An Aggregate Measure of Bicycle Commuting and its Relationship with Heart Disease Prevalence in the United States

Emma C. Todoroff

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Civil Engineering

Frederick E. Paige, Co-chair

Tripp Shealy, Co-chair

Charlotte Baker

Courtney Coughenour

October 22, 2021

Blacksburg, Virginia

Keywords: bicycle commuting, heart disease, urban transport, quantitative analysis

An Aggregate Measure of Bicycle Commuting and its Relationship with Heart Disease Prevalence in the United States

Emma C. Todoroff

ABSTRACT

United States bicycle commuting rates are low compared to similarly developed countries like the Netherlands and Denmark. However, bicycle commuting shows promise for positive health outcomes, especially those related to chronic diseases like obesity, diabetes, and heart disease. Little research has been conducted in the U.S. to study the association between bicycle commuting and heart disease. Furthermore, U.S. cities need guidance on how to increase bicycle commuting rates.

The purpose of this study was to evaluate the association between U.S. bicycle commuting rates and heart disease prevalence and to identify infrastructure and policy factors most significantly associated with bicycle commuting rates in large U.S. cities. This research quantitatively defined infrastructure and policy factors and analyzed ecologic associations across the 50 most populous U.S. cities.

The results of this study are based on an ecologic analysis that evaluated associations at the census tract and city levels. Secondary data from nine sources as used to conduct the analysis. Data sources include the League of American Bicyclists Benchmarking Report, PeopleForBikes bicycle network analysis, the Centers for Disease Control and Prevention Behavioral Risk Factor Surveillance System, the U.S. Census Bureau American Community Survey, and more. A principal components analysis was conducted to identify relevant infrastructure factors for research question one; ordinary least squares regression models were derived to compare associations between infrastructure and policy factors for research question two, and latent class cluster analysis was conducted to calculate the prevalence odds ratios of the association between bicycle commuting rate and heart disease for research question three.

Three factors accounted for 70% of the variation in bicycle commuting rates. Those three factors include the average number of cyclist fatalities, the number of city employees working on bicycle projects, and bicycle network connections to public transit. The results also show that the association between bicycle commuting rate and heart disease prevalence was only statistically significant in census tract populations with predominantly high socioeconomic status, low health risk factors, and white race. The ecologic study design likely masked any positive health outcomes in populations with low socioeconomic status. The findings of this study provide valuable insights for transportation and public health practitioners, and the conclusions set the stage for future research on cycling and chronic disease outcomes in the United States.

An Aggregate Measure of Bicycle Commuting and its Relationship with Heart Disease Prevalence in the United States

Emma C. Todoroff

GENERAL AUDIENCE ABSTRACT

In the U.S., about 1% of the population bicycle commutes to work. Bicycle commuting is uncommon in the U.S. because many people feel it is not a safe or practical form of transportation. However, several U.S. organizations are working to make roadways safer for cyclists. Some of these organizations include The League of American Bicyclists (LAB), PeopleForBikes (PFB), and Smart Growth America (SGA). The LAB has awarded nearly 500 towns and cities with a bicycle-friendly community award, while PFB has created tools to help urban planners examine bicycle networks in their communities. SGA also helps create bicycle-friendly cities by working with elected officials to advocate for policies that will make roadways safer for cyclists. LAB, PFB, and SGA all collect data as part of their work to learn what U.S. communities are doing to support cycling.

I used LAB, PFB, and SGA data to determine the most important factors for bicycle-friendliness in the 50 largest U.S. cities. I evaluated the impact of 14 factors, including, but not limited to, protected bike lanes, network connectedness, and bicycle-friendly policy. I found that three factors had the strongest association with bicycle commuting rates in large U.S. cities: network connections to public transit, the number of city employees working on bicycle projects, and the number of deaths from cyclist fatalities. Cities looking to increase bicycle commuting should use these results to focus their efforts on improving public transit networks, increasing the number of work hours spent on bicycle projects, and identifying strategies to reduce cyclist fatalities.

U.S. cities should work towards improving bicycle-friendliness because of the population health benefits. In the past 20 years, the percentage of Americans with obesity has increased by 40%, and the percentage of Americans with Type 2 diabetes has doubled. Multiple factors contribute to obesity and diabetes, including bicycle commuting, which has been associated with decreases in both obesity and diabetes. Bicycle commuting may also be associated with heart disease, which is the leading cause of death in the U.S., but more research is needed. In the second part of this study, I evaluated the relationship between bicycle commuting rate and the percentage of Americans living with heart disease.

I used data from the Centers for Disease Control and Prevention (CDC) and the U.S. Census Bureau to understand the association between bicycle commuting and heart disease in 50 U.S. cities. I analyzed the data by census tract to understand health outcomes at a population level. I found that the association between bicycle commuting and heart disease was only significant in census tracts that were predominately high-income. I also found that bicycle commuting rates in high-income populations were three times greater than in low-income populations. In other words, health benefits were not visible in low-income populations because of low bicycle commuting rates. Low-income populations have higher rates of obesity and diabetes and would benefit the most from bicycle commuting, yet these populations confront several barriers to cycling. Cities interested in improving bicycle-friendliness should work to engage low-income populations in their work.

Dedication

To the girls who need a little more wind beneath their sails.

You are strong.

You are smart.

You are capable.

Acknowledgements

Thank you to my advisors, Dr. Freddy Paige and Dr. Tripp Shealy. Thank you for seeing my potential as an undergraduate student and for supporting my education over the past seven years. I am especially grateful that you encouraged me to explore new passions along the way. Thank you to my committee members, Dr. Charlotte Baker and Dr. Courtney Coughenour. The knowledge you were willing to share allowed me to grow as a researcher, and your heartfelt advice kept me grounded throughout this process.

I am also thankful to the Charles E. Via, Jr. Department of Civil and Environmental Engineering, the New Horizons Graduate Scholar Program, and the Myers-Lawson School of Construction for their financial support.

Most importantly, I would like to thank my family, past, present, and future, for supporting me and motivating me throughout this process. In the final months of this experience, I was motivated by inspiring stories of my great-grandmother, Ruby Argersinger, and my great-grandfather, John Chapman, who were incredibly resilient despite the educational obstacles and exploitation they faced in their lives. I am also thankful to my parents, Kent and Bobbi Coleman, for breaking the mold of higher education in our family, and for always encouraging me to continue mine. And I must also thank all of my family who is yet to exist-- I was inspired to leave a legacy for you.

And finally, a very special thanks to my husband, Peter. Thank you for being my rock and for being there for me every step of the way. Your encouragement and support mean the world to me.

Table of Contents

List of Tables

List of Figures

CHAPTER ONE: INTRODUCTION

This dissertation reports on an ecologic study of bicycle commuting and heart disease. The study was based upon secondary data from the 50 most populous cities in the United States. The first chapter of the dissertation presents the motivation of the study, specifies the problems addressed by the study, articulates the research questions, and clarifies the scope of the study.

Motivation

Heart disease is the leading cause of death in the U.S.

Heart disease has been the leading cause of death in the United States for the past 50 years (Murphy et al., 2021). In 2017, heart disease was responsible for approximately one-quarter of all deaths in the U.S., whereas cancer accounted for one-fifth (Heron, 2019). In the most recent report from the CDC National Vital Statistics System, heart disease remained the leading cause at approximately 180 deaths per 100,000 population (Murphy et al., 2021). Heart disease's long-standing rank as the number one killer of the U.S. population presents a need for innovative interventions that significantly reduce heart disease mortality.

Public health organizations have worked to reduce heart disease mortality for decades. Individual and organizational approaches to public health interventions have successfully reduced mortality to a certain degree (Castro et al., 2001; Marcus et al., 2000). More specifically, in 1958, heart disease mortality accounted for 500 deaths per 100,000 population, whereas heart disease mortality accounts for about 180 deaths per 100,000 population today (Murphy et al., 2021). The reduction in deaths has likely resulted from public health interventions combined with improvements to emergency medicine, but existing approaches have not been enough to eliminate heart disease as the leading cause of death. On the other hand, policy, environmental, and systems changes can be powerful because of their ability to affect large portions of the population over the long-term (B. Simons-Morton et al., 2012a). Long-term, lifestyle change is particularly vital for the reduction of heart disease because the disease develops over the life course of several decades.

Physical activity is commonly understood by researchers as a lifestyle risk factor for heart disease because 150 minutes of physical activity per week can reduce an individual's risk of contracting heart disease by up to 35% (Olson et al., 2018). Any activity that burns at least 3.5 kcal per minute qualifies as physical activity, but only 50% of the U.S. population accomplishes the recommended 150 minutes per week (Olson et al., 2018). Inadequate physical activity in the U.S. has been associated with about 117 billion dollars in annual health care costs (Olson et al., 2018). Policy, environment, and system changes must be explored for their potential to increase physical activity, decrease healthcare costs, and ultimately reduce heart disease mortality.

