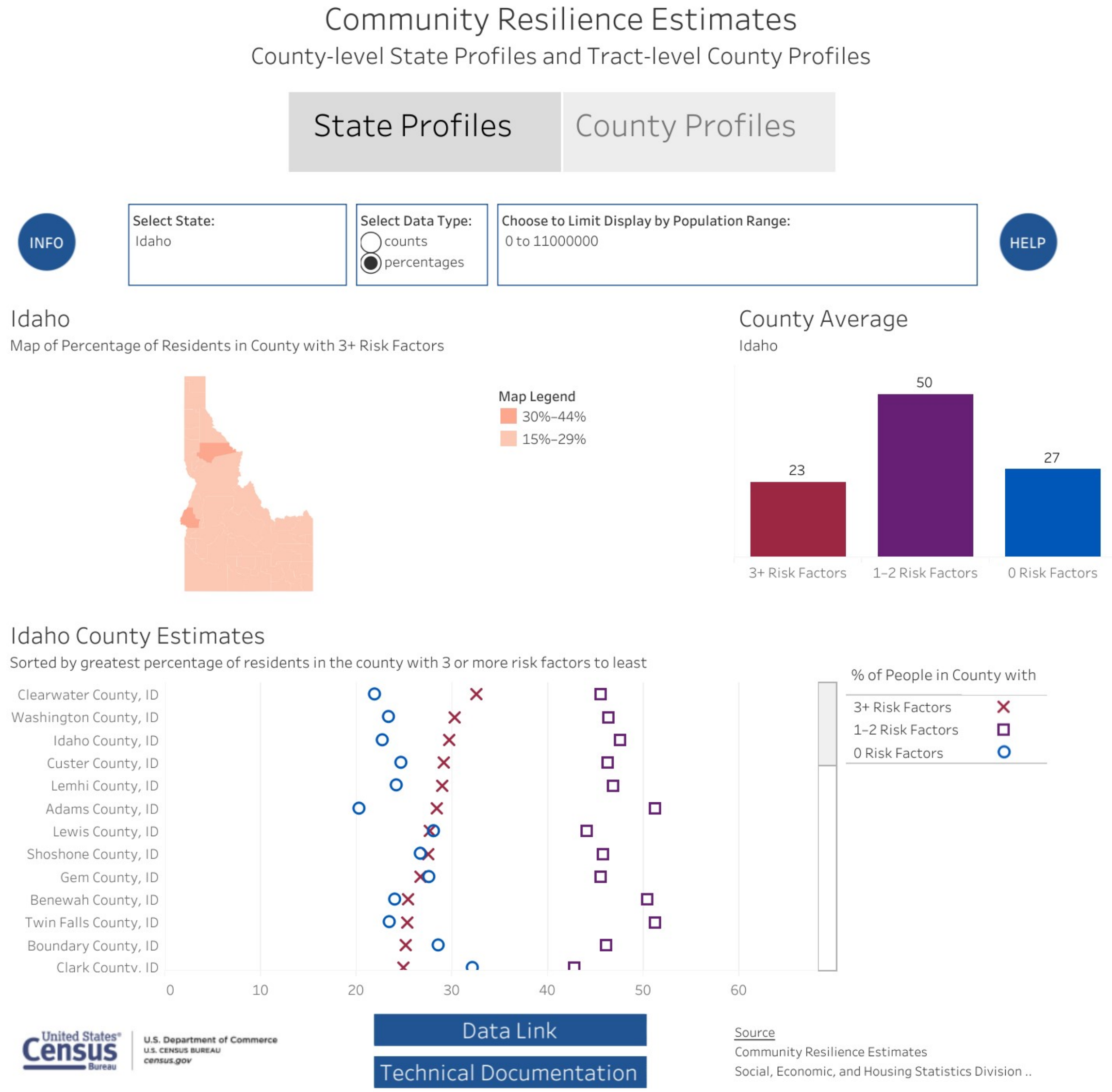


Figure 4.3 County-level community resilience estimates of Idaho

CHAPTER 5: REGRESSION ANALYSIS

This chapter uses a simple linear regression model to quantify the relationship between COVID-19 cases and deaths, released by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [15], and the resilience indices CRE tool from the U.S. Census Bureau CRE tool (census.gov). The COVID-19 data are cumulative values from March 13, 2020, until January 17, 2022.

5.1 Simple Linear Regression

In this research, we choose simple linear regression analysis because it captures the straightforward relationship between two variables. We will use it to determine if the number of COVID-19 cases and deaths depend on Idaho counties risk factors.

Regression can be restrictive because it relies on a fixed set of parameters β_0 and β_1 and assumes a linear relationship between variables, i.e., $y = \beta_0 + \beta_1 x$. An important objective of simple linear regression analysis is to estimate the unknown parameters β_0 (intercept) and β_1 (slope) in the regression model. This process is called fitting the model to the data. We use the Ordinary Least Square (OLS) method with our data to find estimates of β_0 and β_1, denoted by $\hat{\beta}_0$ and $\hat{\beta}_1$ respectively.

To understand the effectiveness of the fitted linear model, we present summary statistics available through the *summary()* command in R. These statistics give us detailed information on the model's performance and coefficients, including standard errors, t-statistics, p-values, and the F-test results. However, we usually cannot detect departures from the underlying assumptions in our model by examining the standard summary statistics, such as t or F statistics or R^2. These are global model properties, and as such, they do not ensure model adequacy. Therefore, we also check the adequacy of our simple linear model.

Regression analysis typically includes model adequacy checking, where the appropriateness of the model is studied, and the quality of the fit is ascertained. Through such analyses, the usefulness of the regression model may be determined. The outcome of adequacy checking may indicate either the model is reasonable or that the original fit must be modified. Thus, regression analysis is an iterative process in which data lead to a model, and a model's fit to the data is produced. A regression model does not imply a cause-and-effect relationship between the variables. To establish the causality, the relationship between the regressors and the response must have a basis outside the sample data- for example, the relationship may be suggested by theoretical considerations. Finally, it is crucial to remember that regression analysis is part of a broader data-analytic approach to problem-solving. That is, the regression equation itself may not be the primary objective of the study. Instead, it is usually more important to gain insight and understanding concerning the system generating the data [26].

5.1.1 Assumptions of a simple linear regression model

If the following assumptions are validated, the simple linear model can acceptably represent the data:

1. The relationship between the response y and the regressor x is linear.
2. The residuals, or errors $\varepsilon = \hat{y} - \hat{\beta}_0 - \hat{\beta}_1 x$ are normally distributed with zero mean.

3. The residuals have a constant variance.

4. The residuals are independent.

Therefore, our regression analysis involves assessing the validity of these assumptions to understand the model's adequacy. Model inadequacies have potentially severe consequences. In particular, gross violations of the assumptions may yield an unstable model in the sense that a different sample could lead to a totally different model with opposite conclusions.

1. Linear relationship between x & y

The first assumption of linear regression is a linear relationship between the independent variable, x, and the independent variable, y. The easiest way to detect this assumption is to create a scatter plot of x vs. y. This allows us to visually see if there is a linear relationship between the two variables. If it looks like the points in the plot could fall along a straight line, then there exists some linear relationship between the two variables, and this assumption is met. Here we run the regression equation and estimate the parameters $\hat{\beta}_0$ and $\hat{\beta}_1$.

After creating a scatter plot, if we see no linear association between the two variables, there is some option to fix this. We can apply a nonlinear transformation to the independent and/or dependent variable by taking the log, the square root, or other data transformations. For example, if the plot of x vs. y has a quadratic shape, it might be possible to add X^2 as an additional independent variable in the model.

2. Residuals are Normally Distributed

The *residuals* play a key role in evaluating model adequacy. Residuals can be viewed as the observed values of the model errors. To check the constant variance and uncorrelated errors assumption, we must first ask ourselves if the residuals look like a random sample from a normal distribution with these properties. After estimating the parameters, the residuals (e) are calculated as the difference between the observed value (y) and the corresponding fitted values ($\hat{y}$); i.e., the ith residual is, $e_i = y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)$. There are two common ways to check if the residuals are normal:

1. Check the assumption visually using Q-Q plots. If the points on the plot roughly form a straight diagonal line, then the normality assumption is met.
2. Check assumptions using formal statistical tests like Shapiro-Wilk, Kolmogorov Smirnov, Jarque-Barre, or D'Agostino-Pearson. However, these tests are sensitive to large sample sizes – that is, they often conclude that the residuals are not normal when the sample size is large. This is why it is often easier to use graphical methods like a Q-Q plot to check this assumption.

However, small departures from the normality assumption do not affect the model greatly, but gross non-normality is potentially more serious as the t or F statistics and confidence and prediction intervals depend on the normality assumption. If the normality assumption is violated, we have a few options:

- Verify that any outliers are not significantly impacting the distribution. If outliers are present, make sure that they are actual values, not some data entry errors.
- If the errors come from a distribution with thicker or heavier tails than the normal, other estimation techniques should be considered, i.e., robust regression methods.

