SOCIAL INFRASTRUCTURE'S IMPACT ON LONELINESS: PROPOSING A NEW LONELINESS INDEX TO PREDICT RELATIONSHIP OF SOCIAL INFRASTRUCTURES AND LONELINESS

BY

SHIRA ZUR

THESIS

Submitted in fulfillment of the requirements for the Honors degree of Geography: Data Science in the Department of Geography at the University of Washington, 2025

Seattle, Washington

Adviser:

Professor Bo Zhao

ABSTRACT

This project aims to continue and grow the conversation about the loneliness epidemic. Loneliness, as experienced both physically and mentally, has major, significant impacts on our individual and societal health. Social infrastructures have been proven to be strong loneliness mitigation tools, as they create a physical space of connectedness and belonging within a community. To see whether the number of and distance to social infrastructures could predict a lower experienced loneliness in a neighborhood, I propose a new loneliness index, composed of 13 social variables that have been proven to be indicative of loneliness, to measure loneliness on a census tract level. This project maps social infrastructures (libraries and parks) and a new loneliness index (created through the Composite Index tool in ArcGIS Pro) in Seattle, Washington, to observe the impact of social infrastructures on loneliness in a metropolitan area. The results show that libraries have little to no correlation to a census tract's reported loneliness on the new loneliness index, while parks have some correlation. This project suggests future research should observe both the quality of social infrastructures as well as the quantity of social infrastructures in a city, as well as implement further individual measures of loneliness, such as a participatory map of reported loneliness, to better understand the impact social infrastructures can have on mitigating and reducing loneliness in a major city.

ACKNOWLEDGEMENTS

I would like to thank the faculty and staff at the Department of Geography for encouraging me to take on and complete this project, especially Professor Bo Zhao who provided insightful feedback and guidance. I would also like to thank Dr. Vivek Murthy for bringing this topic to light, and any and all medical and mental health specialists that are working hard to tackle this difficult and important issue. I hope this project inspires others to discuss loneliness more openly, and that over time, we will see a significant decrease in reported experienced loneliness.

TABLE OF CONTENTS

CHAPTER 1:
AN INTRODUCTION TO LONELINESS

In spring of 2023, the United States Surgeon General Dr. Vivek Murthy published an alarming report about the impacts of loneliness in America. It significantly labeled loneliness as an "epidemic" whose wide-ranging effects can lead to fatal results (Murthy, 2023). As a mitigation technique, Dr. Murthy advised a national response of several steps, with the first one being to improve existing local social infrastructure, such as libraries and parks.

Dr. Murthy's report set off a red warning sign: loneliness is a major public health issue that must be addressed. This study aims to be part of that cause. A missing puzzle piece in better understanding loneliness is that in most previous studies, loneliness is typically measured on an individual level – sampled groups answer survey questions about their subjective experience with loneliness. In this GIS-based analysis of the city of Seattle, I hope to fill this gap by providing a new, neighborhood-level loneliness index scale, which can be used interchangeably across other metropolitan areas in the United States.

A better understanding of loneliness on a neighborhood-level scale can provide valuable knowledge for city governments across the United States. If cities had a better way of spatially measuring loneliness, they could employ specific, targeted mitigation approaches to tackle the issue. Taking Seattle as an example, this could mean identifying specific neighborhoods that need more resources and support to combat loneliness. By utilizing the information from the map this study will create, city planners can be better informed when deciding where to build new parks, libraries, or other community centers.

Beyond Seattle, this map could help inform other city governments and encourage them to create similar loneliness indexes for their cities. It would also overall continue spreading the information that loneliness is a major public health issue that needs to be addressed.

The steps I took to complete this project were to 1) identify 13 vital social factors that are predictive of loneliness, 2) combine them to create a composite loneliness index, and 3) map social infrastructures, specifically parks and libraries, and 4) assess whether proximity to such social infrastructures can predict a lower score on a new loneliness index. As I will discuss later on in this report, I found that libraries had little to no significant correlation to experienced loneliness on my proposed index, while parks had a weak correlation to experienced loneliness on my proposed index. Limitations such as the lack of ground-truthing and assumptions indexes such as these make led to limited findings. Still, ultimately, the goal of this project was to start a more nuanced, broader conversation about loneliness, and this project is just one approach; I hope others will be inspired to follow suit.

CHAPTER 2:
LITERATURE REVIEW

2.1 The Loneliness Epidemic

In Spring of 2023, the United States Surgeon General Dr. Vivek Murthy published a report about the impacts of loneliness on American society. Formally labeling loneliness as an "epidemic", Dr. Murthy highlighted the detriment that social isolation can have on individuals, such as a "29% increased risk of heart disease", a "32% increased risk of stroke", and, overall, a "26% increased risk of premature death" (Murthy, 2023, p.8).

Dr. Murthy's report builds on many research studies that have pointed out the crucial negative toll loneliness can have on our lives (Crowe *et al.*, 2024; Hughes *et al*., 2004; Steptoe *et al*., 2013; Stickley and Ueda, 2022; Mann *et al.*, 2022; Prohaska *et al*., 2022; Algren *et al*., 2020; Kearns *et al*., 2015). Previous literature on the matter discusses how loneliness can both cause physical health problems, such as increased mortality (Steptoe *et al.*, 2013) and engagement in health-risk behaviors (Algren *et al.*, 2020), as well as mental health problems, such as anxiety symptoms (Stickley and Ueda, 2022) and onset of depression (Mann *et al.*, 2022). Physical and mental health problems caused by loneliness are also wide-ranging; they impact both the elderly population (Prohaska *et al*., 2020) as well as young adults (Stickley and Ueda, 2022).

As Dr. Murthy stated in his report, we as a society are quick to recognize the everyday needs of a human being, such as food, water, and a roof over our heads. Social connection is just as much a part of this list of basic human needs for survival, and addressing the lack of it is fundamental for our society's overall public health.

2.1.1 Mitigation and Social Infrastructure

In his report, Dr. Murthy advised a mitigation strategy based on six pillars. The first pillar he highlighted was the need to improve upon our existing local social infrastructure (Murthy, 2023, p.18). He defined social infrastructure as the "physical assets of a community", and discussed how these assets can help a neighborhood and its individuals by providing opportunities for connection and community building (Murthy, 2023, p.18).

Plenty of literature exists demonstrating the positive impacts of third places, which are spaces that are "separate from work or home" and are "places for the community to gather and create relationships" and can be defined as a type of social infrastructure (Reed and Bohr, 2020, p.394). These places, which "facilitate physical activities" and "community building" and "provide options for food" have been proven as a mitigation technique to reduce loneliness (Jing *et al*., 2024, p.14). This finding has been reported across the globe – for instance, a study in Singapore found that individuals living closer to a wet market, which was defined as a third place, reported overall better social health (Lane *et al.*, 2020). Overall, a relationship between strong social infrastructure and a community's ability to overcome loneliness can be observed.

On the opposite end, previous scholarly literature has also examined how the absence of third places (Crowe *et al*., 2024) and poor social infrastructure (Kearns *et al.*, 2015; Satcher, Okfaor, and Dill, 2012) can negatively impact neighborhoods. A study synthesizing previous literature on the connection of built environments and mental health discussed how certain neighborhood characteristics, such as "lack of greenspace … [and] lack of services and resources such as grocery stores, libraries, and recreation centers" have been established as "negative determinants of health" (Satcher, Okafor, and Dill, 2012, p.12). Another study, which looked at self-reported cases of loneliness in neighborhoods in Glasgow, found that residents that viewed their neighborhood and its built environment as being of lower quality were more likely to report feeling lonely (Kearns *et al.,* 2015).

The overall consensus that can be found in several previous studies indicates that social infrastructure can help mediate loneliness (Kim and Kim, 2024; Crowe *et al.*, 2024; Reed and Bohr, 2020; Kearns *et al.*, 2015). This aligns with previous literature that discusses the role social integration and connectedness has in loneliness reduction (Swader, 2018), as well as Dr. Murthy's overall report. When social infrastructure is well-made and preserved, it can be an incredibly strong tool in helping individuals find belonging, make connections, and ultimately feel less lonely.