Bicycle commuting is a solution to physical inactivity and heart disease risk

Physical activity can be recreational or utilitarian in nature. The U.S. population commonly thinks of recreational or gym-based physical activity as "exercise." However utilitarian activities, like actively commuting to work, also contribute towards meeting weekly physical activity recommendations (Garrard et al., 2012). Only 15% of those who bike in the U.S. engage in the activity for commuting purposes, whereas 60% bike for recreation (Pucher & Buehler, 2012). In Europe, the trend is reversed with more individuals biking for commuting purposes than for recreation. Studies of European countries show that bicycle commuters are more likely to achieve 150 minutes of weekly physical activity than those who do not bike (Raser et al., 2018). A geographic study across U.S. states also found that increasing percentages of bicycle commuters were associated with meeting physical activity recommendations (Pucher, Buehler, et al., 2010a).

Bicycle commuting's association with heart disease and heart disease risk is less clear than the association with physical activity, especially in the U.S. The majority of studies examining the association between bicycle commuting and heart disease have been conducted in European countries. For example, a meta-analysis across eight European studies found that active commuting was associated with an 11% reduction in heart disease risk (Hamer & Chida, 2008). A more recent meta-analysis also found that active commuting reduced heart disease risk (Dinu et al., 2019). One of the most recent European studies on bicycle commuting and heart disease was conducted in the United Kingdom and found that bicycle commuting was associated with a reduction in heart disease mortality (Celis-Morales et al., 2017). Based on the positive cardiovascular outcomes of bicycle commuting, as evidenced primarily by European studies, there is a need to evaluate the association between bicycle commuting and heart disease risk in the U.S. To my knowledge, geographic studies evaluating the association between bicycle commuting and heart disease have not been conducted in the United States.

There may be several reasons the association between bicycle commuting and heart disease has not been evaluated in the United States. One reason being the siloed surveillance of public health and transportation data. In the U.S., surveillance of public health outcomes is primarily managed by the Centers for Disease Control and Prevention (CDC) while surveillance of transportation mode is primarily managed by the Department of Transportation (USDOT). For example, the CDC collects data on heart disease and disease risk factors using several surveillance instruments, one being the Behavioral Risk Factor Surveillance System (Centers for Disease Control and Prevention, 2014). While the USDOT collects data on transportation habits using instruments like the National Household Travel Survey (USDOT Federal Highway Administration, 2018). In contrast, some European countries, like Finland, have combined their governmental efforts and recognize transportation habits as a vital aspect of health outcomes. For example, FINRISK was a Finnish population survey that collected information on chronic disease risk factors, including transportation habits, every five years from 1972-2012 (Barengo et al., 2004; *The National FINRISK Study*, 2017). Siloed surveillance of transportation habits and health outcomes have resulted in minimal evaluation of the association between bicycle commuting and heart disease in the U.S. However, it is clear that bicycle commuting can help reduce heart disease risk related to physical inactivity, obesity, and diabetes (Huy et al., 2008; Lusk et al., 2010; Ming Wen & Rissel, 2008; Pucher, Buehler, et al., 2010a). The total cost of heart disease is expected to exceed one trillion dollars by 2035, so transportation-based solutions are needed now more than ever (Benjamin et al., 2019). Environmental, policy, and systems changes to the cycling landscape could be an innovative approach to combat heart disease risk in the U.S.

U.S. cycling infrastructure and policy are far behind

Despite the health benefits of cycling, bicycle infrastructure in the U.S. is far behind European countries who have prioritized cycling in their transportation systems. Bicycle facilities or bicycle infrastructure are elements of the transportation system that make roadways safer for cyclists. The primary types of bicycle infrastructure in the U.S. include bicycle lanes and multi-use paths, but can also include roadway markings, roadway signs, and parking hubs. The social, environmental, economic, and health benefits of bicycle infrastructure are multi-faceted because they can help increase cycling rates and improve cyclist safety (Sallis et al., 2015). For example, marked bicycle lanes can reduce vehicle-bicycle collisions by as much as 50% (Pollack et al., 2012). Furthermore, the benefit-cost ratio of healthcare savings to bicycle facility investment has been estimated at 3.8 to 1 (Gotschi, 2011). However, the U.S. has not fully invested in reaping the benefits of bicycle infrastructure.

Bicycle commuting rates in U.S. cities do not compare to European cities like Copenhagen and Amsterdam where rates are 40% and 28%, respectively (Buehler & Pucher, 2012). On average, cycling accounts for 1% of all trips taken in the U.S while bicycle commuting accounts for approximately 0.6% (Buehler et al., 2020; Buehler & Pucher, 2012). Even in some of the most bicycle-friendly, large cities in the U.S., like Portland, OR, bicycle commuting rates are less than 10%. Several explanations have been speculated for minimal cycling rates in the U.S; one of the most compelling being perceived safety risks due to a lack of bicycle facilities, or infrastructure (Coughenour et al., 2016). Providing adequate bicycle infrastructure is key to improve perceptions of safety and subsequently increase bicycle commuting rates (Adam et al., 2020; Rérat, 2019)

Constructing bicycle infrastructure is a complex issue in the U.S. because the majority of federal transportation funding has been dedicated to motor vehicle service. From 2000 to 2018, cumulative federal government spending on bicycle infrastructure amounted to less than 2% of federal roadway expenditures (Buehler et al., 2020). In other words, funding for bicycle infrastructure projects is highly limited in the U.S. and must be strategically allocated to increase cycling rates. To that end, bicycle infrastructure guidelines are not well established in the U.S. which makes it challenging for municipalities to determine how to strategically allocate their limited funds. The Federal Highway Administration (FHWA) released a memorandum in 2013 to support bicycle infrastructure guidelines established by the American Association of State Highway and Transportation Officials (AASHTO) and the National Association of City Transportation Officials (NACTO), yet the FHWA has not appended these guidelines to their official Manual of Uniform Traffic Control Devices (Federal Highway Administration, 2013). Municipalities desiring to increase cycling rates must overcome multiple barriers to improve bicycle infrastructure because of the limited federal funds and the unofficial design guidelines (Handy et al., 2014).

Research Purpose & Objectives

The purpose of this dissertation was based on the three primary motivations of this study: 1) heart disease is the leading cause of death in the United States, 2) bicycle commuting is a solution to physical inactivity and heart disease risk, and 3) the United States cycling landscape is far behind comparable developed countries, especially European countries. The purpose of this dissertation was: *To evaluate the association between U.S. bicycle commuting rates and heart disease prevalence, and to identify infrastructure and policy factors most significantly associated with U.S. bicycle commuting rates. This research quantitatively defined infrastructure and policy factors and analyzed ecologic associations across the 50 most populous U.S. cities*.

The aim of this dissertation was to understand the infrastructure and policy factors associated with increased bicycle commuting in U.S. cities and to evaluate bicycle commuting as a risk factor of heart disease prevalence. Secondary data from the CDC, U.S. Census Bureau, League of American Bicyclists, and PeopleForBikes was used to analyze the relationships between infrastructure, bicycle commuting, and heart disease. The main objectives of the study were to:

1. Identify quantitative measures of bicycle infrastructure and policy,
2. Determine the significance of these quantitative measures for U.S. bicycle commuting rates,

3. Evaluate the ecologic association between bicycle commuting rates and heart disease prevalence in U.S. cities.

Research Questions

The research questions for this dissertation were derived from the problem statement and reflect the study objectives.

Research Question 1: What variables should be used to quantitatively measure bicycle infrastructure when evaluating the association with U.S. bicycle commuting rates?

Research Question 2: How much of the variation in U.S. bicycle commuting rates can be explained by bicycle infrastructure and policy measures?

> ***Sub-Q2:*** What is the relative significance of bicycle infrastructure versus policy measures when evaluating their association with U.S. bicycle commuting rates?

Research Question 3: What is the significance of association between U.S. bicycle commuting rates and heart disease prevalence, at the census tract level, when accounting for heart disease risk factors?

> ***Sub- Q3:*** How much do obesity, diabetes, and hypertension modify the association?

Research Scope

This dissertation presents the results of an ecologic study. Publicly available, federal data on heart disease and bicycle commuting dictated the ecologic design of the study. Individual data was not available for analysis in this study because, as discussed in the *Motivation* section, United States government organizations like the CDC and USDOT silo federal surveillance of public health data from transportation data. In other words, public health records for individuals who bicycle commute are not federally or publicly available in the United States. However, secondary data provided by the CDC and U.S. Census Bureau do include geographic identifiers.

These geographic identifiers allowed me to conduct an ecologic analysis where bicycle commuting rates were matched with heart disease prevalence by census tract. For this reason, the results of my study were only applied to population health outcomes, not individual health outcomes.

Secondary data availability also dictated the study sample and measures of bicycle commuting and heart disease. The CDC provides heart disease data at the census tract level for 500 of the largest U.S. cities, but the League of American Bicyclists only provides bicycle infrastructure data for 50 large U.S. cities. Therefore, the study sample was limited to the 50 most populous U.S. cities. For the purposes of this study, bicycle commuting rate was defined as the number of commuters per 10,000 population who use cycling as their primary mode of transportation to work; heart disease prevalence was defined as the percent of population in a census tract who have been told by a doctor or health professional that they have angina or coronary artery disease.