- Apply a nonlinear transformation to the independent and/or dependent variable. Common examples include taking the log, the square root, or the reciprocal of the independent and/or dependent variable.

3. Residuals have constant variance

It is an assumption of linear regression that residuals have constant variance at every level of fitted values $\hat{y}$, or *homoscedasticity* occurs. The simplest way to detect non-constant variance is by creating a residual plot against a fitted value ($\hat{y}$) of the model. If the plot resembles that the residuals can be contained in a horizontal band, then there are no obvious model defects. This verifies the assumption that the residuals are randomly distributed and have constant variance. When this is not the case, the residuals are said to suffer from *heteroscedasticity*. Heteroscedasticity indicates that the variance is not constant or the true relationship between x and y is not linear.

There are some ways to fix heteroscedasticity:

- Transform the regressor and/or response variable, i.e., apply log or inverse tarnsformation.
- Use weighted regression.

4. Residual terms are independent

The linear regression model assumes that residuals are independent. This is relevant mainly when working with time-series data. However, as our data is not time-ordered, this assumption check is unnecessary.

5.1.2 Lack of fit of the model

The formal statistical test for the lack of fit of a regression model assumes that the normality, independence, and constant-variance requirements for residuals are met, and only the linear relationship assumption is in doubt. For example, if there is a linear relationship between x and y, we say there is no lack of fit in the simple linear regression model. Otherwise, the true relationship could be quadratic, for example. In this case, what we claim to be a random error could be a systematic departure due to not fitting enough terms. Data transformation of a new model is needed to fix these types of errors.

5.2 Analysis of community resilience data

Estimation and hypothesis testing are complementary inferential processes of a regression model. A hypothesis test is used to determine whether or not a treatment has an effect, while estimation is used to determine how much effect. For example, using the Ordinary Least Squares Method (OLS), we estimate the parameter β_1. Once we are done with estimation, we need to do hypothesis testing to make inferences about the population. Thus, we would like to know how close $\hat{\beta}$ (estimated β) is to true β or how close the variance of $\hat{\beta}$ is to the true variance. If all assumptions of the linear regression are satisfied, OLS gives us the best linear unbiased estimates. In this section, we estimate the coefficient, test of hypothesis, and verify the assumptions of the linear model for six individual cases.

The fitted simple linear regression model is:

$$\hat{y}_{j,k}(i) = \hat{\beta}_{0j,k} + \hat{\beta}_{1j,k} x_k(i) \tag{5.1}$$

where, i represents a county in Idaho (i = 1,2,.....44), j represents number of confirmed cases (j = 1) or confirmed deaths (j = 2) adjusted for population, k indicates low (k = 0), medium (k = 1), or high (k = 2) risk factors. For example, $Y_1(i)$ represents percent of

COVID-19 confirmed cases in Idaho County i, and $X_2(i)$ represents percent of individuals with high risk factors. Values for $Y_j(i)$ were obtained from Johns Hopkins Coronavirus Resource Center [3], while values for the independent variables $X_k(i)$ were obtained from the CRE tool offered by census.gov. The COVID-19 data are cumulative values from March 13,2020 to March 29, 2022 during which time there has been over 500 million confirmed cases and 6.2 million confirmed deaths in the entire state of Idaho.

5.2.1 Case I: confirmed cases vs. low risk in Idaho counties

In this case, our predictor variable is the percent of individuals with low-risk factors in Idaho counties, and the response variable is the percent of confirmed cases in corresponding counties. Our goal is to find if there exists any linear relationship.

First, we created a scatter plot in Figure 5.1 by plotting the percentage of individuals with low-risk factor $x_0(i)$ on the x-axis and the percent of confirmed covid cases $y_1(i)$ on the y-axis. There are 44 points in this plot for 44 Idaho counties. The plot in Figure 5.1 indicates a downhill pattern as we move from left to right, implying a negative relationship between the percent of low risk people and the number of COVID-19 cases.

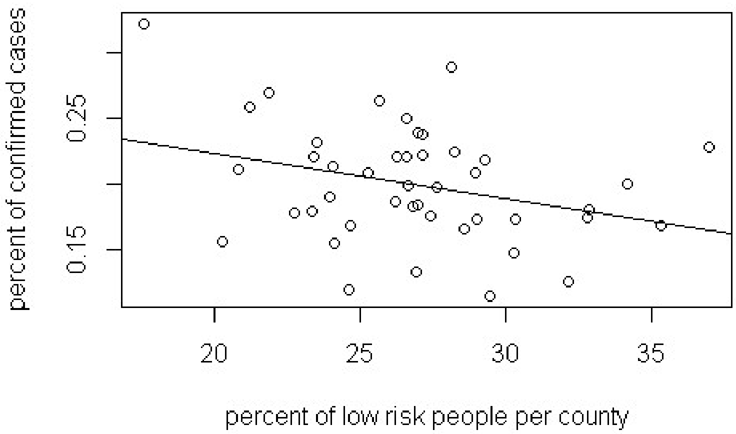


Figure 5.1 Case I: COVID-19 confirmed cases vs low risk for 44 Idaho counties

We then use command lm() in R to fit linear models to data using OLS. The result (Table 5.1) contains coefficients for $\hat{y}_1(i) = \hat{\beta}_{01,0} + \hat{\beta}_{11,0} x_0(i)$, where $\hat{\beta}_{01,0}$ is the x-intercept and $\hat{\beta}_1^{1,0}$ is the slope. From this equation, we could interpret that for each unit increasing rate of individuals with low-risk factors, the rate of confirmed cases decreased by 0.03442.

We use the following null and alternative hypotheses for this t-test:

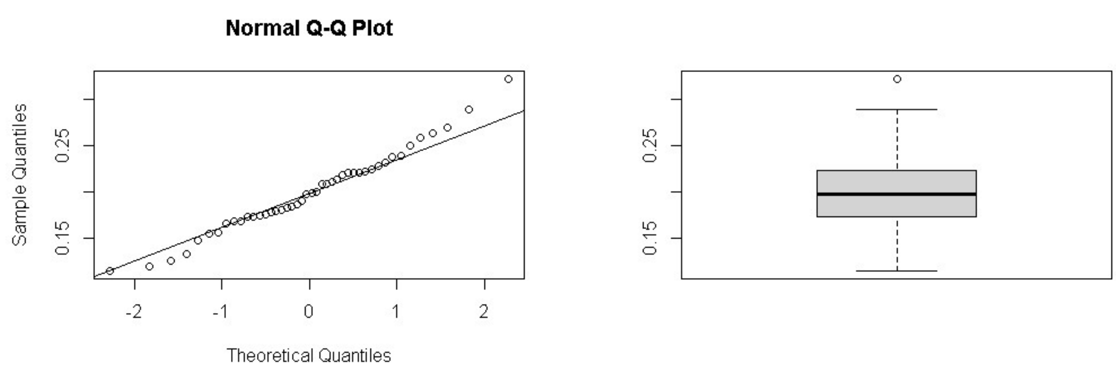
$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0.$$

Our t-statistic value is −2.147 from Table 5.1, and its corresponding p-value is 0.0376 (< significant level, $\alpha = 0.05$). In practice, any p-value below 0.05 is usually deemed as significant. From these results, we can conclude that there is strong evidence that the coefficients in this model are not zero, meaning there is a correlation between our dependent and independent variable; however, the correlation is very weak. From Table 5.1, we see *standard error* (SE) 0.001603, which is close to zero. SE is an indication of the reliability of the mean. This small SE is an indication that the sample mean is a more accurate reflection of the actual population mean. The *correlation coefficient* is $r = -0.314$ indicating the negative relation between the variables. The coefficient of determination $R^2 = 0.090887$ indicates a very weak correlation between our predicted and predictor variables (see Table 5.7).

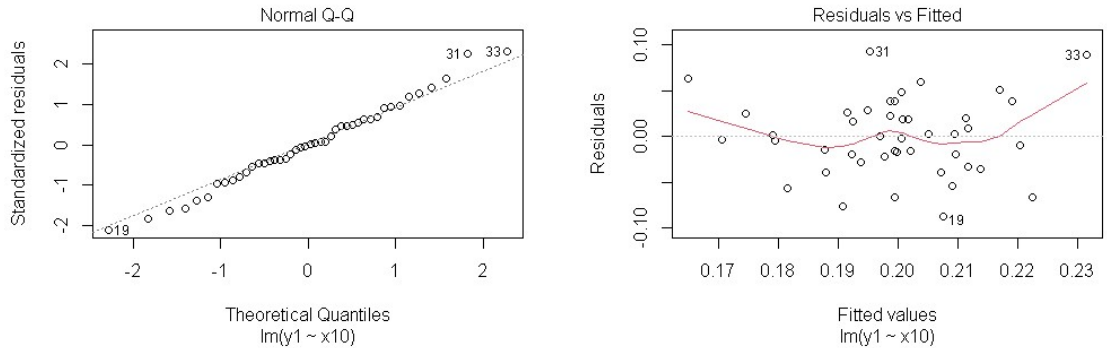
Table 5.1 Least Square Coefficients for case I

	Estimate	Std. Error	t value	Pr (|> |t||)
^ 1,0 β0	0.2922	0.043596	6.7	3.86e-08
^ 1,0 β1	-0.00344	0.0016	-2.147	0.037



(a) Q–Q plot of confirmed cases $Y_1(i)$ (b) Boxplot of confirmed cases $Y_1(i)$

Figure 5.2 Normality and outliers of confirmed cases



(a) Q-Q plot for residual in case I (b) Residual vs. fitted plot for case I

Figure 5.3 Model adequacies of residuals for case I

Now we look at our response variable, confirmed cases $Y_1(i)$. Q–Q plot of the response variable in Figure 5.2a shows that our confirmed cases are samples from a normal distribution. The boxplot in Figure 5.2b gives us a visual representation of the range and outliers of the confirmed cases. We get one outlier, Madison County, having the highest percent of cumulative confirmed cases.