2.1.1.1 The Gap in Measuring and Analyzing Loneliness

While a link can be found in a community's relationship to a social infrastructure or third place and its reduced loneliness, there exist two issues with previous such studies. Firstly, the most frequently used scale to measure loneliness is the UCLA 3-item questionnaire (Das *et al.*, 2021). This scale was established in 2004 by researchers from the University of Chicago and Duke University as a shortened version of the original 1978 20-item scale developed in UCLA in order to allow for larger-scale social surveys (Hughes *et al*., 2004). The 3-item scale asks individuals how often they: 1) feel they lack companionship, 2) feel they are left out, and 3) feel isolated from others, with limited response options of: 1) "hardly ever", 2) "some of the time", or 3) "often" (Hughes *et al.,* 2004). This kind of close-ended, pre-coded survey can be quite limited. The available response options mean that individuals can't categorize themselves as in-between cases – such as not feeling lonely often, which could go between "hardly ever" and "some of the time", or feeling lonely quite often, which could go between "some of the time" and "often" – and therefore its operationalization is not adequately exhaustive.

Secondly, there exists a gap between the discussion of the respondent's geography and their response on the scale. Studies utilizing this scale discuss regional outcomes, such as comparing suburban versus urban neighborhoods (Abshire *et al.*, 2022), or across countries (Swader, 2018). However, this means that there exists no open neighborhood-level data of a loneliness scale, which makes it harder both for researchers to better understand loneliness in a specific location but also for cities to implement policies. To study loneliness on a neighborhood-level scale, a researcher would be required to send out the 3-item survey across the region and hope to get diverse, plentiful responses. This type of probability sampling technique can be a costly and time-consuming method.

Moreover, there exists a major gap when studies reach the step of analyzing and presenting their findings on loneliness. There are not many studies that utilize the tools of GIS (Geographic Information Systems) when researching the relationship between geography and loneliness. There are many advantages to using GIS tools in geography research, but the ones that would especially be beneficial to a loneliness research study is that GIS methods allow us to "examine global patterns but meanwhile attend to local uniqueness" and "to recognize scale-dependent and scale-independent geographic features", as well as to "mediate between nomothetic and idiographic approaches to geographic research" (Yuan, 2021, p.25). A GIS analysis would heavily benefit a loneliness research study as it would allow for a zoomed-in analysis of the measure of loneliness in the neighborhood-level, resulting in more specific, efficient mitigation approaches.

2.1.1.1.1 Proposing a New Index

This study is therefore proposing a new index to measure loneliness, which would then, through the tools of GIS, help identify the number of close social infrastructures and ultimately test whether this proximity can predict a low loneliness score as was seen in previous literature.

Rather than directly ask individuals about their feelings of their own loneliness, this index would look at socioeconomic factors that can cause loneliness on a wider, neighborhood-level scale. As Dr. Murthy's report points out, "social connection is more than a personal issue", and therefore shouldn't be only measured as such (Murthy, 2023, p.18). There are many factors of a community that could place it at a higher risk of loneliness. Dr. Murthy's report highlights that groups with "poor physical or mental health, disabilities, financial insecurity, those who live alone, single parents", and "younger and older populations" are much more likely to report feeling lonely than other groups of people (Murthy, 2023, p.19). Overall, loneliness is more frequent in "disadvantaged socioeconomic groups" (Murthy, 2023, p.42).

These findings directly translate to a community's proximity to social infrastructure as well. A community's socioeconomic status can "create additional barriers related to social access" (Reed and Bohr, 2020, p.393). When viewing a neighborhood's proximity to social infrastructure and its reported loneliness through a socioeconomic lens, it can be seen that these factors can lead to a higher reported loneliness, as disadvantaged groups have less access and opportunities to gather in third places, interact with others, and ultimately find and create strong social bonds (Reed and Bohr, 2020).

My new proposed loneliness index scale would bridge together the knowledge of proximity to social infrastructure and measures to predict loneliness to then fill the gap in the missing smaller, neighborhood-level studies of loneliness.

2.1.1.1.1.1 Research Question

With a new loneliness index scale, when observing the Seattle Metropolitan Area and neighborhoods' results on the scale and proximity to social infrastructure, I aim to answer: Is the proximity to social infrastructures in a neighborhood predictive of loneliness in vulnerable populations?

CHAPTER 3:

METHODS

3.1 Purpose of Research

This is an exploratory study as it is creating a brand new index to measure loneliness, which hasn't been proposed before. This deductive nomothetic research study is aiming to provide a better understanding of the loneliness epidemic through a quantitative approach with Seattle as its case study. Creating a new loneliness index will allow for a deeper, finer neighborhood-level analysis of loneliness. While this new index is being applied to Seattle in this study, the goal is for it to be a helpful tool that can be used to analyze loneliness in other major metropolitan areas. Moreover, mapping out and measuring the number of strong social infrastructures that are in neighborhoods, specifically libraries and parks, helps develop a deeper understanding of the relationship between social infrastructures and loneliness.

The U.S Surgeon General's report signaled to researchers that the loneliness epidemic is a major public health issue that needs to be better studied and understood (Murthy, 2023). The purpose of this research study is to provide a nomothetic exploration of Seattle that could then be replicated on a wider scale. The ultimate goal of this study is to identify patterns that could lay the groundwork for future policies and measures to combat this ongoing crisis.

3.1.1 Mode of Observation and Data Collection

This research experiment will be conducted via a GIS model created in the GIS programming software ArcGIS Pro. This quantitative deductive experiment will look at several datasets to firstly identify variables that are predictive of loneliness and create a loneliness index through ArcGIS Pro's Composite Index Tool. It will secondly map strong social infrastructures (libraries and parks) to identify the geographies of these places in relation to Seattle neighborhoods. Lastly, it will perform a GIS statistical buffer analysis to measure the number of close social infrastructures to neighborhoods to identify whether a neighborhood's closer proximity to social infrastructures is indicative of lower levels of loneliness per the index.

The population being sampled in this model is people, specifically people in Seattle. The element in this study is people as well. Almost all of the data in this study comes from the U.S. Census Bureau's American Community Survey's 5-year estimates collected in 2022. It is important to note that this impacts the population in this study, as the American Community Survey sends its surveys to specific addresses, not people (U.S. Census Bureau, 2024). For the purposes of this study, this means that individuals living in Seattle that are without a permanent home address – for instance, those on military duty, or those who are homeless – are excluded from the survey.

The population size of this data is the number of people living in King County, which is 2,254,371 as of 2022 (U.S. Census Bureau, 2022). The number of addresses selected to be surveyed, or the sample size, is 82,823 (U.S. Census Bureau, 2022). The units of analysis in this data are people, specifically those who live in the jurisdiction of Seattle and were counted in the Census, as discussed above. The units of observation in this study are the various datasets being used to put together the index and map social infrastructures. These include the U.S. Census Bureau's American Community Survey's 5-year estimates collected in 2022, specifically the responses to survey question S0101 about Age and Sex, survey question DP02 about Selected Social Characteristics in the United States, and survey question DP03 about Selected Economic Characteristics (U.S. Census Bureau, 2022). The data of Seattle libraries and their locations comes from the Seattle Public Libraries website (Seattle Public Library, 2024). Finally, the data of Seattle parks and their locations comes from the Seattle Parks and Recreation Department (Griggs, 2024).

The U.S. Census Bureau's American Community Survey uses a probability sampling type, specifically a stratified probability sampling strategy. This sampling strategy is quite intricate, with many different layers and statistical analysis to back it up. This study will break down the sampling strategy into an overly simplified explanation. When sampling its counties, the survey has two phases: firstly, it stratifies all census blocks and addresses into sixteen sampling strata, calculates the sampling rates of the strata, and selects the sample of housing unit addresses from the sample (U.S. Census Bureau, 2024). Secondly, it places these sampled addresses within five sub-frame periods within the current year (in this case, 2022), and then systematically selects samples from within these five sub-frames over time (U.S. Census Bureau, 2024).

Overall, each housing unit address has a chance of about 1-in-480 of being selected to complete the survey; if an address is selected, it can not be selected again for a 5-year period (U.S. Census Bureau, 2024).

3.1.1.1 Variables

This research study will create a composite index using 13 variables measured by the U.S. Census Bureau's American Community Survey's 5-year estimates collected in 2022. These variables can be found below.

Table 1. Variables of Study.