Dissertation Structure

This dissertation is organized into five chapters. The first chapter described the motivation of the study, specified the problems addressed by the study, articulated the research questions, and clarified the scope. Chapter Two establishes a theoretical basis for the study by describing the socioecological model of behavior change. An overview of the literature on cycling, infrastructure, policy, and heart disease is also provided in Chapter Two with special attention to existing knowledge on cycling measures and associations between cycling and heart disease. Chapter Three describes the secondary data sources used in the study, the variables used to measure infrastructure, policy, cycling, and heart disease, and the statistical methods used to analyze the data. Chapter Four presents the results for each research question and provides a discussion of the findings. Finally, Chapter Five presents conclusions and implications for each research question and provides recommendations for future research.

CHAPTER TWO: LITERATURE REVIEW

The Socioecological Model of Behavior Change

Attempts to increase levels of physical activity have focused on changing behavior at the individual level for decades with public health officials often relying on interventions like health education programs and media campaigns (Walsh et al., 2017). These types of interventions have been successfully implemented over the years, but the outcomes are normally short-lived (Castro et al., 2001; Marcus et al., 2000). For example, a health education program called Heart Smart for Women was implemented in Illinois to increase physical activity and reduce CAD risk amongst women living in rural counties. The twelve-week educational program resulted in significant behavior change from baseline to post-intervention, but the changes were no longer significant when post-intervention outcomes were assessed 12 months later. Insignificant behavior change at follow-up suggests that Heart Smart for Women did not provide participants with the skills needed to maintain heightened levels of physical activity post-intervention (Khare et al., 2014). Other types of health interventions like health policy and environmental change may have greater potential for long-term behavior change.

The socioecological model offers a solution for long-term behavior change by placing individual behavior within the larger context of policy, environment, and culture (Bronfenbrenner, 1992). The socioecological model suggests that changes to the social, ecological, and political environment can change health behaviors, at a population level, for the long-term (B. Simons-Morton et al., 2012a). Early views of health behavior theory asserted that personal choices were under the primary control of the individual (B. Simons-Morton et al., 2012a). But Bronfenbrenner, a developmental psychologist, theorized that human behavior occurs within the context of multiple systems ranging from microsystems of the individual to macrosystems of community and culture (Bronfenbrenner, 1981). Bronfenbrenner's theory resulted in the socioecological model of behavior change which addresses five societal levels: individual, interpersonal, organizational, community, and government or public policy (McLeroy et al., 1988; D. G. Simons-Morton et al., 1988). Since origination of the model, levels representing the physical environment and culture have also been added (B. Simons-Morton et al., 2012a). A visual representation of the socioecological model is provided in Figure 2.1.

Figure 2.1: Socioecological model of health behavior

(Simons-Morton et al., 2012a)

The societal levels of the socioecological model are embedded in a nesting framework where individuals' knowledge, attitudes, and behaviors are constantly affected and being affected by the larger system in a phenomenon called reciprocal determinism (B. Simons-Morton et al., 2012a). In other words, all levels interact as part of a larger system, but each level also uniquely contributes to behavior change. The intrapersonal level, or individual level, posits that an individual's knowledge, values, and beliefs effect their behavior. It recognizes the relationship between cognition and behavior, but also acknowledges cognitive dissonance as a common barrier to behavior change. Cognitive dissonance refers to inconsistencies between what an individual knows or believes they should do, and the behaviors they actually exhibit (B. Simons-Morton et al., 2012b). Value-expectancy theories are also commonly applied to individual behavior which suggest that behavior change is more likely when the perceived advantages of a particular action outweigh the costs (B. Simons-Morton et al., 2012b). Health education programs are a common intervention for influencing behavior change at the individual level (Khare et al., 2014; Walsh et al., 2017).

The influence of relationships on behavior change is reflected in the interpersonal level of the socioecological model. Family, friends, acquaintances, neighbors, co-workers, and doctors are just a few examples of relationships that may influence an individual's behavior. Social network theories claim that the strength and structure of relationships are meaningful for behavior change (B. Simons-Morton et al., 2012c). The more direct a relationship, the more influential it will be to the individual (B. Simons-Morton et al., 2012c). For example, an individual may be more influenced by a sibling or parent whom they interact with every day than a doctor they may interact with once or twice a year. Social norms also play a role through relational influence which occurs indirectly through an individual's perception of what is allowed or expected (B. Simons-Morton et al., 2012c). Interventions at the interpersonal level often facilitate social support through the creation of community programs (Khare et al., 2014; Walsh et al., 2017). Organizational partnerships are often sought after to implement community programs. Community organizations like schools, churches, and workplaces are key partners for health behavior change. Interventions at the organizational level often include fostering the adoption of healthful programs and practices. (B. Simons-Morton et al., 2012a)

Individuals, interpersonal relationships, and organizations all exist within a larger community. Communities contain multiple organizations and are considered the fourth level of the socioecological model. In the past, community-based health programs have viewed community as the setting where health promotion occurs (McLeroy et al., 2003), but community efforts to address cardiovascular disease have achieved limited success under this narrow-minded perspective (Shea & Basch, 1990). Community based participatory research (CBPR) has more recently emerged as an approach to health promotion at the community level (*Community Based Participatory Research for Health*, 2003). Under this approach, communities are viewed as equal partners with universities and funding agencies in the development of interventions for healthy behavior (B. Simons-Morton et al., 2012a). Initiatives of CBPR are often effective because they account for social conditions within individual communities like unemployment, poverty, and education (B. Simons-Morton et al., 2012a).

The outermost levels of the socioecological model are public policy, physical environment, and culture. All three levels offer the potential to change the environmental context within which health behavior occurs (B. Simons-Morton et al., 2012a). However, changes to policy and the physical environment do take time because they require community support and consensus amongst policymakers. Cultural change is even more challenging because culture itself is engrained in the identity of a community as a shared system of learned norms, beliefs, values, and behaviors that may differ by region, nationality, ethnicity, or religion (Hruschka & Hadley, 2008). The challenge with implementing policies or intervening in the physical environment is that the bureaucratic systems governing these levels of society must recognize and embrace a need before change can occur (B. Simons-Morton et al., 2012a).

The physical environment, or built environment, and public policy can enable health behavior change with appropriate attention from policymakers, health professionals, and urban planners. For example, in 2006 California was the only state with more than five local/regional Complete Streets policies and today all but 10 states have at least five local/regional Complete Streets policies (*Policy Inventory*, 2019) which aim to increase walking and biking behavior. Cultural change is also possible, but it normally requires long periods of time. For example, sixty years ago cigarette smoking was viewed as a positive, sexually appealing behavior within U.S. culture. Over several decades, strategic health-based interventions at the organizational, community, and policy levels were implemented to change smoking behavior. Now, the majority of the U.S. population views cigarette smoking as a deviant behavior and one that is practiced only by individuals who lack the willpower to quit (B. Simons-Morton et al., 2012a). Cultural change required decades, but it eventually emerged through the power of organizational, community, built environment, and policy interventions.

My dissertation focuses on built environment and public policy as they relate to bicycle commuting behavior. The remainder of my literature review will be divided into two sections. First, I will provide a landscape of the policy and built environment factors that have been studied in prior research on bicycle commuting. Second, I will introduce the literature on heart disease and its associated risk factors, and I will provide an overview of existing research on the associations between bicycle commuting and heart disease. These literature sections will lead to my research questions which seek to 1) identify the policy and built environment factors that impact bicycle commuting in U.S. cities and 2) evaluate the association between bicycle commuting and heart disease in U.S. census tracts.

Cycling, Policy, and the Built Environment

The built environment encompasses buildings, planned outdoor spaces, transportation infrastructure, and all other human-made components of the physical environment (Saelens & Handy, 2008). The Active Living Research (ALR) group recognized the influence of the built environment on physical activity nearly two decades ago, and as a national program of the Robert Wood Johnson Foundation, their mission was to "build evidence about how to create communities that are great for physical activity"(*About Us*, 2019). In a recent effort to summarize the large body of evidence on built environment and physical activity, ALR released a review on the co-benefits of activity-friendly environments (Sallis et al., 2015). In their review, ALR categorized the built environment into five groups: urban design/land use, workplaces/buildings, schools, open spaces/parks/trails, and transportation systems. ALR found that physical activity and its associated health outcomes have been extensively explored in all categories of the built environment except for transportation systems. Understanding the impact of transportation systems on physical activity was identified as a significant gap in knowledge (Sallis et al., 2015).