Lastly, we verify the assumptions for the residuals to check for model adequacy. To validate the normality check for residual, we look at the Q–Q plot of our residual here, Fig 5.3a, which indicates the residual is from a normal distribution. Now we tested if residuals have constant variance by plotting the residual vs. fitted values $\hat{y}$ in Figure 5.3b. Here the pattern of dots is dense around the midline, but the red line deviates from the midline for small and large values of y_1. This reflects some extreme residuals represented by 19 (Custer County), 31 (Lewis County), and 33 (Madison County). However, from the boxplot of our y_1 in Figure 5.2b, we see the only significant outlier of confirmed cases is Madison County. This residual plot indicates either the variance of the residual may not be constant, or these counties are outliers.

We could understand the situation better by applying suitable transformation on the regressor and/or the response variable or the use of weighted least squares. These experiments have been left for future work.

5.2.2 Case II: Confirmed cases vs. medium risk in Idaho counties

In this case, our predictor variable is the percent of individuals with medium risk factors in Idaho counties, and the response variable is the rate of confirmed cases in corresponding counties. The scatter plot in Figure 5.4 shows the percentage of individuals with medium risk factor $x_1(i)$ as the x-axis and the percent of confirmed Covid cases $y_1(i)$ on the y-axis. The plot in Figure 5.4 indicates an uphill pattern as we move from left to right, implying a positive relationship between the number of people in the medium risk category and the number of confirmed cases.

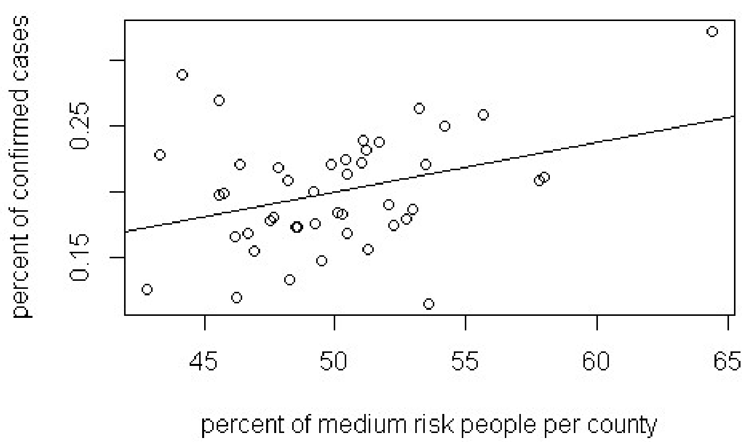


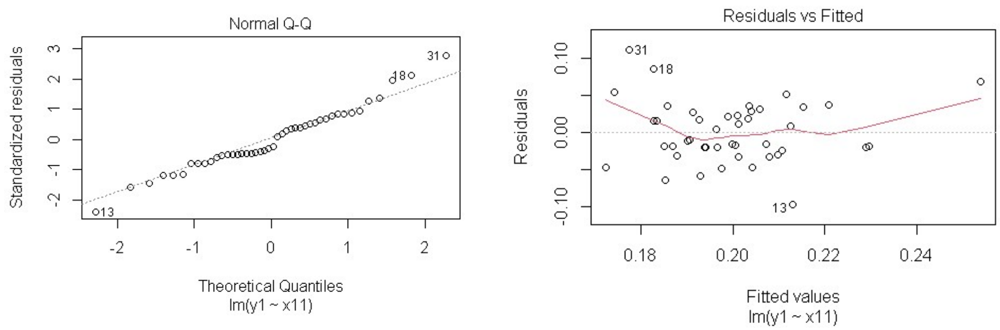
Figure 5.4 Case II: confirmed cases vs medium risk for 44 Idaho counties

We see in Table 5.2 for this case that the rate of confirmed cases increased by 0.003775 for each unit, increasing the rate of individuals with medium-risk factors. Our t-statistic value is 2.456 and its corresponding p-value is 0.018. We see standard error 0.0015, which is close to zero, indicating the reliability of the mean. From these results, we can conclude that there is strong evidence that the coefficients in this model are not zero, meaning there is a correlation between our dependent and independent variable; however, the correlation is very weak. The positive sign of the correlation coefficient r = 0.354 indicates the positive relation. The coefficient of determination (R^2 = 0.125) indicates a very weak correlation between our predicted and predictor variables.

The Q–Q plot of the response variable confirmed cases $Y_1(i)$ is given in Figure 5.2a. The plot indicates that our confirmed cases are samples from a normal distribution. The boxplot in Figure 5.2b is of one outlier, Madison County, having the highest percent of cumulative confirmed cases.

Table 5.2 Ordinary Least Square coefficients for case II

	Estimate	Std. Error	t value	Pr (|> |t||)
$\hat{\beta}_{1,1,0}$	0.0107	0.0771	0.139	0.889
$\hat{\beta}_{1,1,1}$	0.0037	0.0015	2.456	0.018



(a) Q-Q plot for residual in case II (b) residual vs. fitted plot for case II

Figure 5.5 Model adequacies of residuals for case II

Lastly, we verify the assumptions for the residuals. To validate the normality check for residual, we look at the Q–Q plot of our residual here, Fig 5.5a, which indicates the residual is from a normal distribution. We tested if residuals have constant variance by plotting the residual vs. fitted values $\hat{y}$ in Figure 5.5b. Here the pattern of dots is dense around the midline, but the red line deviates from the midline for large and small values for y, as in case I. In this case, some extreme residuals are 13 (Camas County), 18 (Clearwater County), and 31 (Lewis County). However, from the boxplot of our y_1 in Figure 5.2b, we see the significant outlier of confirmed cases is Madison County. Again we conclude the variance of the residual is not constant, or these counties are considered outliers, and future work indicates applying transformation or weighted least squares.

<u>5.2.3 Case III: Confirmed cases vs. high risk in Idaho counties</u>

In this case, our predictor variable is the percent of individuals with high-risk factors in Idaho counties, and the response variable is the rate of confirmed cases in corresponding counties. The scatter plot in Fig 5.6 shows the percentage of individuals with high-risk factor $x_2(i)$ as the x-axis and the percent of confirmed Covid cases $y_1(i)$ on the y-axis. Unfortunately, the plot in Fig 5.6 does not indicate any significant pattern as we move from left to right; we have to look at the summary statistics for further information.

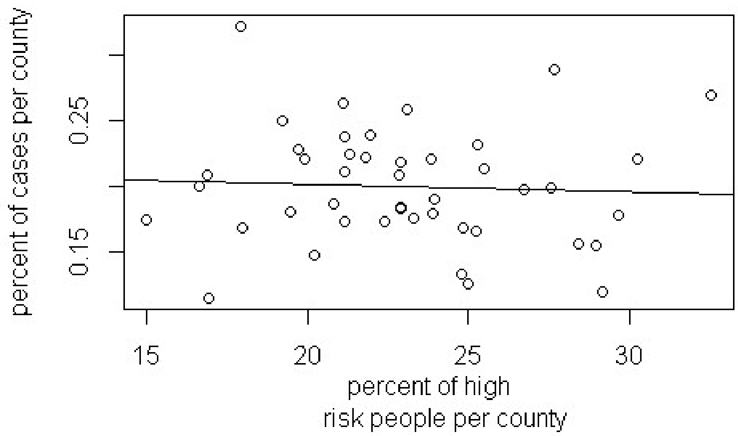
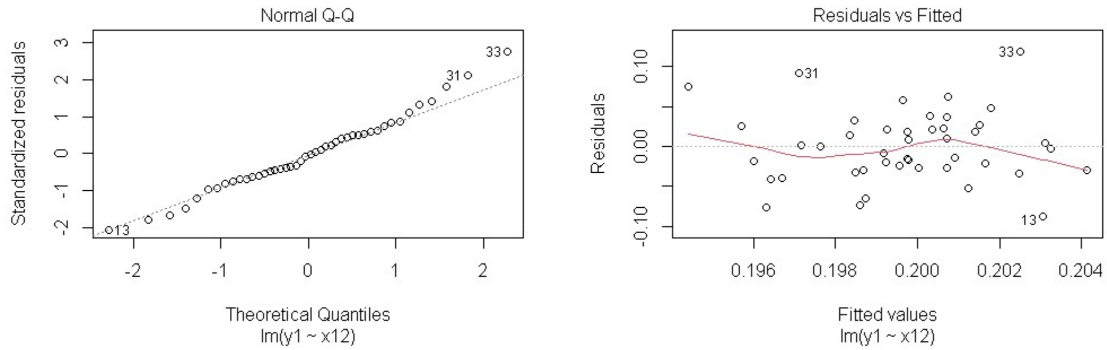


Figure 5.6 Case III: confirmed cases vs high risk for 44 Idaho counties

We see in Table 5.3 for this case that the rate of confirmed cases decreased by 0.000553 for each unit, increasing rate of individuals with high-risk factors. Our t-statistic value is −0.326 and its corresponding p-value is 0.746. we see standard error 0.0016, which is very small, indicating the reliability of the mean. From these results, we can conclude that there is insufficient evidence of a correlation between our dependent and independent variables. The positive sign of the correlation coefficient $r = -0.05$ indicates the positive relation. The coefficient of determination $R^2 = 0.0025$ indicates a very weak correlation between our predicted and predictor variables.