Variable	Conceptualization
Total households with a broadband Internet subscription	This indicator variable measures Internet access.
Total households with a male householder, no spouse/partner present, with children of the householder under 18 years	This indicator variable measures the number of single fathers in a population.
Total households with a female householder, no spouse/partner present, with children of the householder under 18 years	This indicator variable measures the number of single mothers in a population.
Total households with a male householder, no spouse/partner present, with householder living alone	This indicator variable measures the number of males living alone.
Total households with a female householder, no spouse/partner present, with householder living alone	This indicator variable measures the number of females living alone.
Total households with a male householder, no spouse/partner present, with householder living alone and 65 years and over	This indicator variable measures the number of elderly males living alone.
Total households with a female householder, no spouse/partner present, with householder living alone and 65 years and over	This indicator variable measures the number of elderly females living alone.
Total Civilian Noninstitutionalized Population with a disability	This indicator variable measures the number of people in a population with a disability.
Total population 18 to 24 years	This indicator variable measures the number of young people in a population.
Total population 65 years and over	This indicator variable measures the number of elderly people in a population.
Percentage of all people whose income in the past 12 months is below the poverty level	This indicator variable measures the percentage of people living below the poverty level.

Table 1 (cont.)

Civilian labor force unemployment rate	This indicator variable measures the percentage of people who are able to work who are unemployed, which is predicative of loneliness.
Total households that are a married-couple household	This indicator variable measures the number of married people in a population, which is predicative of loneliness.

Indicator variables that are measuring the total number of that variable in the population (Variables 1-9 and 13 in *Table 1*) were measured by counting the number of responses that indicated this variable as applicable to the respondent. Indicator variables that are measuring the percentage or rate of that variable in the population (Variables 11 and 12 in *Table 1*) were measured by calculating the number of responses that indicated this variable as applicable to the respondent divided by the total number of respondents of the survey. Each variable listed is measured on the Census tract level in King County.

Variable 1 in *Table 1*, which is measuring household Internet access, is an independent variable measured on a nominal scale; respondents either indicated that they do or do not have a broadband Internet subscription. Previous literature has shown that the lack of Internet access indicates a higher risk of loneliness (Shovestul *et al*, 2020; Crowe *et al*, 2024; Jing *et al*, 2024). The face validity of this variable is high as it is a common sense way to measure Internet access. The construct validity of this variable is high, as the theoretical concept of Internet access is fully measured by the existence or lack thereof of a broadband subscription. The content validity of this variable is low, as it is a pre-coded, close-ended question that doesn't allow for a wide range of responses – for instance, households that don't have home Internet access might still have access through a local library or work, resulting in lower reported loneliness regardless of this measure. The variable is clear and concise, making it reliable between researchers, between the researcher and researched, and if used in future research. As mentioned in the content validity of this variable, the limited response options of this variable make it not precise enough. However, in a broad, nomothetic context, its accuracy is high as it can reflect wider Internet access patterns outside of King County.

Both Variable 2 and 3 in *Table 1* are measuring the number of single parents (fathers and mothers, respectively) in each Census tract. These variables are independent and measured on a nominal scale – respondents either indicated that they are a male householder living alone with children or not (Variable 2), and that they are a female householder living alone with children or not (Variable 3). While single parenthood is a less common indicator of loneliness, some previous literature has pointed to this trend (Murthy, 2023; Mah *et al*, 2023). The face validity of these variables is high as it is a common sense way to measure a demographic vulnerability. The construct validity of this variable is high, as it addresses multiple aspects of a person's identity – whether they are a male householder, whether they are single, and whether they have children that are younger than 18 years old, which fully captures the theoretical framework of single parenthood. The content validity of this variable is low, as respondents could fall in gray-area cases that aren't covered by the pre-coded options given to them; for instance, respondents could be part-time single parents, with a partner some of the time. Due to this variable's specificity, its precision is quite high; it is also most likely accurate as it can reflect wider trends of single parenthood.

Variables 4 and 5 in *Table 1* measure the number of people (male and female, respectively) that are living alone. With both physical and mental health implications, living alone has become a significant indicator of loneliness in recent literature (Murthy, 2023; Shaw *et al*, 2021; Swader, 2018; Algren *et al*, 2020; Stickley and Ueda, 2022; Lazzari and Rabottini, 2022; Mann *et al*, 2022; Prohaska *et al*, 2020; Rohr *et al*, 2021). In this study, this variable is independent and measured on a nominal scale – respondents either indicated that they are a single male living alone or not (Variable 4), and that they are a single female living alone or not (Variable 5). The face validity of this variable is high as it is a common sense way to measure single living. Similarly to Variables 2 and 3, the construct validity of this variable is high as it fully captures the larger theoretical measure of living alone as it breaks down multiple parts of the respondent's identity – being a male or female householder, with no spouse or partner present, who is living alone. The content validity of this variable is low, as it doesn't leave room for other case scenarios, such as people who live alone some of the time but not all of the time. Overall, this variable is quite precise and specific; its accuracy is high as well as it can reflect wider national trends.

Both Variables 6 and 7 in *Table 1* are similar to Variables 4 and 5 except that they also take age into account, as they measure the singlehood of the elderly population which is defined as aged 65 and older. The combination of these two demographic factors (living alone and elderly adults) has not been as widely studied; however, some literature exists suggesting a higher risk of loneliness in older adults living alone (Kim and Kim, 2024). Variables 4 and 5 are independent and measured on a nominal scale – respondents either indicated that they are an elderly male living alone or not (Variable 6), and that they are an elderly female living alone or not (Variable 7). These variables' face validity is high as they are a common sense way to measure the number of older adults living alone. Their construct validity of these variables is high as they fully encompass the conceptualization of the measure of older adults living alone, breaking down multiple parts of the respondent's identity – being age 65 or older, and living alone. Their content validity is low as, similarly to Variables 4 and 5, they don't leave room for other scenarios, such as an elderly person who is living alone some but not all of the time. These variables are quite specific and therefore precise; they are also accurate as they can reflect larger trends of older adults living alone.

Variable 8 in *Table 1* is measuring the number of people with a disability by Census tract. Disability has been widely studied as a major risk factor of high loneliness (Murthy, 2023; Chien *et al*, 2024; Algren *et al*, 2020; Stickley and Ueda, 2022; Mann *et al*, 2022; Rohr *et al*, 2022; Reed and Bohr, 2020). This variable is an independent variable measured on a nominal scale – respondents either indicated that they do or do not have a disability. It has a high face validity as it is a clear question which makes sense – respondents can self-identify their disability status. Its construct validity is high as it fully captures the specific theoretical concept of a disability status. Its content validity is low, as it is a black-and-white measure of disability; a respondent that has a minor learning disability might not face the same loneliness that a person with a physical chronic disability might face. This variable is a specific and therefore precise measure, also making it accurate as it is a good indicator of wider trends.

Both Variable 9 and 10 in *Table 1* are measuring the age of the respondents, but through specific age range buckets meant to measure the number of young adults and the number of elderly people in a Census tract. Plenty of previous literature has indicated that both young adult are more likely to report loneliness (Stickley and Ueda, 2022; Crowe *et al*, 2022; Satcher, Okafor, and Dill, 2012; Mann *et al*, 2022; Prohaska *et al*, 2020; Shovestul *et al*, 2020; Rohr *et al*, 2021; Reed and Bohr, 2020; Kim and Kim 2024) and older adults are more likely to report loneliness (Prohaska *et al*, 2020; Haggerty, Minotti, and Bouharaoui, 2023; Mah *et al*, 2023; Hughes *et al*, 2004; Rohr *et al*, 2021; Villeneuve *et al*, 2023; Lane *et al*, 2020; Jing *et al*, 2023). These independent variables are measured on an ordinal scale, as each age bucket has a specific range (18-24 years old for young age, and 65 and older for old age). They have a high face validity as they are a common sense way to measure age. They have a high construct validity as they fully capture the straight-forward measure of young and old age. They have a high content validity as the rest of the measure, which isn't being used in this study, provides the full scale of age range buckets that a person could fall under – every participant can indicate that they belong into one of the age buckets provided. They are precise age ranges that are accurate and reflect wider trends of young and old age ranges.