Physical activity facilitated by transportation systems is frequently called "active transportation." Active transportation is defined as any self-propelled, human-powered mode of transport and includes both walking and biking (Saelens & Handy, 2008). The built environment's influence on walking has been widely applied in practice (Saelens & Handy, 2008; Stewart et al., 2016). For example, urban planners commonly apply the walkability index when designing cities conducive to walking. The walkability index includes four elements of the built environment: residential density, intersection density, land use mix, and retail floor to area ratio (Frank et al., 2009). Other built environment features that have been correlated with walkability are proximity to parks, access to transit, and density or proximity to points of interest (Clifton & Dill, 2005; Gauvin et al., 2005; Handy, S & Transportation Research Board, 2005; Hoehner et al., 2005; Lee & Moudon, 2006). Contrastingly, a "bike-ability" index has not been well-established (Buehler & Dill, 2016).

Infrastructure Influences on Cycling

Built environment influences on cycling levels have been examined by several researchers (Buehler & Dill, 2016; Pucher et al., 2010). Bike lanes and bike paths have been most widely investigated in the literature. Bike lanes are most commonly indicated by a striped line on U.S. roads while bike paths are separate from roads and often allow pedestrian traffic. Most studies have found a positive association between bike lanes or bike paths and cycling levels, but it is unclear whether one has a more significant impact (Buehler & Dill, 2016). A study that predicted aggregate cycling levels for 90 large U.S. cities concluded that both bike paths and bike lanes have statistically similar positive associations with cycling levels (Buehler & Pucher, 2012). Other studies have revealed a preference for bike lanes or bike paths relative to demographic characteristics. For example, women and inexperienced cyclists tend to prefer bike paths while experienced bike commuters prefer bike lanes (Garrard et al., 2008; Stinson & Bhat, 2003; Tilahun et al., 2007).

The interplay between bike paths and bike lanes could be better understood, but regardless, researchers must consider the availability of bike paths and bike lanes when evaluating cycling levels (Buehler & Dill, 2016). Studies have shown that both cyclists and non-cyclists feel safer on roadways where bicycles travel separate from motor vehicles (Akar & Clifton, 2009; Fishman et al., 2015; Sanders & Cooper, 2013). For example, cycle tracks separate bicycle traffic from motor traffic with a curb or sidewalk and act like multi-use paths along major roadways. Cycle tracks are heavily implemented in Denmark but are rare within the United States (Buehler & Dill, 2016; Pucher et al., 2010). Experimental cycle tracks in Copenhagen and Washington, D.C. have significantly increased cycling levels (Goodno et al., 2013; Snizek et al., 2013).

Buffered bike lanes also separate bicycle traffic from motor traffic, and they are more common in the U.S. than cycle tracks. Studies have shown that buffered lanes increase perceptions of safety for both cyclists and non-cyclists (Coughenour et al., 2016; Waygood et al., 2019), an important factor considering that perceived safety often dictates cyclists' route choices. For example, a study in Portland, OR found that roadways separating bicycle traffic from motor traffic accounted for 50% of cyclists' chosen routes, even though only 8% of the roadway network provided separate space for the two transportation modes (Broach et al., 2012; Dill, 2009). In other words, roads with infrastructure separating bicycle traffic from motor traffic were disproportionately chosen by cyclists.

Bikeshare stations and bike parking have also been correlated with cycling levels (Buehler & Dill, 2016; Pucher et al., 2010). There is strong evidence that the *proximity* of bike parking to metro and transit stops is correlated with cycling levels (Brunsing, 1997; Hegger, 2007; Martens, 2007; Pucher & Buehler, 2009; Rietveld, 2000). The *availability* of bikeshare and bike parking has been cited in fewer studies (Hunt & Abraham, 2007; Wardman et al., 2007). In addition to infrastructure elements like bike lanes, bike paths, and bike parking, land use attributes like street density, street connectivity, bike network density, and bike network directness have been correlated with cycling levels (Beenackers et al., 2012; Cervero et al., 2009; Dill et al., 2014; Dill & Voros, 2007; Schoner & Levinson, 2014).

Bike networks are challenging to define in the U.S. because 50 to 90% of the roadways used by cyclists do not contain infrastructure separating bike traffic from motor traffic (Buehler & Dill, 2016). In fact, the majority of U.S. cyclists ride on roadways without separate facilities for at least a portion of their route. Prior researchers have used measures of density, directness, fragmentation, and connectivity to evaluate bike networks (Schoner & Levinson, 2014). Bike network density has been most significantly associated with cycling levels when density accounted for roadways that contain infrastructure separating bike traffic from motor traffic (i.e. bike lanes and bike paths). A change in bike network density by one standard deviation corresponded to a 77% increase in bicycle commuting rates (Schoner & Levinson, 2014).

Policy Influences on Cycling

Public policy in the U.S. has heavily influenced a car-centric transportation system, and bike networks are arguably less developed in the U.S. than in countries where cycling rates are high. In 1991, the U.S. Congress passed the Intermodal Surface Transportation Efficiency Act (ISTEA), to encourage alternatives to motor vehicles (Dill et al., 2017). The Act required states to include biking and walking in their transportation plans, opened new funding sources for bike and pedestrian facilities, and required each state DOT to have a bike and pedestrian coordinator (Pucher et al., 1999). The ISTEA successfully increased federal spending on bicycle and pedestrian projects (Cradock et al., 2009). However, in 1995, a follow-up study found that bike and pedestrian transportation plans were being created by localities but were not being effectively implemented (Moe et al. 1997).

Decades have passed since ISTEA was passed, yet the U.S. has not updated legislative policy regarding active transportation since 1991. In 2010, the USDOT attempted to update national policy with a policy statement on Bicycle and Pedestrian Accommodation Regulations.

The policy statement asserts, “every transportation agency, including the DOT, has the responsibility to improve conditions and opportunities for walking and biking” (USDOT, 2010). However, the statement does not articulate requirements for taking action. The USDOT statement includes some recommended actions, but none are required. As the two existing national policies that address active transportation, the ISTEA and the USDOT policy statement were the basis for defining policy variables in my dissertation.

Municipalities have identified funding as one of the largest barriers to implementation of cycling projects (Assunçao-Denis & Tomalty, 2019; Dill et al., 2017). From 2000 to 2018, cumulative U.S. federal government spending on bicycle projects amounted to less than 2% of federal roadway expenditures (Buehler et al., 2020). In 1991, the ISTEA encouraged regional spending on bicycle and pedestrian projects (S. L. Handy et al., 2009), yet, 30 years later, the percent of transportation budget allocated to cycling projects remains minimal (Arellana et al., 2020; Buehler et al., 2020).

In addition to funding, the implementation of cycling projects relies on successful partnerships between government agencies and community groups (Marsden & Stead, 2011). Collaboration between city transportation staff and cycling advocacy organizations has facilitated the creation of bike and pedestrian plans as well as the adoption of Complete Streets policies (Aytur et al., 2013; Moreland-Russell et al., 2013). More specifically, the efforts of DOT bike and pedestrian coordinators combined with the political pressure created by cycling interest groups have been two of the strongest motivating factors for states and localities to adopt stricter bicycle policies (Dill et al., 2017).

Cycling interest groups have played a key role in U.S. cities with high bicycle commuting rates like Portland, Minneapolis, Chicago, and San Francisco (Pucher et al., 2011). These advocacy organizations disseminate information about the benefits of cycling to generate public and political support, and they lobby for funding (Pucher et al., 2011). The strength of these organizations can be measured by the number of full-time equivalent city staff working on cycling projects. For example, San Francisco has the strongest bike advocacy in the U.S. with the most funding, and they have six times the number of full-time staff as New York City (Pucher et al. 2011). The quality and sustainability of cycling initiatives is limited when cities do not prioritize staff for cycling projects (Zieff et al., 2013).

City staff are instrumental in passing local transportation policy like Complete Streets policy. Many states, MPOs, and localities have passed Complete Streets policies to institutionalize the USDOT policy statement (Biton et al., 2014). A study across 48 U.S. cities found that the existence of Compete Streets policy has been associated with an increase in cycling levels (Suminski et al., 2014). Complete Streets policies assert that streets should be designed for all potential users, not just motor vehicles, but their scope varies by locality (Moreland-Russell et al., 2013). Smart Growth America, a non-profit organization working with policymakers to improve public policy, derived a scale to rate the quality of Complete Streets policies (*Policy Inventory*, 2019). The scale contains ten ranking categories that quantitatively compare policies between states, MPOs, and localities. The categories of the ranking scale include: vision and intent, diverse users, commitment, design, land use and context sensitivity, project selection criteria, jurisdiction, exceptions, performance measures, and implementation steps (Riveron, 2019). One of the main benefits of the Complete Streets policy rating is that it weighs implementation steps most heavily. Rating implementation steps most heavily may hold localities accountable for taking actionable steps towards built environment change. Case studies in the literature have referenced the importance of policy in promoting cycling since 2010, but

Complete Streets policy has not been quantitatively evaluated in a model of cycling levels (Buehler & Dill, 2016; Pucher et al., 2010).