(a) Q-Q plot for residual in case III (b) residual vs. fitted plot for case III

Figure 5.7 Model adequacies of residuals for case III

The Q–Q plot of the response variable is given in Figure 5.2a. The plot indicates that our confirmed cases are samples from a normal distribution. The boxplot in Figure 5.2b is of one outlier, Madison County, having the highest percent of cumulative confirmed cases.

Table 5.3 Least Square Coefficients for case III

	Estimate	Std. Error	t value	Pr (|> |t||)
$\hat{\beta}^{1,2}_0$	0.2124	0.0396	5.364	0.0000032
$\hat{\beta}^{1,2}_1$	-0.0005	0.0016	-0.326	0.746

Lastly, we verify the assumptions for the residuals. To validate the normality check for residual, we look at the Q–Q plot of our residual here, Fig 5.7a, which indicates the residual is from a normal distribution. We tested if residuals have constant variance by plotting the residual vs. fitted values $\hat{y}$ in Figure 5.7b. Here the pattern of dots is dense around the midline, but the red line deviates from the midline for large and small values for y_1 as in case I. In this case, some extreme residuals are numbers 13 (Camas County), 31 (Lewis County), and 33 (Madison County). However, from the boxplot of our y_1 in Figure 5.2b, we see the significant outlier of confirmed cases is Madison County. Again we conclude the variance of the residual is not constant, or these counties are considered outliers, and future work indicates applying transformation or weighted lease squares.

5.2.4 Case IV: Confirmed deaths vs. low risk in Idaho counties

In this case, our predictor variable is the percent of individuals with low-risk factors in Idaho counties, and the response variable is the percent of confirmed deaths in corresponding counties. The scatter plot in Figure 5.8 shows the percentage of individuals with low-risk factor $x_0(i)$ as the x-axis and the percent of confirmed deaths $y_2(i)$ on the y-axis. The plot in Figure 5.8 indicates a downhill pattern as we move from left to right, implying a negative relationship between the number of people in the medium risk category and the number of confirmed deaths.

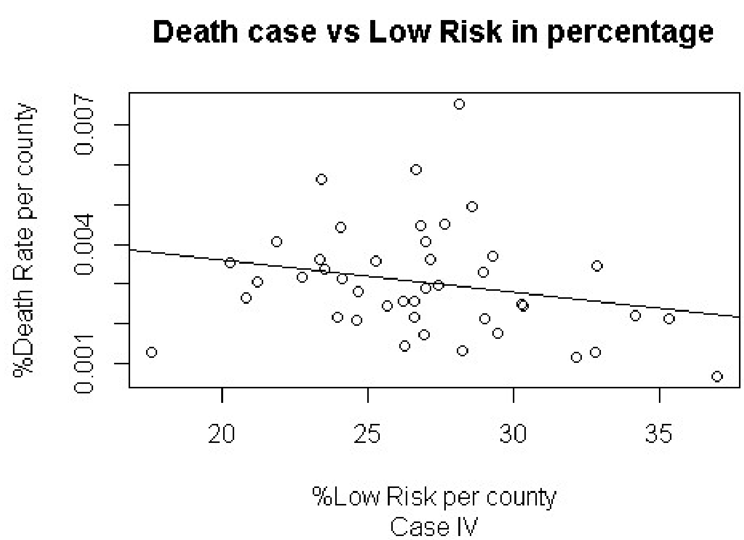
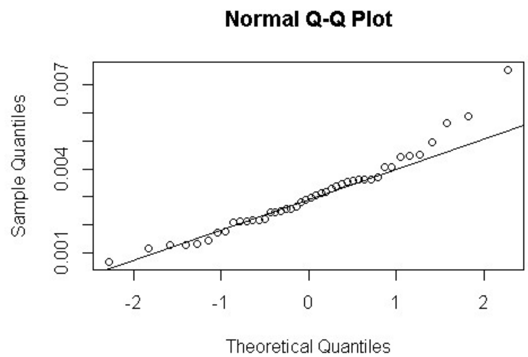
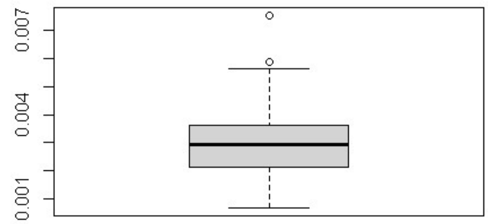


Figure 5.8 Case IV: confirmed deaths vs low risk for 44 Idaho counties



(a) Q–Q plot of confirmed deaths $Y_2(i)$



(b) Boxplot of confirmed deaths $Y_2(i)$

Figure 5.9 Pattern and outliers of confirmed deaths

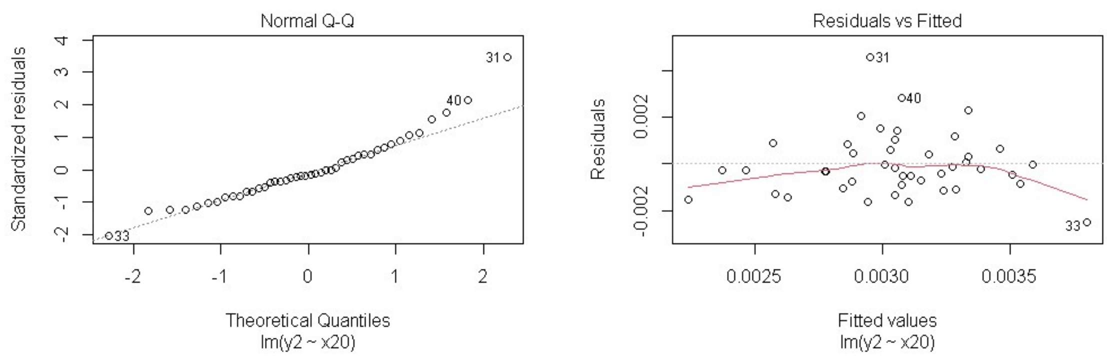
We see in Table 5.4 for this case that the rate of confirmed deaths decreased by $-8.055e-05$ for each unit, increasing the rate of individuals with low-risk factors. Our t-statistic value is -1.582 and its corresponding p-value is 0.121. We see standard error 0.00005. SE is close to zero, indicating that the sample mean accurately reflects the actual population mean. The negative sign of the correlation coefficient $r = -0.237$ indicates the negative relation. The coefficient of determination $R^2 = 0.056$ indicates that if there exists any linear relationship between our predicted and predictor variable, it is very weak. From these results, we can conclude that there is no solid evidence to conclude a correlation between our dependent and independent variable.

The Q–Q plot of the response variable confirmed deaths $Y_2(i)$ is given in Figure 5.9a. The plot indicates departure from a normal distribution. The boxplot in Figure 5.9b is of two outliers, Shoshone County and Lewis County, having the highest percent of cumulative confirmed deaths.

Lastly, we verify the assumptions for the residuals. To validate the normality check for residual, we look at the Q–Q plot of our residual here, Fig 5.10a, which indicates the residual is from a normal distribution. We tested if residuals have constant variance by plotting the residual plot in figure 5.10b. Here the pattern of dots is dense around the midline, but the red line deviates from the midline for large and small values for y_2.

Table 5.4 Least Square Coefficients for case IV

	Estimate	Std. Error	t value	Pr (|> |t||)
^ 2,0 β_0	5.2e-03	1.38e-03	3.76	0.0005
^ 2,0 β_1	-8.05e-05	5.09e-05	-1.582	0.121



(a) Q-Q plot for residual in case IV (b) residual vs. fitted plot for case IV

Figure 5.10 Model adequacies of residuals for case IV

In this case, some extreme residuals are 31 (Lewis county), 33 (Madison County), and 40 (Shoshone County). However, from the boxplot of our y_2 in Figure 5.9b, we see the significant outliers of confirmed deaths are Shoshone and Lewis County. Again we conclude the variance of the residual is not constant, or these counties are considered outliers and our future work indicates applying transformation or weighted least squares.