Variable 11 in *Table 1*, which is measuring the percentage of people earning an income below the poverty level, is an independent variable being measured on a ratio scale – there exists a "true zero", as no Census tract can report having less than zero percent of people living below the poverty level. Previous literature has indicated that low socioeconomic status is a risk of higher reported loneliness (Stickley and Ueda, 2022; Algren *et al*, 2020; Murthy, 2023; Chien *et al*, 2024; Mah *et al*, 2023; Kim and Kim, 2024). The variable's face validity is high as it is a common sense way to measure socioeconomic status. Its criterion validity is high as it is relying on the well-known federal-based poverty level scale. Its construct validity is high as it fully captures the theoretical concept of whether or not an individual is facing poverty. Its content validity is low, since it is lacking a more nuanced approach to socioeconomic status; respondents that have only recently been living under the poverty level are most likely facing different loneliness rates than those who have lived in chronic poverty all of their lives. This variable is a precise, accurate measure of those living below the poverty line.

Variable 12 in *Table 1*, which is measuring the percentage of people who are unemployed, is an independent variable being measured on a ratio scale – there exists a "true zero", because no Census tract can report an unemployment rate of less than zero percent. The employment status of an individual is an indicator of loneliness, with specifically unemployment placing an individual under high risk of loneliness (Algren *et al*, 2020; Kim and Kim, 2024; Hanibuchi *et al*, 2012; Stickley and Ueda, 2022; Rohr *et al*, 2022; Villeneuve *et al*, 2024; Lane *et* *al*, 2020; Reed and Bohr, 2021; Shaw *et al*, 2021). The variable's face validity is high as it is a common sense way to measure socioeconomic status. Its construct validity is high as it fully captures the theoretical concept of employment, with individuals being able to indicate whether they are or aren't employed. Its content validity is low, since it isn't capturing a complete measure of an individual's employment status; an individual who was recently laid off and is actively looking for work could be facing different rates of loneliness than an individual who has never been employed. This makes it an accurate measure, but not a precise one.

Lastly, Variable 13 in *Table 1* is measuring the number of married couples and is an independent variable measured on a nominal scale – a coupled household either indicated that they are married or are not married. Previous literature has identified marital status as an indicator of reported loneliness (Murthy, 2023; Steptoe *et al*, 2013; Prohaska *et al*, 2020; Villeneuve *et al*, 2023; Lane *et al*, 2020; Reed and Bohr, 2020; Kim and Kim, 2024; Shaw *et al*, 2021). The face validity of this variable is high as it is a common sense way to measure marital status. The construct validity of this variable is high as it fully captures the conceptualization of marital status, with survey takers being able to self-report whether they are or aren't married. The content validity of this variable is low, as it is a pre-coded, close-ended question that doesn't allow for a wide range of responses – for instance, individuals that are engaged and soon-to-be married aren't captured by this variable. This makes it an accurate but not precise measure.

All of the variables listed in *Table 1* have a high reliability between researchers and between researcher and future self. The American Community Survey is a well-known, well-established survey that has conceptualized and operationalized its variables to a high degree; these variables are well-understood among researchers, especially those in the field of geography. Furthermore, the U.S. Census Bureau does not change the meaning or concepts of its survey variables over time. This means that if this research were conducted again in the future, these variables would most likely get the same results.

All of the variables that make up the composite index are independent in terms of the index; however, once they are all calculated into the single index, the loneliness index itself would be a dependent variable measured on an ordinal scale of low, medium, and high vulnerability. Each Seattle neighborhood will be ascribed this index scale and will be labeled based on its loneliness score on the scale.

The independent variables whose impact on the loneliness index is being studied in this research are parks and libraries. Previous literature has shown that the proximity to third places is predicative of loneliness, specifically parks (Murthy 2023; Hanibuchi *et al*, 2012; Villeneuve *et al*, 2023; Lane *et al*, 2020; Reed and Bohr, 2020; Coll-Planas *et al*, 2024; Finlay *et al*, 2019; Jing *et al*, 2024) and libraries (Satcher, Okafor, and Dill, 2012; Lane *et al*, 2020; Reed and Bohr, 2020; Kim and Kim, 2024; Finlay *et al*, 2019; Jing *et al*, 2024). The parks dataset comes from the City of Seattle's Department of Parks and Recreation database (Griggs, 2024), while the library dataset comes from the City of Seattle's Libraries website (Seattle Public Library, 2024). Both of these variables are being measured on a ratio scale. Their face validity is high, as the numbers of parks and libraries is a common sense way of measuring third places in a city. Their construct validity is high as they are specific, singular measures – there either is a park or library in a neighborhood, or there isn't. Due to the same reason, the content validity for these variables is high. They are both a precise and accurate measure, making them also a reliable one.

3.1.1.1.1 Mode of Analysis

To analyze whether a neighborhood's proximity to parks and libraries makes it predicative of a low score on the new loneliness index scale, I will be creating a model using ArcGIS Pro's analysis tools. I will use ArcGIS Pro's Calculate Composite Index tool to combine the 13 variables listed in *Table 1* into a singular variable. This tool works in multiple steps: first, I'll preprocess the variables by deciding whether scoring highly on a variable would indicate more or less loneliness (for instance, having a higher percentage of unemployment would indicate higher loneliness, while having a lower marital rate would indicate higher loneliness). Secondly, I'll need to specify the weighting of each variable to decide which variables I'd want to be taken into a larger consideration in the index. For instance, I will choose to weigh the percentage of those living below the poverty line as a more significant weight, as there exists significantly more literature demonstrating the impact of this measure on loneliness. Lastly, I'll use the z-score scaling method to scale all of the variables so they fall between 0 and 1. This will make it easier to create three distinct class values of a low, medium, and high rate of loneliness, which will then be ascribed to each Seattle neighborhood based on where each neighborhood falls on the scale.

I will add in my independent variables by mapping them in ArcGIS Pro. I will geocode my parks and libraries dataset so that I can plot each park and library spatially on my map. I will then utilize the Select By Location tool in ArcGIS Pro to identify which neighborhoods are near the most parks and libraries, and which aren't. I'll create a buffer radius to help visualize this finding, and layer both this buffer and the choropleth results of the loneliness index scale of each neighborhood together in an interactive, web-based map for wider accessibility.

CHAPTER 4:
DISCUSSION OF FINDINGS

4.1 Discussion of Findings

This research paper aimed to answer the following research question: Is the proximity to social infrastructures in a neighborhood predictive of loneliness in vulnerable populations? To answer this question, two loneliness indexes were created – one unweighted (Figure 1) and one weighted (Figure 5) – to predict the vulnerability to loneliness of Seattle neighborhoods on a smaller scale than previously available. The two indexes revealed similar insights, finding that the proximity to libraries have no impact on a neighborhood's reported loneliness on the new index, both weighted and unweighted, while the proximity to parks have some impact on a neighborhood's reported loneliness on the new index, both weighted and unweighted. Finally, there was not a significant difference across the weighted and unweighted indexes, leading to further questions about which variables proved to be the most influential in the index and which new variables could be measured to calculate a more accurate, predictive loneliness index.

When looking at the unweighted index, meaning every variable of the 13 measured took on the same weight of 1 when calculating the index, it appears that many census tracts have similar reported levels of loneliness, mostly leaning towards little to low reported loneliness. In North Seattle and the University District, a few census tracts report higher loneliness and have little to no parks. The opposite is true in Downtown Seattle, where there are multiple census tracts reporting high rates of loneliness but there are numerous parks in each tract. A full breakdown of the distribution of the unweighted index is seen in Figure 2. The mean report falls on 0, and most census tracts fall in the categories of very low reported loneliness to somewhat low reported loneliness. Most census tracts fall into a loneliness of -0.3 on the index, meaning a low reported loneliness. It is worth noting that hundreds of census tracts fall above the mean of 0 on the index, meaning there is some prediction of loneliness in many tracts.

This raises the question of how each variable is impacting the unweighted index. The relationships between the 13 variables and their influence on the unweighted index mean can be observed in Figure 3 below. The variables with the most influence include the percentage of females living alone in a census tract, as well as the percentage of married couples in a census tract, with correlations of 0.62 and 0.73 respectively. The variables with the least influence include the percentages of single fathers and mothers in a census tract, as well as the percentages of people above the age of 65 years old.

A weighted index was created to make a comparison between the benchmark unweighted index and to apply the findings from the literature review which indicated some variables having more impact than others on an individual's loneliness. The weights assigned to each variable in the weighted index can be seen in Figure 4.