To summarize, several single-case, disaggregated studies have provided evidence of an association between the built environment, public policy, and cycling levels. Single case studies are valuable because they often provide time-series evidence on cycling levels pre and post built environment or policy interventions. However, single case studies rarely evaluate the interplay between built environment and policy factors (Buehler & Pucher, 2012). There are far fewer studies that evaluate cycling levels with an aggregate measure. The following section summarizes four cross case studies that have evaluated the impact of built environment on U.S. cycling levels with an aggregate measure. The comparison between the four studies highlights the built environment and policy measures evaluated by prior researchers.

Measuring Socioecological Influences on Cycling

Over the past two decades, a series of regression models has evaluated aggregate bicycle commuting rates across U.S. cities. These models have included variables measuring bicycle infrastructure, bicycle networks, and covariates like urban sprawl, vehicle ownership, amount of rainfall, and college student population. A complete account of variables considered in prior models, and their statistical significance are listed in Table 2.1. The table codes statistically significant variables with a "✓", non-significant variables with a "×", and variables not included in prior studies with a "-." It should be noted that several infrastructure variables (e.g., miles of bicycle lanes, miles of bicycle paths, bicycle network density, and bicycle network directness) were statistically significant, while evaluation of policy-related variables has been neglected.

Table 2.1: Socioecological variables that have been evaluated in bicycle commuting regressions[1]

Independent Variable	1997	2003	2011	2014
Miles of bicycle lanes	-	✓	✓	-
Miles of bicycle paths	✓	×	✓	-
Bicycle network size factor	-	-	-	×
Bicycle network density factor	-	-	-	✓
Bicycle network connectivity factor	-	-	-	×
Bicycle network fragmentation factor	-	-	-	×
Bicycle network directness factor	-	-	-	✓
Percent of college students in the population	✓	×	✓	✓
Percent of workers by industry category	-	×	-	-
Number of workers	-	×	-	×
Average number of vehicles per household	-	✓	-	×
Percent of households with zero vehicles	-	×	✓	-
Percent of households with kids	-	×	-	×
Gasoline price	-	×	×	-
Transit vehicle revenue miles per mi. service area	-	×	-	-
Annual vehicle miles of transit supply	-	-	×	-
Percent of housing units built before 1950	-	×	-	-
Sprawl index	-	-	✓	-
Population density	-	×	-	×
Average household income	-	×	-	-
Median household income	-	×	-	×
Percent of people older than 18 in poverty	-	×	-	-
Average number of days of rainfall per year	✓	✓	-	-
Average annual inches of rainfall	-	×	×	-
Number of hot and cold weather days	-	-	×	-
Mean high temperature	×	-	-	-
Average bicyclist fatalities per 10,000 commuters	-	-	✓	-
Average annual per capita state spending on bike or pedestrian improvements	-	✓	-	-

[1] The regressions summarized in this table are from Buehler and Pucher 2012, Dill and Carr 2003, Nelson and Allen 1997, and Schoner and Levinson 2014.

Nelson and Allen pioneered bicycle commuting regression models in 1997 by evaluating four variables across 18 U.S. cities. They found bike path supply, percent of college student population, and average number of annual rainfall days were significantly associated with bicycle commuting rates (Nelson & Allen, 1997). Dill and Carr (2003) contributed to bicycle commute regressions in 2003 when they evaluated bike lane supply in their 35-city model. They considered a total of 18 variables, those that Nelson and Allen (1997) found significant along with an additional 15 variables. However, findings from Dill and Carr's (2003) model were inconsistent with Nelson and Allen (1997). The average number of annual rainfall days was the only variable they found was similarly significant (Dill & Carr, 2003). In contrast to Nelson and Allen's (1997) model, bike path supply and percent of college student population were not statistically significant. Three of the variables evaluated by Dill and Carr (2003) were statistically significant: bike lane supply, average number of vehicles per household, and average annual per capita state spending on bike/pedestrian improvements. Dill and Carr's (2003) model of bicycle commuting rates evaluated several new variables, but the R^2 value (0.304) was significantly lower than Nelson and Allen's (1997) model (0.825). The lower R^2 value could be related to doubling the number of cities analyzed.

In 2011, Buehler and Pucher tripled the number of cities analyzed to derive a regression model across 90 cities (Buehler & Pucher, 2012). However, they did not evaluate three of the variables that were statistically significant in Nelson and Allen (1997) and Dill and Carr's (2003) previous models (see Table 2.1). The average number of annual rainfall days was not included, even though it was a significant variable in both prior models. Buehler and Pucher (2011) considered ten variables and found that six had statistically significant associations with the number of bike commuters per 10,000 population. Similar to Nelson and Allen's (1997) model, but dissimilar to Dill and Carr's (2003) model, bike path supply and percent of college student population were significant. Similar to Dill and Carr's (2003) model, bike lane supply and vehicle availability were significant. However, Buehler and Pucher (2011) measured vehicle availability by the percent of households with zero vehicles instead of the average number of vehicles per household. Sprawl index and average bicyclist fatalities per 1,000 commuters were also significantly associated with bicycle commuting in Buehler and Pucher's model and had not been evaluated in prior models. The R^2 value of their model was higher than Dill and Carr (2003) at a value of 0.65 which is likely attributable to a greater number of statistically significant variables.

The most recent regression model of U.S. bicycle commuting rates, published in 2014, was significantly different from the previous three models because bicycle network measures were the focus. Five measures of the bicycle network were calculated from 74 city's bicycle path and bicycle lane data (Schoner & Levinson, 2014). Network density and network directness were the only two network measures significantly associated with the number of bike commuters. In total, Schoner and Levinson (2014) included eleven variables in their model, but the two network measures along with the percent of college student population were the only three statistically significant variables. Perhaps due to the rigor involved with calculating bicycle network measures, Schoner and Levinson (2014) failed to include five of the statistically significant variables from the prior models (see Table 2.1). The knowledge Schoner and Levinson contributed for network measures is valuable. However, future models should include network directness and density in addition to those variables that were statistically significant in prior models. Regardless of the limitations, the R^2 value of Schoner and Levinson's (2014) model was impressive ($R^2 = 0.804$), considering they found only three statistically significant variables across a large number of cities. The strength of their model alludes to the power of network density and network directness for predicting the number of bicycle commuters.

Cycling and Heart Disease

The built environment has a significant impact on physical activity. A study in Montreal Canada found that street connectivity and residential density were better predictors of physical activity than participation in organized sports or fitness (Ross & Hermann, 2019). In the

Canadian study, populations living in areas with high street connectivity and residential density participated in active transportation which allowed them to meet or exceed weekly recommendations for 150 minutes of physical activity (Ross & Hermann, 2019). This study is one of many where the built environment was positively associated with physical activity (Ding & Gebel, 2012; Durand et al., 2011; Hankey et al., 2012; Salvo et al., 2018).

Bicycle commuting is a form of habitual physical activity that can help prevent chronic conditions like obesity, hypertension, and diabetes (Berger et al., 2018; Hu et al., 2002; Huy et al., 2008; Laverty et al., 2013; Pucher, Buehler, et al., 2010; von Huth Smith et al., 2007). In the United States, published studies predominately evaluate bicycle commuting as a means of obesity prevention. Lusk et al. (2010) found that bicycling reduced weight gain over time, especially among those who were overweight or obese. Wojan and Hamrick (2015) found that bicycle commuting was associated with lower body mass index (BMI); and Suminski et al. (2014) found that U.S. cities with higher rates of bicycle commuting had lower rates of obesity. Health outcomes associated with bicycle commuting, such as obesity reduction, likely translate to the prevention of atherosclerotic heart disease, also known as coronary artery disease (CAD).

Heart Disease and its Risk Factors

There are multiple types of cardiovascular disease, or heart disease, including congenital heart disease, syphilitic cardiovascular disease, pulmonary heart disease, rheumatic heart disease, bacterial endocarditis, and coronary artery disease (Duchosal et al., 1958). The World Health Organization (WHO) states that coronary artery disease (CAD) is responsible for the majority of cardiovascular deaths (Duchosal et al., 1958). WHO released a statement on the urgency of CAD research when cardiovascular deaths were at their peak in 1960, and heart disease was responsible for more than 350 deaths per 100,000 amongst the U.S. population (Dalen et al., 2014). Death rates from heart disease have declined since 1960 (200/100K) based on a reduction in deaths from acute myocardial infarction (heart attack) and sudden death (Dalen et al., 2014). However, heart disease remains responsible for the greatest percentage of deaths in the United States and CAD is responsible for the majority of those cases. As the long-standing number one killer in the U.S., coronary artery disease requires continued attention in epidemiologic research.