5.2.5 Case V: confirmed deaths vs. medium risk in Idaho counties

In this case, our predictor variable is the percent of individuals with medium-risk factors in Idaho counties, and the response variable is the rate of confirmed deaths in corresponding counties. The scatter plot in Figure 5.11 shows the percentage of individuals with medium risk factor $x_1(i)$ as the x-axis and the percent of confirmed deaths $y_2(i)$ on the y-axis. The plot in Figure 5.11 indicates a downhill pattern as we move from left to right, implying a negative relationship between the number of people in the medium risk category and the number of confirmed deaths.

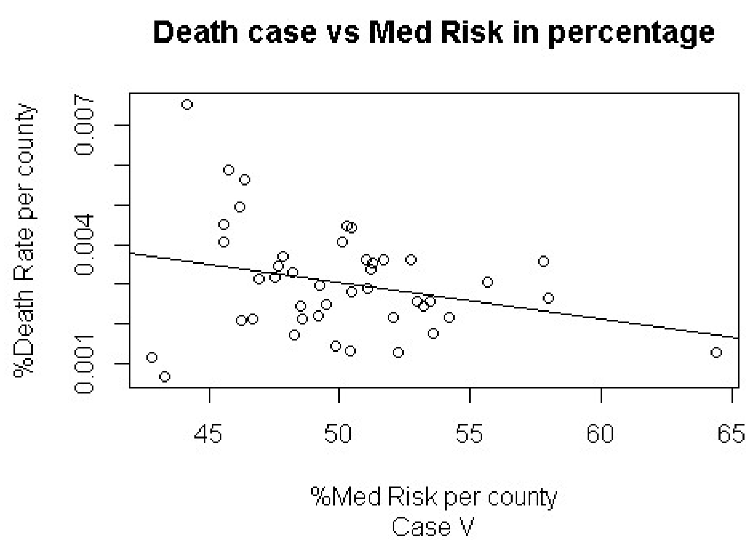
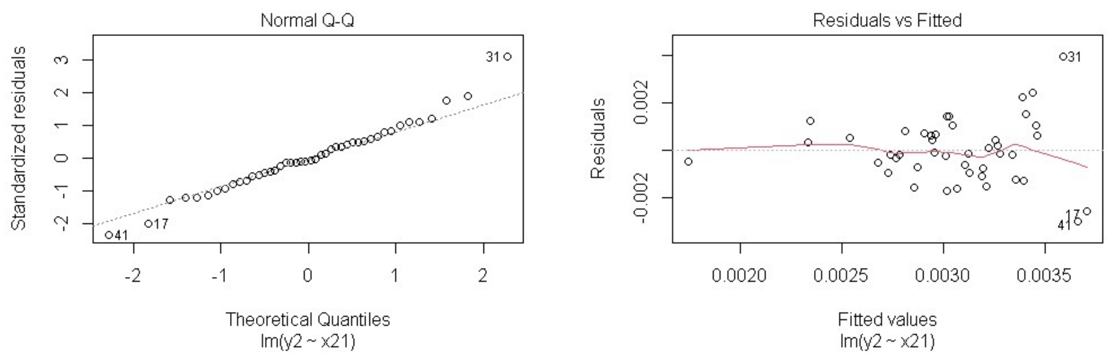


Figure 5.11 Case V: confirmed deaths vs medium risk for 44 Idaho counties

We see in Table 5.5 for this case that the rate of confirmed deaths decreased by 0.00009097 for each unit, increasing the rate of individuals with medium-risk factors. Our t-statistic value is −1.855, and its corresponding p-value is 0.07. We see the standard error is 0.000049, which is very small. The small SE value indicates that the sample mean is a more accurate reflection of the actual population mean. From these results, we can conclude that there is no solid evidence to conclude a correlation between our dependent and independent variable. The negative sign of the correlation coefficient $r = -0.275$ indicates the negative relation. The coefficient of determination $R^2 = 0.075$ indicates a very weak correlation between our predicted and predictor variables.



(a) Q-Q plot for residual in case V (b) residual vs. fitted plot for case V

Figure 5.12 Model adequacies of residuals for case V

The Q–Q plot of the response variable confirmed cases $y_2(i)$ is given in Figure 5.9a. The boxplot in Figure 5.9b gives us two outliers, Shoshone County and Lewis County, having the highest percent of cumulative confirmed deaths.

Table 5.5 Least Square Coefficients for case V

	Estimate	Std. Error	t value	Pr (|> |t||)
ˆ 2,1 β0	7.6e-03	2.4e-03	3.089	0.0035
ˆ 2,1 β1	-9.097e-05	4.9e-05	-1.855	0.07

Lastly, we verify the assumptions for the residuals. To validate the normality check for residual, we look at the Q–Q plot of our residual here, Fig 5.12a, which indicates the residual is from a normal distribution.

We tested if residuals have constant variance by plotting the residual vs. fitted values $\hat{y}$ in Figure 5.12b. Here the pattern of dots is dense around the midline, But the red line deviates from the midline for some small and large values for y_2. In this case, some extreme residuals are 17 (Clark County), 31 (Lewis County), and 41 (Teton County). However, from the boxplot of y_2 in Figure 5.9b, we see the significant outliers of confirmed deaths are Lewis and Shoshone County. Again we conclude the variance of the residual is not constant, or these counties are considered outliers and our future work indicates applying transformation or weighted least squares.

5.2.6 Case VI: Confirmed deaths vs. high risk in Idaho counties

In this case, our predictor variable is the percent of individuals with high-risk factors in Idaho counties, and the response variable is the rate of confirmed deaths in corresponding counties. The scatter plot in Figure 5.13 shows the percentage of individuals with high-risk factor $x_2(i)$ as the x-axis and the percent of confirmed Covid deaths $y_2(i)$ on the y-axis. The plot in Figure 5.13 indicates an uphill pattern as we move from left to right, implying a positive relationship between the number of people in the high risk category and the number of confirmed deaths.

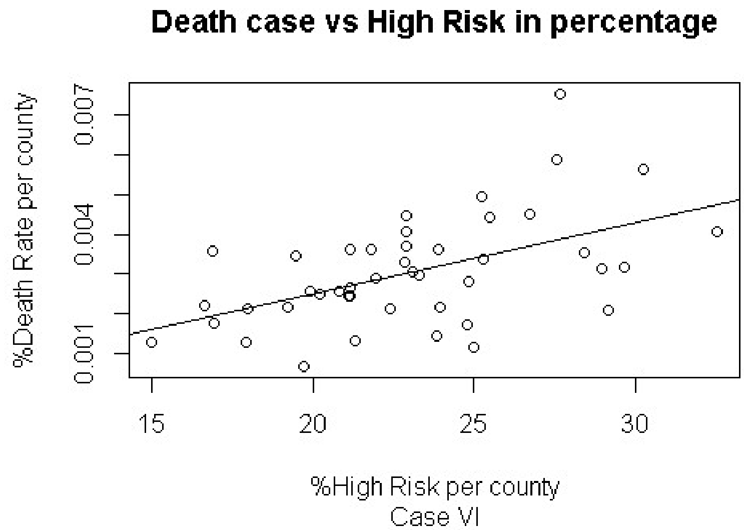


Figure 5.13 Case VI: confirmed deaths vs high risk for 44 Idaho counties

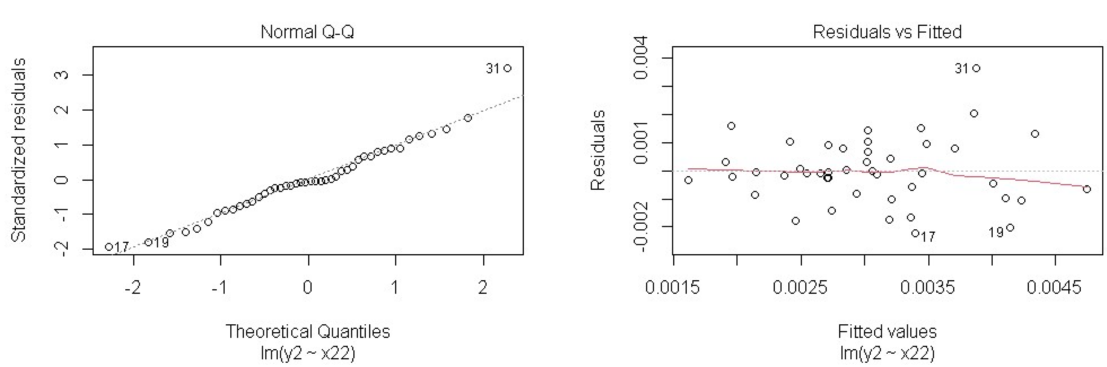
We see in Table 5.2 for this case that the rate of confirmed deaths increased by 0.00017 for each unit, increasing the rate of individuals with high-risk factors. Our t-statistic value is 3.96 and its corresponding p-value is 0.00027. We see standard error 0.0000448, which indicates that the sample mean accurately reflects the actual population mean. From these results, we can conclude that there is strong evidence that the coefficients in this model are not zero, meaning there is a correlation between our dependent and independent variable; however, the correlation is very weak. The positive sign of the correlation coefficient r = 0.522 indicates the positive relation. The coefficient of determination R^2= 0.272 indicates a very weak correlation between our predicted and predictor variables.