The 5 variables assigned a higher weight than the standard 1 were the percentage of females living alone, the percentage of males living alone, the percentage of people living in poverty, the percentage of females above the age of 65 living alone, and the percentage of males above the age of 65 living alone; the former three received a weight of 2 while the latter two received a weight of 3. The plotted weighted index can be seen below in Figure 5.

As observed in the figures below, there are little to no significant differences between the weighted and unweighted indexes. The one observable different that can be made is that the census tracts that were already reported a higher rate of loneliness, such as in North Seattle and in Downtown Seattle, have a higher reported loneliness indicated by a darker shade of purple, and other census tracts that were reporting a somewhat low rate of loneliness are now falling under a higher loneliness rate on the weighted index. Similarly to the observation of the unweighted index, the census tracts reporting high loneliness rates are not in proximity to many parks, with Downtown Seattle being the exception. Still, the correlation to a library is not clear as these census tracts are in proximity to a library. A breakdown of the distribution of the weighted index can be seen below in Figure 6.

Similarly to the distribution of the unweighted index, the majority of the census tracts fall either slightly above or slightly below the mean. However, in the weighted index, there are more census tracts falling above the mean, with more outliers in higher reported loneliness. Still, most census tracts are falling on -0.3 of the index, meaning most are reporting little to low loneliness rates. The relationships between the distinct variables and their impact on the weighted index can be seen below in Figure 7.

There are more variables correlated to the weighted index than the unweighted index. The variables with the highest correlation to the index mean are the percentage of females living alone, the percentage of males living alone, the percentage of females above the age of 65 living alone, and the percentage of males above the age of 65 living alone, with correlations of 0.68, 0.57, 0.63, and 0.61 respectively. Similarly to the unweighted index, the percentage of people with a disability also has a high correlation to the weighted index, although less so than the unweighted index. Still, the variable with the highest correlation to the weighted index is the percentage of married couples in a census tract.

Overall, the findings reveal that a loneliness index is a complex idea that requires more nuance and discussion than this research project was able to provide. While some findings revealed that the lack of proximity to parks resulted in census tracts reporting higher levels of loneliness, this wasn't consistent for the entire city. Moreover, the distribution of both weighted and unweighted indexes revealed that many census tracts are similar and fall under little to low loneliness reported, while there are some outliers that report higher levels of loneliness (made more clear by the weighted index). This reveals that the usage of the index can be applied to very specific, targeted approaches of mediating loneliness. By identifying these singular vulnerable census tracts, the mitigation techniques are narrowed down and more manageable than approaching individual-based reports of loneliness.

Additionally, this analysis revealed that not all 13 variables identified earlier in this study are strongly correlated in predicting a loneliness index, and that more analysis can be used to narrow down the distinct social characteristics that could result in reported loneliness. For instance, the single status of females or males in a census tract had little to no correlation to both weighted and unweighted indexes, while the percentage of married couples had a significant correlation to reported loneliness. This was also true of the unemployment rate of a census tract, which had little to no correlation to the weighted and unweighted indexes, while the percentage of people living in poverty had significant correlation to the indexes. By narrowing down the social health variables that impact the likelihood of higher reported loneliness on the index, the mitigation approaches to loneliness can again be more specific and consequently helpful in reducing experienced loneliness. This can also inspire further research into the linkage of specific variables to loneliness, for instance by observing the relationship of census tracts with high marriage rates, low marriage rates, and loneliness.

Ultimately, the two indexes produced are only the beginning of a larger discussion about loneliness and how people experience it across the United States, and are riddled with many issues that can be addressed in future research. As seen in the figures below, several variables in the index were found not to be correlated to loneliness, and would need to be removed from future indexes to gain a more accurate index. Furthermore, the specific relationship between social infrastructure and reported loneliness is still unclear, especially with public social infrastructures such as libraries and parks. There are some ways to mitigate these issues, however, and continue to build upon this research; these ideas are detailed in the following section.