The principle manifestations of CAD include: 1) angina pectoris, 2) myocardial infarction, 3) sudden death, and 4) congestive heart failure (Duchosal et al., 1958). Manifestations of CAD are notably different between the sexes because CAD in females manifests primarily as angina pectoris (70%) and primarily as myocardial infarction or sudden death in males (Kannel et al., 1961). These results were based on the Framingham Study, one of the most trusted cohort studies ever conducted on cardiovascular disease in the U.S. (Friedman, 1994; Jekel et al., 1996). During the Framingham Study, researchers conducted biennial comprehensive examinations for 20 years on over 5000 subjects. Researchers collected data from physical examinations, laboratory tests, and medical histories. Results of the Framingham Study have been presented in a series of reports since 1951 with the major findings being a set of risk factors for CAD. The risk factors include: male sex, advancing age, high blood pressure, diabetes or glucose intolerance, obesity, cigarette smoking, high serum lipid concentrations, enhanced blood clotting factors, and electrocardiographic abnormalities (Friedman, 1994). Furthermore, results of the Framingham study suggest that risk factors interact synergistically creating an inflation effect. For example, an individual with risk factors of smoking, glucose intolerance, hypertension, high cholesterol, and hypertrophy of the left ventricle would have a 78% risk of developing CAD in 8 years while an individual with none of these risk factors would have a 2.2% risk. The likelihood of disease development would have only been 30% if the risk factor percentages were simply added, yet the observed risk of 78% was more than 2.5 times the expected risk, demonstrating an inflation effect on CAD risk (Jekel et al., 1996).

Other studies have examined the genetic, psychosocial, and eco-social factors associated with CAD, although the mechanisms that cause these associations are unknown and often intertwined (Krieger, 2001; Lovallo & Gerin, 2003; Schwartz et al., 2003). Genetic predisposition can have varied effects on CAD risk based on the environment in which the genes are expressed (Schwartz et al., 2003). For example, a study was conducted on a group of nuns living in secluded quarters (Timio et al., 1997). At baseline, the age, blood pressure, body mass index, and family history of hypertension did not differ between the nuns and a group of women living in the same area. Over a 30 year follow up period, none of the nuns developed diastolic blood pressure greater than 90 mm Hg even though 19% of the nuns did have a family history of hypertension. Contrastingly, blood pressure increased with age amongst the control group. Blood pressure differences between the nuns and control group could not be explained by physical activity, diet, childbearing, or BMI (Timio et al., 1997). Results of the study demonstrate the power of psychosocial factors, like chronic stress, on CAD risk. Job strain, marital stress, and low socioeconomic status are sources of chronic stress that have been associated with CAD risk (Kaplan & Keil, 1993; Karasek et al., 1988; Orth-Gomér et al., 2000).

Eco-social factors like social class also have complex interactions with CAD risk. A cross-sectional study conducted in Evans County, Georgia found that social class, measured in terms of occupation, education, and income, was associated with CAD prevalence (Cassel, 1971; Hames, 1971). The age-adjusted CAD prevalence rates per 1000 population were 97, 40, and 21 for high social class whites, low social class whites, and blacks, respectively (McDonough et al., 1965). CAD prevalence differences between social classes could not be explained by blood pressure, cholesterol levels, body weight, smoking, or diet. However, the researchers estimated habitual physical activity by occupation type and found an inverse association with CAD prevalence (McDonough et al., 1965). In other words, Evans County populations with jobs requiring high levels of physical activity, like manual labor, had a lower risk of CAD prevalence. Studies in the UK have found opposing results where those working highly physical jobs were more likely to die of CAD than those working decision-heavy, administrative positions (Marmot & Theorell, 1988). Occupational effects on CAD prevalence and death rates could be related to influences on blood pressure, smoking, or obesity or it could be related by more direct stress mechanisms (Marmot & Theorell, 1988). The associations between social class and CAD prevalence are mixed, but it is clear that a significant influence exists in some capacity. Measures of social class can be challenging to define, but ethnicity, occupation, education, and socioeconomic status are typically used to measure socioeconomic factors in epidemiological studies (Kaplan & Keil, 1993; MacMahon & Trichopoulos, 1996).

Interactions between genetic, socioeconomic, and eco-social factors increase the risk of CAD incidence. These interactions are showcased in Figure 2.2 in a web of causation for myocardial infarction, more commonly known as heart attack, a common manifestation of CAD incidence (Friedman, 1994). At the top of the web are genetic adaptation, social pressure, and industrial society. Social pressure and industrial society are analogous to the physical environment and cultural levels of the socioecological model. More specifically, industrial society might reflect a city's infrastructure or food environment. Most factors on the second level of the web would be classified as lifestyle risk factors (i.e. lack of exercise, dietary excess, cigarette smoking, and chronic stress). In the third level of the web, genetic factors, in combination with lifestyle risk factors, have subsequent effects on obesity, hypertension, diabetes, and other medical conditions. With the exception of coronary artery distribution, the medical conditions leading to CAD incidence are directly linked to lifestyle factors of diet, stress, smoking, and physical activity.

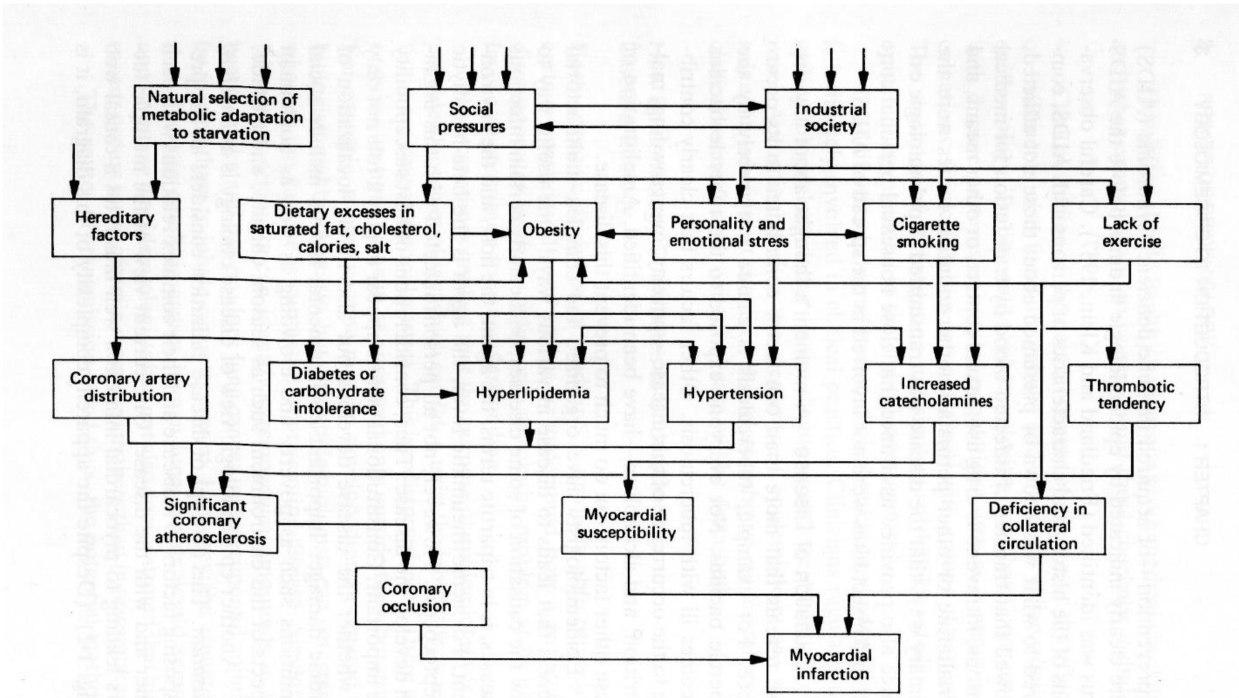


Figure 2.2: Web of causation for myocardial infarction

(Friedman, 1994)

Associations Between Cycling and Heart Disease

Physical activity, and more specifically active transportation, has been associated with CAD risk factors like obesity, diabetes, and hypertension (Bassett et al., 2008; Dons et al., 2018; Hu et al., 2002, 2003; Huy et al., 2008; Pucher, Buehler, et al., 2010b). The majority of studies evaluating the associations between active transportation and CAD have been conducted in European countries (Barengo et al., 2004; Celis-Morales et al., 2017; Hu et al., 2007; Wennberg et al., 2006). In countries like Sweden, England, and Finland, bicycle commuting has been associated with a lower risk of CAD incidence and reduced CAD mortality (Barengo et al., 2004;

Celis-Morales et al., 2017; Wennberg et al., 2006). These European associations could translate to similar associations in the United States (Hamer & Chida, 2008). However, CAD risk factors do vary by country. A study conducted in England found that manual occupation was a risk factor of CAD (Marmot & Theorell, 1988) while studies conducted in Sweden, Finland, and the U.S. found the opposite (Barengo et al., 2004; McDonough et al., 1965; Wennberg et al., 2006).

Fewer studies have evaluated the association between bicycle commuting and CAD in the United States. One U.S. study found statistically significant associations between active transportation, obesity, and diabetes at the state and city level (Pucher et al., 2010). Another U.S. study found statistically significant associations between bicycle commuting and obesity amongst a population of women, but the study did not consider diabetes or hypertension (Lusk et al., 2010). Table 2.2 summarizes existing literature on the association between active transportation, bicycle commuting, and CAD risk. The table lists six studies that have evaluated associations between active transportation, CAD, and CAD risk factors; and six studies that have evaluated associations between bicycle commuting, CAD, and CAD risk factors.