The Q–Q plot of the response variable is given in Figure 5.9a. The boxplot in Figure 5.9b is of two outliers, Shoshone County and Lewis County, having the highest percent of cumulative confirmed deaths.

Table 5.6 Least Square Coefficients for case VI

	Estimate	Std. Error	t value	Pr (|> |t||)
$\hat{\ }_{2,2}\beta_0$	-1.051e-03	1.05e-03	-1.002	0.322
$\hat{\ }_{2,2}\beta_1$	1.78e-04	4.48e-05	3.966	0.00027

Lastly, we verify the assumptions for the residuals. To validate the normality check for residual, we look at the Q–Q plot of our residual here, Figure 5.14a, which indicates the residual is from a normal distribution. We tested if residuals have constant variance by plotting the residual vs. fitted values $\hat{y}$ in Figure 5.14b. Here the pattern of dots is dense around the midline, but the red line deviates from the midline for large and small values for y. In this case, some extreme residuals are 17 (Clark County), 19 (Custer County), and 31 (Lewis County). However, from the boxplot of our y_2 in Figure 5.9b, we see the significant outliers of the confirmed deaths are Lewis and Shoshone County. Again we conclude the variance of the residual is not constant, or these counties are considered outliers and our future work indicates applying transformation or weighted least squares.



(a) Q-Q plot for residual in case VI (b) residual vs. fitted plot for case VI

Figure 5.14 Model adequacies of residuals for case VI

5.3 Discussion

P-values and correlation coefficients in regression analysis tell us which relationships in our model are statistically significant and the nature of those relationships. The coefficients describe the mathematical relationship between independent and dependent variables for six cases. The p-values help determine whether the relationships that we observe in the sample also exist in the population. Table 5.7 summarizes view of the linear relationship between our predicted and predictor variables.

From Table 5.7 we can see that for the case I, the significant p-value gives us evidence of a relationship between the number of cases and those in low risk categories, and a weak and negative correlation exists. For case II, the p-value is significant, and there exists a weak correlation between a number of cases and those in medium risk categories. For case III, the p-value is not significant, and a negligible correlation exists between a number of cases and those in high risk categories. For case IV, the p-value is not significant, as well as a negligible correlation exists between the number of deaths and those in low risk categories. For case V, the p-value is not significant, and a negligible correlation exists between the number of deaths and those in medium risk categories. For case VI, the p-value is significant, and there is a moderate positive correlation between the number of deaths and those in high-risk categories.

We also checked the model adequacy. All of the cases show negligence or weak relationship between variables. Also, cases III, IV, and V give us some insignificant results. We ran the same analysis excluding outliers we find from y_1 and y_2 and verified that the outliers are not significantly impacting the distributions. Furthermore, we are inconclusive if our variance of residual is constant in each case.

Table 5.7 Correlation Coefficients and p-values for six cases

Cases	Correlation Coefficient, *r*	p-values(< 0.05)
I	-0.3144	0.037
II	0.354	0.018
III	-0.237	0.121
IV	-0.2768076	0.07
V	-0.2267045	0.1389
VI	0.522	0.000279

CHAPTER 6: CONCLUSIONS AND FUTURE WORK

As the world is diverse, its community characteristics are discrete, and people are vulnerable to different factors depending on geography, race, ethnicity, etc. Thus, identifying a worldwide community resilience that could apply to multiple hazards is difficult to measure. But understanding the overall structural environment of global, regional, and local community resilience is crucial for public health strategies. Severity and frequency of a hazard, numbers of people and assets exposed to the hazard, and their vulnerability are some components of disaster risk. Reducing vulnerability is one of the most effective ways to reduce disaster risk.

After analyzing the current list of community resilience indicators, we choose to work with CRE tool. CRE tool designates at-risk populations by determining if a person has three or more factors, that make them particularly vulnerable. This available data giving insight on vulnerability, which is important for emergency management organizations. The CRE can help determine outreach services, the number and type of personnel to deploy, and disaster assistance programs that can be activated to aid affected areas. CRE also provides insight into a community's capacity to recover from a disaster like tornadoes, flooding and severe storm. However, our analysis shows that this metric is not very compatible with the COVID-19 John Hopkins data.

On the other hand, COVID-19 pandemic creates a multidimensional crisis, affecting different socioeconomic group and environmental area in different way. Unfortunately, Covid-19 has a more significant impact on poor areas. Poverty is both a driver and a consequence of disaster risk because economic pressures force people to live in unsafe conditions. Thus income-to-poverty ratio, crowded house, and full-time year-round employment are most likely relatable/significant indicators for COVID-19.

It is difficult to estimate the actual effect of the pandemic due to several reasons. Early in the pandemic, people who died of COVID-19 may not have been recognized because of inadequate knowledge of the symptom or lack of testing availability. Many people who die while infected with COVID-19 are never tested for it and do not enter the official totals. Some countries and organizations choose to hide the true toll of COVID-19 due to political or other issues. Conversely, some people whose deaths have been attributed to COVID-19 had other ailments that might have ended their lives in a similar timeframe anyway.

There are different methods to count the fatality rate, i.e., crude mortality rate, case fatality rate, infection fatality rate, and excess death rate. Various organizations follow different methods; thus, the results vary. Also, modeling epidemiologic years instead of calendar years would reduce the excess deaths estimate.

Rural and urban areas vary by the health infrastructure/number of hospitals and health services, and these could also influence the rate of COVID-19 cases and deaths. The health and safety between urban counties (9) and rural counties (35) in Idaho are not equal. The residents of rural and underserved areas tend to experience higher rates of poverty, lower per capita income, and distant from hospitals, as compared to their urban counterparts, thus more exposed to vulnerability [29]. For example, Washington County is showing high risk and having a greater number of confirmed death rate. Also, misinformation and poor communication disproportionally affect individuals with less access to information channels. These people are therefore more likely to ignore government health warnings.

Sample size affects the generalizability of the results. It is a genuine problem because a small sample size is associated with low statistical power. In this research, we are dealing with only 44 counties, and it is hard to get any kind of meaningful result from these few data points. And this could also applicable when dealing with small populations of the counties. For example, Lewis County shows outliers in both confirmed and death cases and its population size (3864) is relatively small compared to other counties. So, small sample size effect may affect Lewis County.

Our analysis did not get an enough correlation between the CRE tool and COVID-19 cases and deaths to use the data for further estimations and predictions. Residual vs. fitted plots of a few cases show heteroscedasticity may occur, but graphs and plots are very subjective. We ran the analysis with and without outliers, which did not create a huge difference. At the same time, we do not have enough logic to delete the COVID-19 variables just because the rates are high, as these are the true values from reliable sources. Also, we stopped updating our COVID-19 data on March 29, 2022, for our research analysis. There is a possibility that the future data could give a better relationship for our variables.

The CRE tool is the weighted aggregate of the ten risk factors. All these ten factors get equal importance in this measure, but in reality, some risk factors are more vulnerable than others. For example, age is a significant factor in this pandemic compared to households without internet access. This tool could be more useful if the data of the ten risk factors were available alongside with this CRE data, we could look at the separate risk factors as separate indicators. These individual indicators can recommend to which resources be directed, according to the reality of each county. Motivation behind this aggregate-level metric need further explanation. Overcoming these limitations could serve a better analysis and help us predict and create a resilient community during hazards and disasters.

For future research, there are several options available.

- Deleting the outlier and performing the same analysis to find a better relationship between the variables. However, strong reasoning is needed to delete the variables in this case.
- Performing data transformation, redefining dependent variables to remove heteroscedasticity for each case, and running the analysis to get a better relationship between variables.
- If there is no strong correlation, chances are there are non-linear models. Here, fitting different non-linear models to different cases can be a solution. However, these need a lot of trial-and-error processes.
- Confirmed deaths (Y_2) showing departure from normal distribution. Usually, this could happen for 2 reasons,
    1. Our dependent variable confirmed deaths come from a non-normal distribution.
    2. Existence of a few outliers or extreme values which disrupt the model prediction.

In this case, several normality tests are there to run, for example, Shapiro-Wilk test, Kolmogorov–Smirnov test.

- Run Logistic Model instead of simple linear regression, for each case.
- Research the indicators further and improve the method.
- Analyze other community resilience indicators to get the best fit for COVID-19 data.

REFERENCES

[1] David Adam. The pandemic's true death toll: millions more than official counts. *Nature*, 601(7893):312–315, 2022.

[2] Jo-An Atkinson, Andrew Vallely, Lisa Fitzgerald, Maxine Whittaker, and Marcel Tanner. The architecture and effect of participation: a systematic review of community participation for communicable disease control and elimination. implications for malaria elimination. *Malaria Journal*, 10(1):1–33, August 2011.