4.2 Figures and Tables

Figure 1. A Map of the Unweighted Loneliness Index

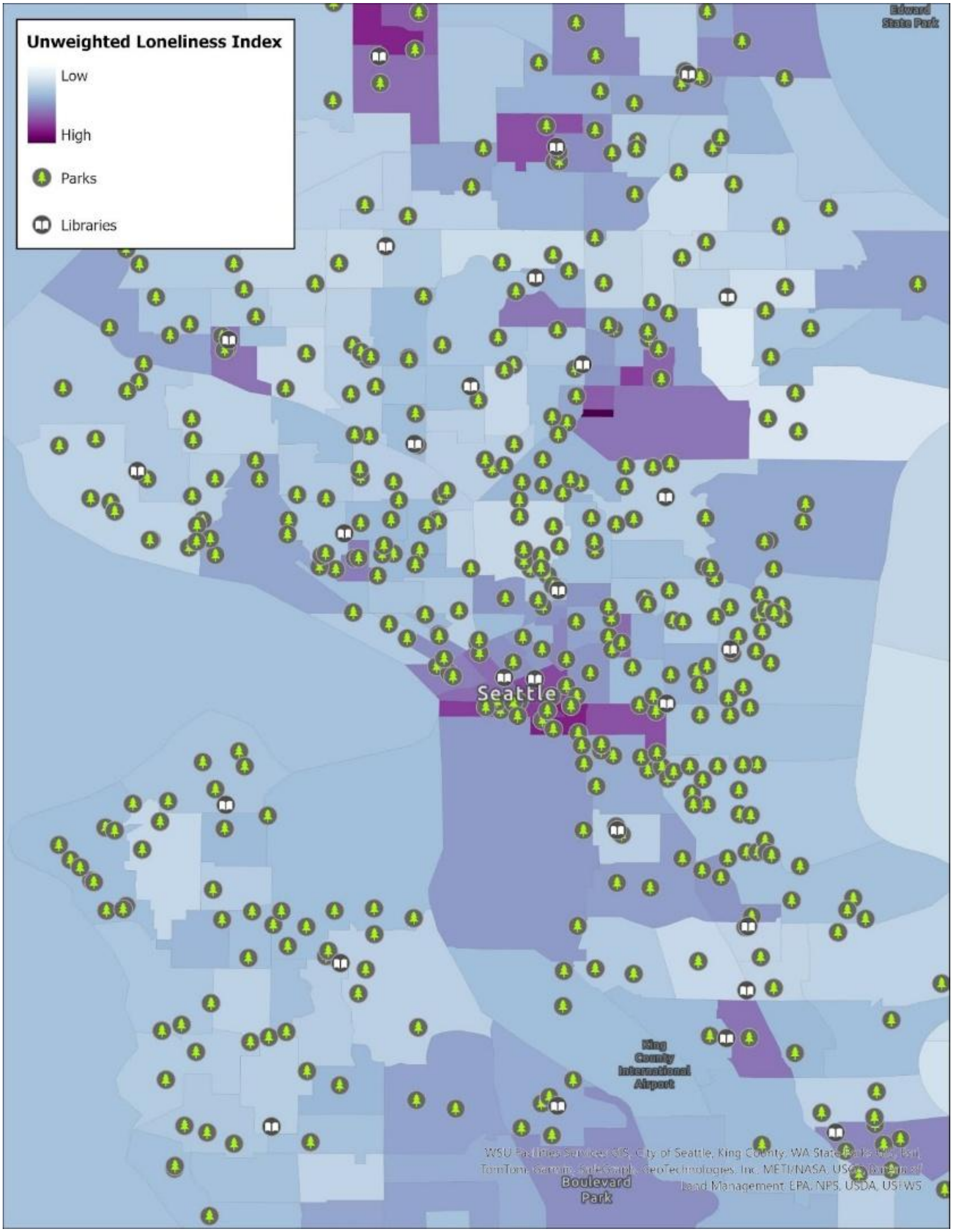


Figure 2. A Chart of the Distribution of the Unweighted Index

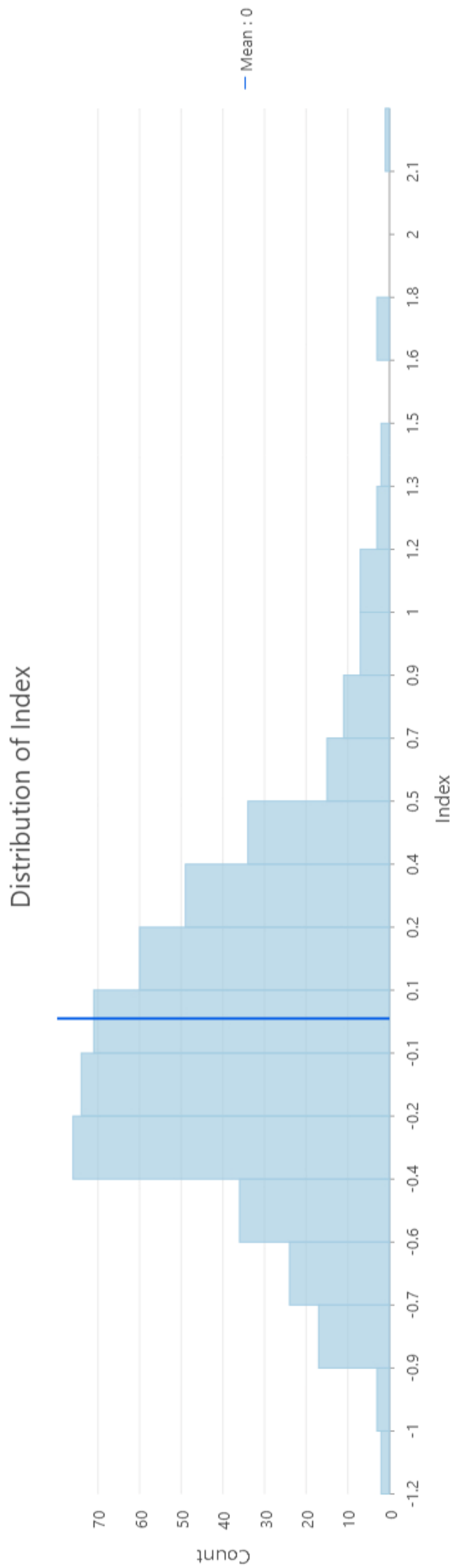


Figure 3. A Chart of the Relationships of Scaled Variables and Unweighted Index

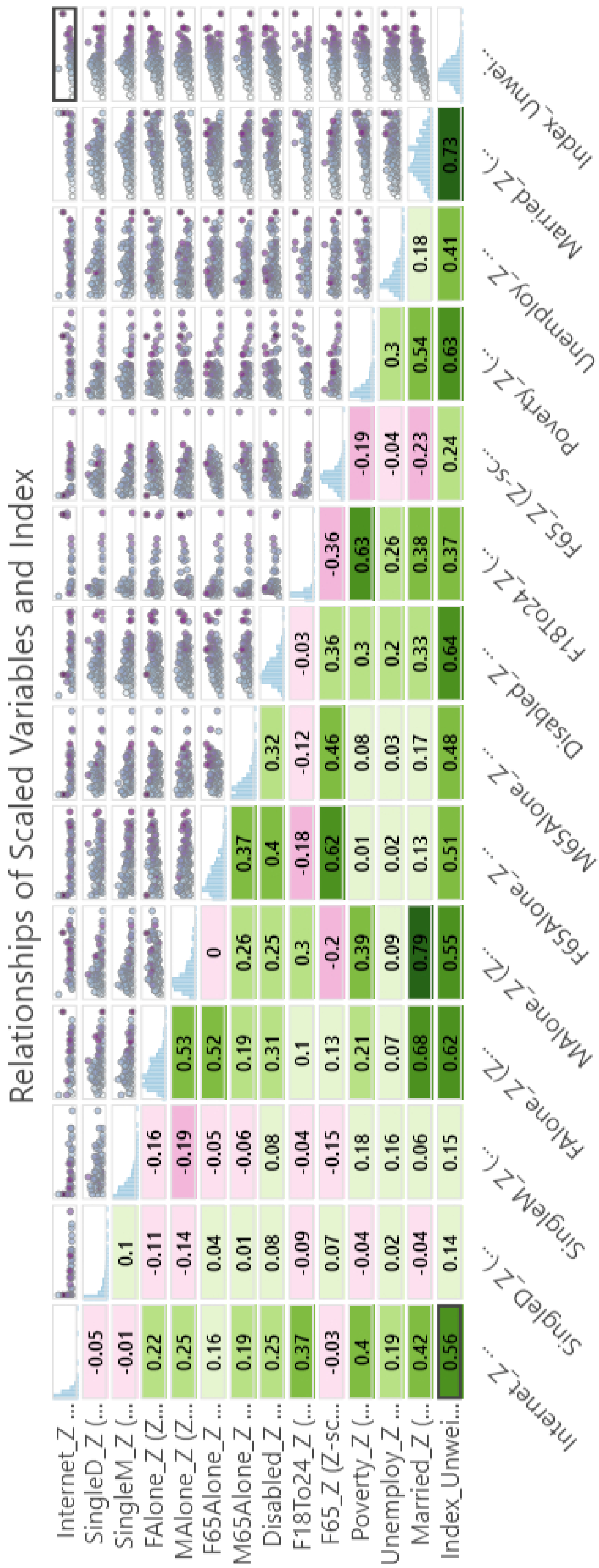


Figure 4. A Table of the Variables and their Corresponding Weights

Field	Weight
Internet_Z	1
SingleD_Z	1
SingleM_Z	1
FAlone_Z	2
MAlone_Z	2
F65Alone_Z	3
M65Alone_Z	3
Disabled_Z	1
F18To24_Z	1
F65_Z	1
Poverty_Z	2
Unemploy_Z	1
Married_Z	1