Table 2.2: Research studies evaluating the association between active transportation, bicycle commuting, and CAD risk factors

CAD/ CAD risk factors	**Active transportation**	**Bicycle commuting**
Obesity	U.S.- Pucher et al. 2010 [1]Bassett et al. 2008 China- Hu et al. 2002	[2]Dons et al. 2018 U.S.- Lusk et al. 2010 Australia- Ming Wen & Rissel 2008 Germany- Huy et al. 2008
Diabetes	U.S.- Pucher et al. 2010 Finland- Hu et al. 2003	Germany- Huy et al. 2008
Hypertension	China- Hu et al. 2002	Germany- Huy et al. 2008
Heart disease incidence	Finland- Hu et al. 2007	
Heart disease mortality	Finland- Barengo et al. 2004	U.K.- Celis-Morales et al. 2017
Myocardial infarction		Sweden- Wennberg et al. 2004

[1]Associations across 12 countries: U.S., Australia, Canada, Ireland, France, Denmark, Finland, Germany, Sweden, Spain, Netherlands, Switzerland
[2]Associations across 7 European cities: Antwerp, Barcelona, London, Oerebro, Rome, Vienna, Zurich

Variables that are risk factors of CAD and also associated with bicycle commuting would be considered confounders in an epidemiologic analysis. There are three criteria for confounders: 1) they are associated with the independent variable; 2) they are associated with the dependent variable; and 3) they are not an intermediate step in the causal pathway (Friis & Sellers, 2015). I conducted a review of potential confounders of the association between bicycle commuting and CAD prevalence. The results of the review are cited in Table 2.3. The first column lists risk factors of CAD; the second column indicates whether the risk factor has been associated with bicycle commuting and the direction of association. The "+" means the risk factor was positively associated with bicycle commuting, the "–" means the risk factor was inversely associated with bicycle commuting, and "0" means there was no statistically significant association between the risk factor and bicycle commuting. A symbol of association is provided for each study that is cited in the third column of the table. Probable confounders include male sex, age, smoking, race, household income, and college education.

Table 2.3: Potential confounders of the association between bicycle commuting rate and CAD prevalence[1]

CAD risk factors	Association with bicycle commuting[2]	Citations
Male sex (%)	++ + + + + + + + + +	Buehler et al., 2020; Cole-Hunter et al., 2015; Donaire-Gonzalez et al., 2015; Goldsmith, 1992; Parkin et al., 2008; Reis et al., 2013; Sallis et al., 2013; Stinson & Bhat, 2003; Wardman et al., 2007; Williams & Larson, 1996)
Age (mean)	– – – – – – + 0 0	Buehler et al., 2020; Cole-Hunter et al., 2015; Donaire-Gonzalez et al., 2015; Goldsmith, 1992; Plaut, 2005; Reis et al., 2013; Sallis et al., 2013; Stinson & Bhat, 2003; Tribby & Tharp, 2019; Wardman et al., 2007)
Smoking (%)	– –	Donaire-Gonzalez et al., 2015; Kaczynski, 2008
Chronic Stress proxies:		
Manual labor occupation (%)	0	Dill & Carr, 2003
Married (%)	– 0 0	Donaire-Gonzalez et al., 2015; Reis et al., 2013; Sallis et al., 2013
Socioeconomic Status proxies:		
White (%)	+ + +	Parkin et al., 2008; Plaut, 2005; Sallis et al., 2013
Income (mean)	– – – – – + 0 0	Cole-Hunter et al., 2015; Goldsmith, 1992; Parkin et al., 2008; Plaut, 2005; Shafizadeh & Niemeier, 1997; Stinson & Bhat, 2003; Tribby & Tharp, 2019; Wardman et al., 2007)
No college education (%)	– – – – – –	Buehler et al., 2020; Cole-Hunter et al., 2015; Donaire-Gonzalez et al., 2015; Plaut, 2005; Reis et al., 2013; Sallis et al., 2013)

[1] The format of this table was adapted from (S. L. Handy & Xing, 2011)
[2] The symbols indicate direction of association, + is positive association, – is inverse association, and 0 is no significant association. A symbol is provided for each study that is cited in the third column.

CHAPTER THREE: METHODOLOGY

Data Sources

Table 3.1 provides a summary of the data sources that I used to answer the research questions in this dissertation. The table lists the data sources by research question with some sources being used for more than one question. All data sources are secondary and were collected by a variety of organizations and governmental agencies. I curated the data sets and combined them to answer my research questions. I provide details about each data source in the following subsections. The details I provide include how each organization or agency collected the data, the frequency with which the data was collected, and how I accessed the data.

Table 3.1: Summary of data sources by research question

Data Source	**RQ1**	**RQ2**	**RQ3**
League of American Bicyclists Benchmarking Report (K. McLeod et al., 2019)	✓	✓	
PeopleForBikes Bicycle Network Analysis (Bicycle Network Analysis [Data], 2020)	✓	✓	
Smart Growth America’s Best Complete Streets Initiatives (Atherton et al., 2018)		✓	
League of American Bicyclists Lifting the Veil on Bicycle and Pedestrian Spending (K. McLeod, 2017)		✓	
USDOT Fatality Analysis Reporting System (*Fatality Analysis Reporting System [CSV Data File]*, 2017)		✓	
Sprawl Indices (Ewing & Hamidi, 2013)		✓	
NOAA U.S. Climate Normals (NOAA National Centers for Environmental Information, 2010)		✓	

U.S. Census Bureau American Community Survey (American Community Survey 5-Year Estimates [CSV Data File], 2017)	✓	✓
Centers for Disease Control and Prevention PLACES (500 Cities Project Data [CSV Data File], 2018)		✓

Research Question 1

I used two data sets to answer research question one. The data sets include the League of American Bicyclists Community Survey (K. McLeod et al., 2019) and PeopleForBikes Bicycle Network Analysis (Bicycle Network Analysis [Data], 2020). A description of each dataset is provided below.

Community Survey, League of American Bicyclists

The League of American Bicyclists (LAB) distributes a biennial Community Survey to the 50 most populous U.S. cities to gather data on bicycle demographics, infrastructure, policies, and programs (K. McLeod et al., 2019). The LAB Community Survey is a self-reported data source collected from city representatives applying for the Bicycle Friendly Community award. The LAB distributed the 2017 Community Survey to all cities that had responded to the Community Survey in prior years and to city contacts who had submitted Bicycle Friendly Community applications within the previous year. The LAB publishes a Benchmarking Report biennially to share the results of the Community Survey. The 50 most populous cities with data published in the 2019 report are mapped in Figure 3.1.

Figure 3.1: The fifty most populous cities in the U.S. based on 2016 ACS estimates (K. McLeod et al., 2019)

The most recent Benchmarking Report was published by the LAB in 2019 and includes data collected as recently as 2017. When compiling data for the 2019 Benchmarking Report, the LAB chose to include as much data as possible even if a community did not provide a response to the 2017 Community Survey. For this reason, the data collection year for data published in the 2019 Benchmarking Report varies by city and may precede 2017. A complete record of data collection year is provided in Table 3.2.

Table 3.2: Year of bicycle infrastructure data included in the 2018 LAB Benchmarking Report

Year	City Count	City Names
2017	26	Atlanta GA, Boston MA, Denver CO, Fort Worth TX, Houston TX, Jacksonville FL, Kansas City MO, Las Vegas NV, Long Beach CA, Louisville KY, Mesa AZ, Oakland CA, Omaha NA, Oklahoma City OK, Philadelphia PA, Phoenix AZ, Portland OR, Raleigh NC, Sacramento CA, San Antonio TX, San Francisco CA, San Diego CA, Tucson AZ, Tulsa OK, Washington DC, Wichita KS

2016	6	Albuquerque NM, Charlotte NC, Cleveland OH, Colorado Springs CO, Los Angeles CA, Miami FL
2015	13	Arlington TX, Austin TX, Columbus OH, Dallas TX, El Paso TX, Fresno CA, Indianapolis IN, Milwaukee WI, Minneapolis MN, New York NY, San Jose CA, Seattle WA, Virginia Beach VA
2014	4	Baltimore MD, Chicago IL, Memphis TN, Nashville TN,
2013	1	Detroit MI

I downloaded the following data from the 2018 LAB Community Survey: the number of miles of paved public paths, number of miles of protected & buffered bike lanes, and number of miles of "other" bike lanes (i.e., non-protected, striped, etc.).