[3] Hamada S. Badr, Benjamin F. Zaitchik, Gaige H. Kerr, Nhat-Lan Nguyen, Yen-Ting Chen, Patrick Hinson, Josh M. Colston, Margaret N. Kosek, Ensheng Dong, Hongru Du, Maximilian Marshall, Kristen Nixon, Arash Mohegh, Daniel L. Goldberg, Susan C. Anenberg, and Lauren M. Gardner. Unified real-time environmentalepidemiological data for multiscale modeling of the covid-19 pandemic. *medRxiv*, 2021.

[4] Jamie Bedson, Mohamed Jalloh, Danielle Pedi, Saiku Bah, Katharine Owen, Allan Oniba, Musa Sangarie, James Fofanah, Mohammed Jalloh, Paul Sengeh, Laura Skrip, Benjamin Althouse, and Laurent Hebert-Dufresne. Community engagement´ in outbreak response: Lessons from the 2014-2016 ebola outbreak in sierra leone. *BMJ Global Health*, 5:e002145, 08 2020.

[5] United States Census Bureau. Community resilience estimates. www.census.gov/data/academy/webinars/2021/community-resilience-estimates.html, June 2000.

[6] UK Cabinet Office. Strategic national framework on community resilience. *London: UK Government*, pages 1–25, 2011.

[7] Kenneth Rochel de Camargo Jr. Trying to make sense out of chaos: science, politics and the covid-19 pandemic. *Cadernos de Saude P´ ublica´* , 36:e00088120, 2020.

[8] Matthew Castleden, Martin McKee, Virginia Murray, and Giovanni Leonardi. Resilience thinking in health protection. *Journal of public health*, 33(3):369–377, 2011.

[9] Shawn C Chiang, Holly H Fisher, Matthew E Bridwell, Silvia M Trigoso, Bobby B Rasulnia, and Sachiko A Kuwabara. Applying the ready, willing, and able framework to assess agency public health emergency preparedness: the cdc perspective. *Health security*, 18(2):75–82, 2020.

[10] Matteo Cinelli, Walter Quattrociocchi, Alessandro Galeazzi, Carlo Michele Valensise, Emanuele Brugnoli, Ana Lucia Schmidt, Paola Zola, Fabiana Zollo, and Antonio Scala. The covid-19 social media infodemic. *Scientific reports*, 10(1):1–10, 2020.

[11] Eve Coles and Philip Buckle. Developing community resilience as a foundation for effective disaster recovery. *Australian Journal of Emergency Management, The*, 19(4):6–15, 2004.

[12] Robin S Cox and Karen-Marie Elah Perry. Like a fish out of water: reconsidering disaster recovery and the role of place and social capital in community disaster resilience. *American journal of community psychology*, 48(3-4):395–411, 2011.

[13] Dmitry M Davydov, Robert Stewart, Karen Ritchie, and Isabelle Chaudieu. Resilience and mental health. *Clinical psychology review*, 30(5):479–495, 2010.

[14] Paolo Di Giamberardino and Daniela Iacoviello. Epidemic modeling and control of hiv/aids dynamics in populations under external interactions: a worldwide challenge. In *Control Applications for Biomedical Engineering Systems*, pages 197–249. Elsevier, 2020.

[15] Ensheng Dong, Hongru Du, and Lauren Gardner. An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, 20(5):533–534, 2020.

[16] Bao V Duong, Puchanee Larpruenrudee, Tianxin Fang, Sheikh I Hossain, Suvash C Saha, Yuantong Gu, and Mohammad S Islam. Is the sars cov-2 omicron variant deadlier and more transmissible than delta variant? *International Journal of Environmental Research and Public Health*, 19(8):4586, 2022.

[17] Department for International Development. Defining disaster resilience:a dfid approach paper, 2011.

[18] Carl Gibson. An integrated approach to managing disruption-related risk: life and death in a model community. *Journal of business continuity & emergency planning*, 4(3):246—261, 2010.

[19] Brynne Gilmore, Rawlance Ndejjo, Adalbert Tchetchia, Vergil de Claro, Elizabeth Mago, Alpha A Diallo, Claudia Lopes, and Sanghita Bhattacharyya. Community engagement for covid-19 prevention and control: a rapid evidence synthesis. *BMJ Global Health*, 5(10), 2020.

[20] OCHA IFRC. Covid-19: Community insights from the asia pacific region - indonesia, malaysia, myanmar, and pakistan, September 2020.

[21] Ariel Karlinsky and Dmitry Kobak. Tracking excess mortality across countries during the covid-19 pandemic with the world mortality dataset. *Elife*, 10:e69336, 2021.

[22] Carolyn Kousky. *Impacts of Natural Disasters on Children*. Number 1. Princeton University, 2016. http://www.jstor.org/stable/43755231.

[23] Kristen Magis. Community resilience: An indicator of social sustainability. *Society and natural resources*, 23(5):401–416, 2010.

[24] Siambabala Bernard Manyena. The concept of resilience revisited. *Disasters*, 30(4):434–450, 2006.

[25] Joseph Stephen Mayunga. *Measuring the measure: A multi-dimensional scale model to measure community disaster resilience in the US Gulf Coast region*. Texas A&M University, 2009.

[26] Douglas C. Montgomery, Elizabeth A. Peck, and Geoffrey G. Vining. *Introduction to Linear Regression Analysis (5th ed.)*. John Wiley & Sons, 2012.

[27] U Nations. Report of the world conference on disaster reduction. *Geneva: United Nations*, pages 1–42, 2005.

[28] Fran H Norris, Susan P Stevens, Betty Pfefferbaum, Karen F Wyche, and Rose L Pfefferbaum. Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness. *American journal of community psychology*, 41(1-2):127–150, mar 2008.

[29] Idaho Divison of Public Health. Bureu of rural health and primary care brief. february 2020. ruralhealth.dhw.idaho.go.

[30] World Health Organization et al. Covid-19 global risk communication and community engagement strategy, december 2020-may 2021: interim guidance, 23 december 2020. Technical report, World Health Organization, 2020.

[31] World Health Organization et al. Estimating mortality from covid-19. Technical report, World Health Organization, 2020.

[32] World Health Organization et al. Pandemic fatigue: reinvigorating the public to prevent covid-19: policy framework for supporting pandemic prevention and management: revised version november 2020. Technical report, World Health Organization. Regional Office for Europe, 2020.

[33] Abbas Ostad Taghizadeh, Ali Ardalan, Douglas Paton, Hossain Jabbari, and Hamid Reza Khankeh. Community disaster resilience: a systematic review on assessment models and tools. *PLoS Currents*, 7, 04 2015.

[34] Sonny Patel, Brooke Rogers, Richard Amlot, and G. Rubin. What do we mean byˆ ’community resilience’? a systematic literature review of how it is defined in the literature. *PLoS Currents*, 9, 02 2017.

[35] Douglas Paton, Marian Millar, and David Johnston. Community resilience to volcanic hazard consequences. *Natural hazards*, 24(2):157–169, 2001.

[36] Virendra Proag. The concept of vulnerability and resilience. *Procedia Economics and Finance*, 18:369–376, 2014. 4th International Conference on Building Resilience, Incorporating the 3rd Annual Conference of the ANDROID Disaster Resilience Network, 8th – 11th September 2014, Salford Quays, United Kingdom.

[37] Olha Puhach, Kenneth Adea, Nicolas Hulo, Pascale Sattonnet, Camille Genecand, Anne Iten, Fred´ erique Jacqu´ erioz Bausch, Laurent Kaiser, Pauline Vetter, Isabella´ Eckerle, et al. Infectious viral load in unvaccinated and vaccinated individuals infected with ancestral, delta or omicron sars-cov-2. *Nature Medicine*, pages 1–1, 2022.

[38] Disaster Risk Reduction. Unisdr terminology on disaster risk reduction. *UNISDR*, page 30, 2009.

[39] Mark D Slivkoff, Catherine Johnson, and Sean Tackett. First-year medical student experiences adjusting to the immediate aftermath of covid-19. *Medical Science Educator*, 31(2):557–564, 2021.

[40] Anuradha Tomar and Neeraj Gupta. Prediction for the spread of covid-19 in india and effectiveness of preventive measures. *Science of The Total Environment*, 728:138762, 2020.

[41] John S Tregoning, Katie E Flight, Sophie L Higham, Ziyin Wang, and Benjamin F Pierce. Progress of the covid-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness and escape. *Nature Reviews Immunology*, 21(10):626–636, 2021.

[42] John Twigg. Characteristics of a disaster-resilient community: A guidance note. august 2007.

[43] Haidong Wang, Katherine R Paulson, Spencer A Pease, Stefanie Watson, Haley Comfort, Peng Zheng, Aleksandr Y Aravkin, Catherine Bisignano, Ryan M Barber, Tahiya Alam, et al. Estimating excess mortality due to the covid-19 pandemic: a systematic analysis of covid-19-related mortality, 2020–21. *The Lancet*, 2022.