Add another

Figure 5. A Map of the Weighted Loneliness Index

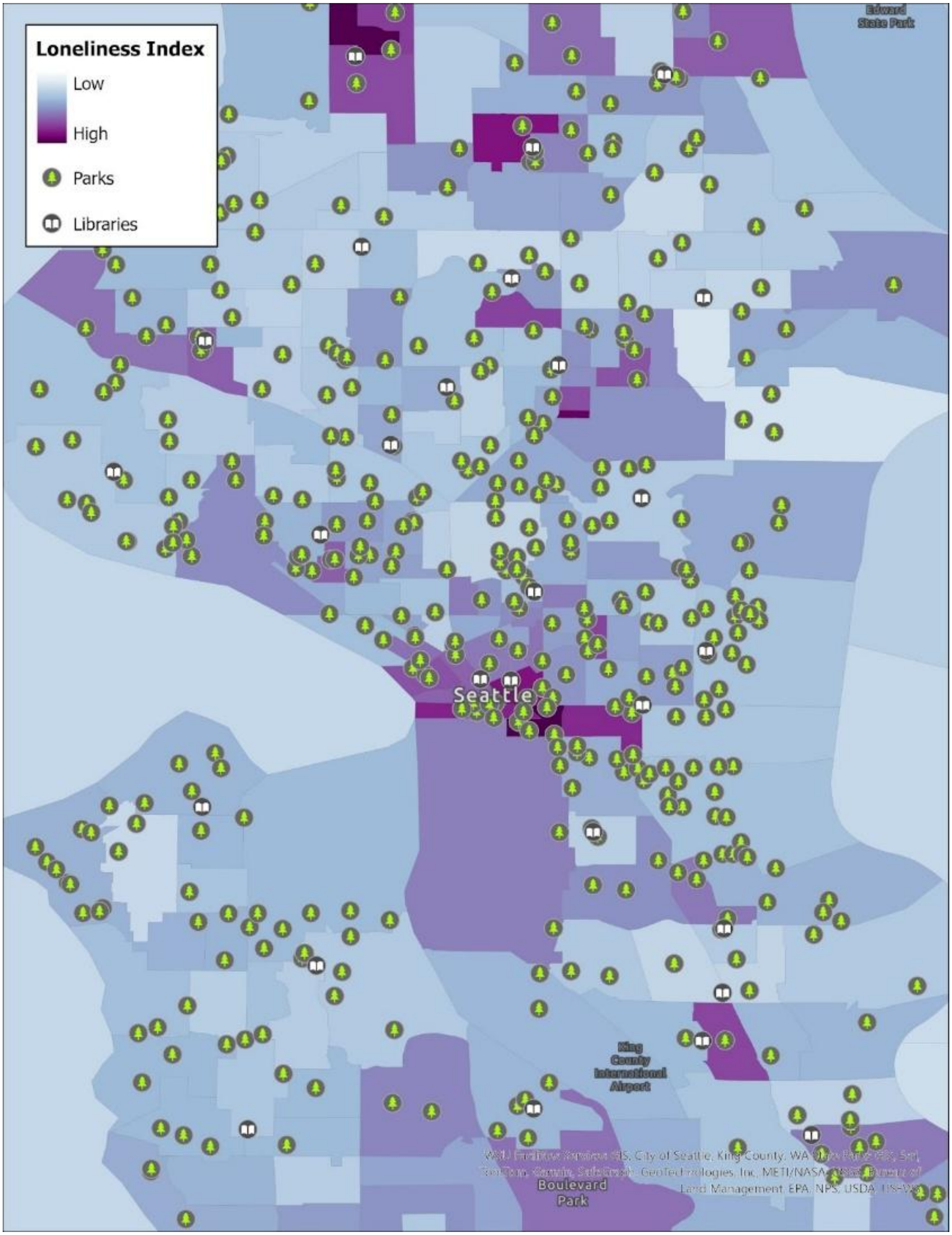


Figure 6. A Chart of the Distribution of the Weighted Index

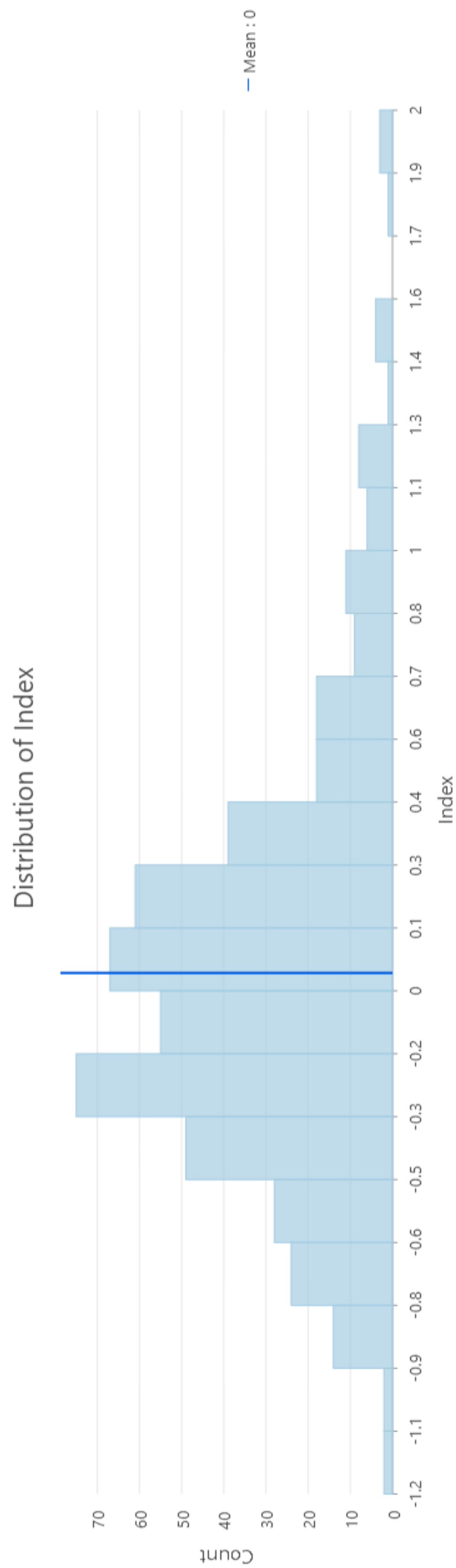
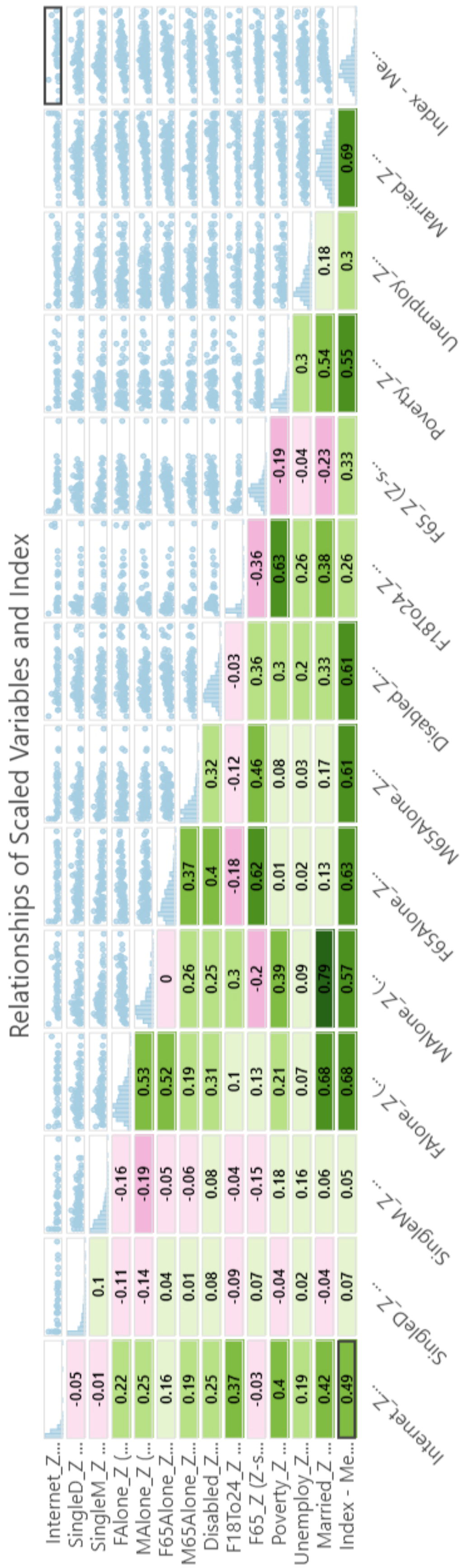


Figure 7. A Chart of the Relationships of the Scaled Variables and Weighted Index



CHAPTER 5:
CONCLUSION

Overall, this project was meant to be a building block in a larger discussion about loneliness and its impact on our society in relation to our geography. A strong intervention technique meant to lower reported loneliness is finding a physical place that helps create a sense of community and connectedness, which could be social infrastructures. This project aimed to map these social infrastructures, as well as propose a new loneliness index: a new way to measure loneliness on a smaller, census tract scale, to further the discussion about loneliness and demonstrate a new way to measure the serious public health issue. While loneliness is experienced individually and varies from person to person, previous literature has demonstrated that there are universal measures that could result in a higher experienced loneliness. Identifying these root causes and measuring their impact in a composite index could help identify the specific geographical locations in need of mitigation techniques and resources and produce a more specific, relevant approach to the matter.

Beyond narrowing down the variables used in the index, finding new ways to weight the index appropriately, measuring other social infrastructures beyond parks and libraries, and specifying how to measure the "proximity" to social infrastructures, future research should include ground-truthing into its methodology. Ground-truthing is the popular social science and humanities research process of physically visiting the geographical sites a researcher is studying to check their analysis against. In this case, future researchers should visit the census tracts reporting higher levels of loneliness to understand how the communities in these census tracts are actually experiencing the public health issues on an individual and communal level. How is loneliness showing up in the lived experiences of these communities, and how are they mitigating loneliness? Is loneliness a major issue they would name, or are there other public health issues they find more pressing? These questions would significantly improve the findings of this research as they can help check whether the findings revealed in this research are accurate, and can help identify specific mitigation techniques that would truly benefit these communities. Ground-truthing is deeply important to this kind of social science geographical analysis and should not be ignored or pushed to the side.

Ultimately, the indexes proposed in this analysis are just one step in creating a larger, more nuanced conversation about loneliness. As previous research has shown, loneliness is a major public health issue that needs to be addressed sooner rather than later. When we predict what could make communities lonely before they reach their breaking point, we can integrate and execute mitigating strategies that are effective and see truly transformational results. The goal of this study was to inspire future loneliness indexes to be made around other major metropolitan areas in the United States. As we continue to analyze and better understand this complex public health issue, we should hope to see decreased reports of loneliness.