[44] Katharine Wulff, Darrin Donato, and Nicole Lurie. What is health resilience and how can we build it? *Annual review of public health*, 36:361–374, 2015.

[45] John Zarocostas. How to fight an infodemic. *The lancet*, 395(10225):676, 2020.

APPENDIX A

List of Counties with Covid-19 Confirmed Cases

No.	County Name	Population	Low Risk	Medium Risk	High Risk	Percent of confirmed cases
13	Camas County	1127	29.46	53.59	16.95	0.1153505
19	Custer County	4271	24.61	46.24	29.15	0.1203465
17	Clark County	852	32.16	42.84	25.00	0.1255869
8	Boise County	7600	26.95	48.28	24.78	0.1335526
4	Bear Lake County	6050	30.30	49.47	20.23	0.1484298
30	Lemhi County	7949	24.15	46.89	28.96	0.1557429
2	Adams County	4231	20.28	51.29	28.43	0.1567005
11	Boundary County	11928	28.60	46.17	25.23	0.1656606
21	Franklin County	13713	35.34	46.69	17.98	0.1680158
32	Lincoln County	5360	24.68	50.47	24.85	0.1684701
36	Oneida County	4488	30.35	48.48	21.17	0.1733512
22	Fremont County	12751	29.04	48.54	22.42	0.1735550
29	Latah County	37379	32.78	52.22	15.01	0.1749646
34	Minidoka County	20816	27.45	49.27	23.28	0.1757302
25	Idaho County	16105	22.78	47.54	29.68	0.1782676
24	Gooding County	15153	23.39	52.72	23.88	0.1791724
12	Butte County	2609	32.85	47.68	19.47	0.1809122
37	Owyhee County	11614	26.85	50.28	22.88	0.1831410
9	Bonner County	44625	26.97	50.13	22.90	0.1840672
39	Power County	7748	26.20	52.98	20.82	0.1862416
16	Cassia County	23730	23.99	52.03	23.97	0.1904762
23	Gem County	17581	27.67	45.59	26.74	0.1975428
40	Shoshone County	12736	26.65	45.76	27.59	0.1985710
26	Jefferson County	29359	34.18	49.20	16.63	0.2003134
15	Caribou County	7009	25.31	57.83	16.86	0.2080183
6	Bingham County	46114	28.97	48.18	22.86	0.208852
27	Jerome County	23956	20.85	57.97	21.18	0.2112623
5	Benewah County	9217	24.06	50.46	25.47	0.2128675
35	Nez Perce County	39879	29.28	47.82	22.90	0.2178089
20	Elmore County	27122	26.58	53.49	19.93	0.2204483
43	Valley County	11002	26.29	49.87	23.84	0.2206871
44	Washington County	10117	23.40	46.35	30.26	0.2207176
38	Payette County	23496	27.15	51.02	21.83	0.2219527
7	Blaine County	22601	28.26	50.41	21.32	0.2237954
41	Teton County	11640	36.97	43.31	19.73	0.2281787
42	Twin Falls County	85624	23.51	51.21	25.28	0.231523
28	Kootenai County	160924	27.15	51.67	21.18	0.237174
3	Bannock County	85629	26.97	51.09	21.94	0.2385056
1	Ada County	461076	26.61	54.17	19.22	0.2496465

14	Canyon County	221400	21.23	55.67	23.10	0.13930894
10	Bonneville County	116497	25.68	53.21	21.11	0.2582114
18	Clearwater County	8086	21.88	45.56	32.56	0.2687361
31	Lewis County	3846	28.16	44.15	27.69	0.2888716
33	Madison County	38730	17.63	64.42	17.94	0.3215337

APPENDIX B

List of Counties with Covid Confirmed Deaths

No.	County Name	Population	Low Risk	Medium Risk	High Risk	Percent of confirmed deaths
41	Teton County	11640	36.97	43.31	19.73	0.0006872852
17	Clark County	852	32.16	42.84	25.00	0.0011737089
29	Latah County	37379	32.78	52.22	15.01	0.0012841435
33	Madison County	38730	17.63	64.42	17.94	0.0012909889
7	Blaine County	22601	28.26	50.41	21.32	0.0013273749
43	Valley County	11002	26.29	49.87	23.84	0.0014542810
8	Boise County	7600	26.95	48.28	24.78	0.0017105263
13	Camas County	1127	29.46	53.59	16.95	0.0017746229
19	Custer County	4271	24.61	46.24	29.15	0.0021072348
21	Franklin County	13713	35.34	46.69	17.98	0.0021147816
22	Fremont County	12751	29.04	48.54	22.42	0.002117481
1	Ada County	461076	26.61	54.17	19.22	0.0021861906
16	Cassia County	23730	23.99	52.03	23.97	0.0021913190
26	Jefferson County	29359	34.18	49.20	16.63	0.0022139719
10	Bonneville County	116497	25.68	53.21	21.11	0.0024378310
36	Oneida County	4488	30.35	48.48	21.17	0.0024509804
4	Bear Lake County	6050	30.30	49.47	20.23	0.0024793388
20	Elmore County	27122	26.58	53.49	19.93	0.00258093
39	Power County	7748	26.20	52.98	20.82	0.0025813113
27	Jerome County	23956	20.85	57.97	21.18	0.0026715645
32	Lincoln County	5360	24.68	50.47	24.85	0.0027985075
3	Bannock County	85629	26.97	51.09	21.94	0.0028845368
34	Minidoka County	20816	27.45	49.27	23.28	0.0029784781
14	Canyon County	221400	21.23	55.67	23.10	0.0030578139
30	Lemhi County	7949	24.15	46.89	28.96	0.0031450497
25	Idaho County	16105	22.78	47.54	29.68	0.0031667184
6	Bingham County	46114	28.97	48.18	22.86	0.0033178644
42	Twin Falls County	85624	23.51	51.21	25.28	0.0033752219
12	Butte County	2609	32.85	47.68	19.47	0.0034495975
2	Adams County	4231	20.28	51.29	28.43	0.0035452612
15	Caribou County	7009	25.31	57.83	16.86	0.0035668426
38	Payette County	23496	27.15	51.02	21.83	0.0036176370
24	Gooding County	15153	23.39	52.72	23.88	0.0036296443
28	Kootenai County	160924	27.15	51.67	21.18	0.0036414705
35	Nez Perce County	39879	29.28	47.82	22.90	0.0036861506
9	Bonner County	44625	26.97	50.13	22.90	0.0040784314
18	Clearwater County	8086	21.88	45.56	32.56	0.0040811279
5	Benewah County	9217	24.06	50.46	25.47	0.0044483021
37	Owyhee County	11614	26.85	50.28	22.88	0.0044773549

23	Gem County	17581	27.67	45.59	26.74	0.0044934873
11	Boundary County	11928	28.60	46.17	25.23	0.0049463447
44	Washington County	10117	23.40	46.35	30.26	0.0056340812
40	Shoshone County	12736	26.65	45.76	27.59	0.0058888191
31	Lewis County	3846	28.16	44.15	27.69	0.0075403016

APPENDIX C

Covid-19 Data Sources

Terms for COVID-19 data:

Confirmed Cases: Confirmed cases are counts of individuals whose coronavirus infections were confirmed by a laboratory test and reported by a federal, state, territorial or local government agency. Only tests that detect viral RNA in a sample are considered confirmatory. These are often called molecular or RT-PCR tests [3].

Confirmed Deaths: Confirmed deaths are individuals who have died and meet the definition for a confirmed COVID-19 case. Some states reconcile these records with death certificates to remove deaths from their count where COVID-19 is not listed as the cause of death. We follow health departments in removing non-COVID-19 deaths among confirmed cases when we have information to unambiguously know the deaths were not due to COVID-19, i.e. in cases of homicide, suicide, car crash or drug overdose [3].

Source	Description	Level
JHU	Johns Hopkins University CSSE	Global County/State, United States
CTP	The COVID Tracking Project	State, United States
NYC	New York City Department of Health and Mental Hygiene	ZCTA/Borough, New York City
NYT	The New York Times	County/State, United States
UVA	University of Virginia School of Medicine	Municipality/State, South America
SES	Monitoring COVID-19 Cases and Deaths in Brazil	Municipality/State/Country, Brazil
DPC	Italian Civil Protection Department	NUTS 0-3, Italy
RKI	Robert Koch-Institut, Germany	NUTS 0-3, Germany
JRC	Joint Research Centre	Global NUTS 0-3, Europe
ERA5	The fifth generation of ECMWF reanalysis	All levels
NLDAS	North American Land Data Assimilation System	County/State, United States
CIESIN	C. for International Earth Science Information Net.	Global gridded population
OxCGRT	Oxford COVID-19 Government Response Tracker	National (global) subnational (US, UK)
CRC	Johns Hopkins Centers for Civic Impact	National (global) subnational (US)
IHME	Institute for Health Metrics and Evaluation	National (global) subnational (US)