Works Cited

Abshire, A. D. *et al.* (2020) 'Differences in Loneliness Across the Rural-Urban Continuum Among Adults Living in Washington State', *The Journal of Rural Health*, 38, pp. 187-193. doi: https://doi.org/10.1111/jrh.12535

Algren, M. H. et al. (2020) 'Social isolation, loneliness, socioeconomic status, and health-risk behaviour in deprived neighbourhoods in Denmark: A cross-sectional study', SSM - Population Health, 10, pp. 1-10. doi: https://doi.org/10.1016/j.ssmph.2020.100546

Chien, L-C. et al. (2024) 'Exploring the role of the social vulnerability index in understanding COVID-19 immunization rates', PLOS ONE, 19(6), pp. 1. doi: https://doi.org/10.1371/journal.pone.0302934

Coll-Planas, L. *et al*. (2024) 'Nature-based social interventions to address loneliness among vulnerable populations: a common study protocol for three related randomized controlled trials in Barcelona, Helsinki, and Prague within the RECETAS European project', *BMC Public Health*, 24, pp. 172. doi: https://doi.org/10.1186/s12889-023-17547-x

Crowe, C. L. *et al.* (2024) 'Loneliness prevention and the role of the Public Health system', *Perspectives in Public Health*, 144, pp. 31-38. doi: https://doi.org/10.1177/17579139221106579

Das, A. *et al.* (2021) 'A systematic review of loneliness and social isolation scales used in epidemics and pandemics', *Psychiatry Research*, 306, pp. 1. doi: https://doi.org/10.1016/j.psychres.2021.114217

Finlay, J. *et al.* (2019) 'Closure of 'third places'? Exploring potential consequences for collective health and wellbeing', *Health & Place*, 60, pp. 1-5. doi: https://doi.org/10.17269/s41997-023-00841-x

Griggs, K. (2024) 'Seattle Parks and Recreation Park Addresses'. Available at: https://data.seattle.gov/Community-and-Culture/Seattle-Parks-And-Recreation-Park-Addresses/v5tj-kqhc/about_data (Accessed November 2024).

Haggerty, J., Minotti, S. C., and Bouharaoui, F. (2023) 'Development of an individual index of social vulnerability that predicts negative healthcare events: a proposed tool to address healthcare equity in primary care research and practice', *International Journal for Equity in Health*, 22, pp. 157. doi: https://doi.org/10.1186/s12939-023-01965-9

Hanibuchi, T. *et al*. (2012) 'Does walkable mean sociable? Neighborhood determinants of social capital among older adults in Japan', *Health & Place*, 18, pp. 229-239. doi: https://doi:10.1016/j.healthplace.2011.09.015

Hughes, M. E. *et al*. (2004) 'A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies', *Res Aging*, 26(6), pp. 655-672.

Jing, J. *et al.* (2024) ‘The Role of Third Place concerning Loneliness in the Context of Ageing in Place: Three Neighbourhoods in Stockholm’, *Health & Social Care in the Community*, 2024, pp. 1-17. doi: https://doi.org/10.1155/2024/4172682

Kearns, A. *et al*. (2015) “‘Lonesome Town’? Is Loneliness Associated With The Residential Environment, Including Housing And Neighborhood Factors?’, Journal of Community Psychology, 43, pp. 849-867. doi: https://doi.org/10.1002/jcop.21711

Kim, Y.-K. and Kim, D. (2024) ‘Role of Social Infrastructure in Social Isolation within Urban Communities’, *Land*, 13, pp. 1-13. doi: https://doi.org/10.3390/land13081260

Lane, A. P. *et al.* (2020) ‘Cross-sectional associations of neighborhood third places with social health among community-dwelling older adults’, *Social Science & Medicine*, 258, pp. 1-8. doi: https://doi.org/10.1016/j.socscimed.2020.113057

Lazzari, C. and Rabottini, M. (2021) ‘COVID-19, loneliness, social isolation and risk of dementia in older people: a systematic review and meta-analysis of the relevant literature’, *International Journal of Psychiatry in Clinical Practice*, 26(2), pp. 196-207. doi: https://doi.org/10.1080/13651501.2021.1959616

Mah, J. C., *et al*. (2023) ‘Social vulnerability indices: a scoping review’, *BMC Public Health*, 23, pp. 1253. doi: https://doi.org/10.1186/s12889-023-16097-6

Mann, F. *et al*. (2022) ‘Loneliness and the onset of new mental health problems in the general population’, Social Psychiatry and Psychiatric Epidemiology, 57, pp. 2161-2178. doi: https://doi.org/10.1007/s00127-022-02261-7

Murthy, V. H. (2023) *Our Epidemic of Loneliness and Isolation: The U.S. Surgeon General’s Advisory on the Healing Effects of Social Connection and Community*. Available at: https://www.hhs.gov/sites/default/files/surgeon-general-social-connection-advisory.pdf (Accessed: 6 October 2024).

Prohaska, T. *et al*. (2020) ‘Consensus statement: loneliness in older adults, the 21st century social determinant of health?’, BMJ Open, 10, pp. 1-6. doi: https://doi.org/10.1136/bmjopen-2019-034967

Reed, S. and Bohr, J. (2020) ‘The Influence of Local Built Environments on Social Wellbeing: a Community’s Experience with Social Isolation, Social Loneliness, and Social Belonging’, *International Journal of Community Well-Being*, 4(3), pp. 393-413. doi: https://doi.org/10.1007/s42413-020-00091-6

Rohr, S. *et al*. (2021) ‘Social factors and the prevalence of social isolation in a population-based adult cohort’, *Social Psychiatry and Psychiatric Epidemiology*, 57, pp. 1959-1968. doi: https://doi.org/10.1007/s00127-021-02174-x

Russell, D., Peplau, L. A., and Ferguson, M. L. (1978) ‘Developing a measure of loneliness’, *Journal of Personality Assessment*, 42, pp. 290-294.

Satcher, D., Okafor, M., and Dill, L. J. (2012) ‘Impact of the Built Environment on Mental and Sexual Health: Policy Implications and Recommendations’, ISRN Public Health, 2012, pp. 1-7. doi: https://doi.org/10.5402/2012/806792

Shaw, R. J. *et al*. (2021) ‘Living alone, loneliness and lack of emotional support as predicators of suicide and self harm: A nine-year follow up of the UK Biobank cohort’, *Journal of Affective Disorders*, 297, pp. 316-323. doi: https://doi.org/10.1016/j.jad.2020.10.026

Shovestul, B. *et al*. (2020) ‘Risk factors of loneliness: The high relative importance of age versus other factors’, *PLOS ONE*, 15(2), pp. 1. doi: https://doi.org/10.1371/journal.pone.0229087

Steptoe, A. *et al*. (2013) ‘Social isolation, loneliness, and all-cause mortality in older men and women’, PNAS, 110(15), pp. 5797-5801. doi: https://doi.org/10.1073/pnas.1219686110

Stickley, A. and Ueda, M. (2021) ‘Loneliness in Japan during the COVID-19 pandemic: Prevalence, correlates and association with mental health’, Psychiatry Research, 3071, pp. 1-6. doi: https://doi.org/10.1016/j.psychres.2021.114318

Swader, C. S. (2018) ‘Loneliness in Europe: Personal and Societal Individualism-Collectivism and Their Connection to Social Isolation’, Social Forces, 97(3), pp. 1307-1336. doi: https://doi.org/10.1093/sf/soy088

Swader, C. S. *et al.* (2023) ‘Social Infrastructure and the Alleviation of Loneliness in Europe’, *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 75, pp. 387-414. doi: https://doi.org/10.1007/s11577-023-00883-6

The Seattle Public Library (2024) *Hours & Locations*. Available at: https://www.spl.org/hours-and-locations (Accessed November 2024).

U.S. Census Bureau (2024) *Household Pulse Survey for U.S., Washington State, and Seattle Metropolitan Area*. Available at https://www.census.gov/programs-surveys/household-pulse-survey/data/tables.html (Accessed: 20 October 2024).

Villeneuve, P. J. *et al*. (2024) ‘Does urban greenness reduce loneliness and social isolation among Canadians? A cross-sectional study of middle-aged and older adults of the Canadian Longitudinal Study on Aging (CLSA)’, *Canadian Journal of Public Health*, 115, pp. 282-295. doi: https://doi.org/10.17269/s41997-023-00841-x

Yuan, M. (2020) ‘GIS research to address tensions in geography’, *Singapore Journal of Tropical Geography*, 42, pp. 13-30. doi: https://doi.org/10.1111/sjtg.12344