Bicycle Network Analysis, PeopleForBikes

PeopleForBikes created a bicycle network tool called the Bike Network Analysis (BNA) that measures the quality of bike networks and their connectivity to community amenities. The BNA was derived from data provided by the U.S. Census and OpenStreetMap (PeopleForBikes, 2019). The unit of analysis for the BNA was census blocks, as delineated by the U.S. Census Bureau's 2010 Decennial Census. In addition to U.S. Census block-level data, PeopleForBikes used spatial data from OpenStreetMaps (OSM) to map bicycle facilities and community amenities and, ultimately, derived bicycle network scores by city. The BNA bike network score accounts for bike network connectivity to six amenity types: people (access to other people based on population distribution), opportunity (access to jobs and educational institutions), core services (access to critical services like health care and grocery stores), recreation (access to public recreation like parks and trails), retail (access to shopping), and transit (access to major transit hubs). OSM is a crowdsourced data set that is constantly updated by the public. The BNA downloaded the most recent OSM data for the area within a city's boundary to calculate bike network scores. I downloaded the BNA bike network scores, BNA transit score, and the number of low-stress miles from bna.peopleforbikes.org in January 2020.

Research Question 2

I used ten data sets to answer research question two. The data sets include the League of American Bicyclists Community Survey, PeopleForBikes Bicycle Network Analysis, and eight additional datasets. A description of each dataset is provided below.

Community Survey, League of American Bicyclists

The League of American Bicyclists (LAB) distributes a biennial Community Survey to the 50 most populous U.S. cities to gather data on bicycle demographics, infrastructure, policies, and programs (K. McLeod et al., 2019). A detailed description of the LAB Community Survey is provided under the *Research Question 1* section above. I downloaded the following bicycle infrastructure data from the LAB Community Survey for research question two: the number of miles of paved public paths, number of miles of protected & buffered bike lanes, number of miles of "other" bike lanes (i.e., non-protected, striped, etc.), the number of full-time city employees working on bike or pedestrian issues per 100,000 population, and the number of League of American Bicyclist member organizations per 100,000 population.

Bicycle Network Analysis, PeopleForBikes

PeopleForBikes created a bicycle network tool called the Bike Network Analysis (BNA) that measures the quality of bike networks and their connectivity to community amenities. A detailed description of BNA data sources is provided under the *Research Question 1* section above. I downloaded the BNA bike network scores, BNA transit score, and the number of low-stress miles from bna.peopleforbikes.org in January 2020.

American Community Survey, U.S. Census

The American Community Survey (ACS) is a national survey that collects social, economic, housing, and demographic data across the United States (*Guidance for Data Users*, 2020). The U.S. Census Bureau distributes the ACS monthly via mail and collects a sample size of about 3.5 million households per year. ACS 1-year estimates are data that have been collected over a 12 month period while ACS 5-year estimates are data that have been collected over a 60 month period.

The U.S. Census Bureau provides aggregate ACS data to the general public for analysis purposes. Aggregate data is estimated for a range of geographic areas including states, counties, metropolitan statistical areas, cities and small-area estimates for census tracts and block groups. Census tracts are small subdivisions of counties with 2500 to 8000 residents and are designed to follow the boundaries of neighborhoods, encompassing areas that are homogenous with regard to population characteristics, economic status, and living conditions (U.S. Census Bureau, 2020).

I downloaded ACS data from factfinder.census.gov in January 2020. The ACS data I downloaded for research question two includes city estimates of bicycle commuting rates, vehicle ownership, and college student population. The ACS city estimates of bicycle commuting rates, vehicle ownership, and college student population have commonly been used when deriving aggregate models of bicycle commuting (Buehler & Pucher, 2012; Dill & Carr, 2003; Nelson & Allen, 1997; Schoner & Levinson, 2014).

Best Complete Streets, Smart Growth America

The National Complete Streets Coalition, a program of Smart Growth America, evaluates Complete Streets policies annually. The National Complete Streets Coalition began evaluating policies in 2006 and publishes ranked scores for hundreds of cities as they pass Complete Streets legislation. Smart Growth America publishes the Complete Streets policy scores in annual reports. I used the scores published in two reports: The Best Complete Streets Policies of 2016 and The Best Complete Streets Policies of 2018 (Atherton et al., 2018). The scores were provided in a pdf table in the reports, but I requested a spreadsheet directly from Smart Growth America's Program Manager in August 2019.

As of August 2019, Smart Growth America had scored policies for 48 of the 50 most populous cities in my analysis. The scores for Baltimore and Milwaukee were retrieved from the 2018 report while the scores for the remaining cities were retrieved from the 2016 report. The policy rubric was updated in 2018 with stricter requirements for equity and implementation. Therefore, the scores for Baltimore and Milwaukee were rated on a different scale than the other cities in my analysis. According to Smart Growth America's Program Manager, the organization evaluates policies annually, as the legislation is passed, so the organization does not plan to re-grade old policies using the new rubric.

Lifting the Veil, League of American Bicyclists

The League of American Bicyclists (LAB) analyzed Statewide Transportation Improvement Programs (STIP) in 2015 to report on the distribution of transportation budgets for bicycle and pedestrian projects (K. McLeod, 2017). The LAB analyzed transportation budgets by state and provided a breakdown of the budget allocated to bicycle and pedestrian projects with various project types. The Appendix of the LAB report provides a detailed list of the STIP data source used for each state. The report reveals that STIP data was publicly accessible on the DOT website for some states while STIP data for other states was provided by a DOT staff member. The LAB reported the results of their STIP analysis in the form of a pdf scorecard for each state, but I requested a spreadsheet directly from the League of American Bicyclists Policy Director.

The spreadsheet included data on the total cost of bicycle projects, the percentage of total cost, and the average project cost.

U.S. Climate Normals, NOAA

The National Centers for Environmental Information has generated U.S. Climate Normals since 1950 to meet requirements of the World Meteorological Organization and the National Weather Service. The U.S. Climate Normals dataset provides information about climate conditions for thousands of locations across the United States. The official normals are calculated every 30 years and consist of annual, monthly, daily, and hourly averages of temperature, precipitation, and other climatological variables from 15,000 weather stations. The 1991-2020 U.S. Climate Normals are the most up-to-date normals available for the U.S. The U.S. Climate Normal data is accessible at ncei.noaa.gov/products/us-climate-normals. I downloaded annual precipitation data from airport weather stations for the 50 most populous cities in January 2020. Buehler and Pucher (2012) also used U.S. Climate Normals in their model of bicycle commuting rates.

Sprawl Indices, Ewing & Hamidi

Reid Ewing at the University of Utah developed a sprawl index in 2002 that combines 22 variables of urban form, land use mix, network density, and network connectivity (Ewing & Hamidi, 2013). The sprawl index was approximated by metropolitan statistical area, but the measure is also useful for comparing land-use characteristics of central cities because it considers several measures of downtown strength and overall urban compactness (Buehler & Pucher, 2012). The most updated sprawl indices were calculated using population estimates from the U.S. Census Bureau’s 2010 Decennial Census. A full account of the methodology used to generate the sprawl index is provided at https://gis.cancer.gov/tools/urban-sprawl/sprawl-report-short.pdf. I downloaded the 2010 MSA sprawl index dataset in January 2020 from gis.cancer.gov/tools/urban-sprawl. Buehler and Pucher (2012) also used Ewing's sprawl index in their model of bicycle commuting rates.

Fatality Analysis Reporting System, USDOT

The Fatality Analysis Reporting System (FARS) is a national census that provides the National Highway Traffic Safety Administration with annual data on fatal injuries resulting from motor vehicle crashes (USDOT Federal Highway Administration, 2018). FARS reports on the annual number of fatal injuries involving bicyclists and pedestrians per state. Data is publicly available from 1975 to 2019 at nhtsa.gov/research-data/fatality-analysis-reporting-system-fars. I downloaded the FARS datasets from 2015, 2016, and 2017 in January 2020 (*Fatality Analysis Reporting System [CSV Data File]*, 2017). Buehler and Pucher (2012) also used FARS data to measure cyclist fatalities in their model of bicycle commuting rates.

Research Question 3

I used two data sets to answer research question three. The data sets include the American Community Survey and the Behavioral Risk Factor Surveillance System, compiled via the PLACES data project. A description of each dataset and the PLACES project is provided below.

American Community Survey, U.S. Census

The American Community Survey (ACS) is a national survey that collects social, economic, housing, and demographic data across the United States (*Guidance for Data Users*, 2020). A detailed description of the ACS is provided under the *Research Question 2* section above. The ACS data I downloaded for research question 3 includes socioeconomic and demographic data on sex, age, race, ethnicity, occupation type, education, and marital status. I downloaded the ACS data from factfinder.census.gov in March 2020.

Behavioral Risk Factor Surveillance System, CDC

The Behavioral Risk Factor Surveillance System (BRFSS) is a self-reported, national household survey administered biennially by the CDC via telephone (Centers for Disease Control and Prevention, 2014). The BRFSS collects data from over 400,000 adults on health-related risk behaviors, chronic health conditions, and use of preventative services. The dataset has been weighted by the CDC to be nationally representative. I compiled health measures from the PLACES dataset which the CDC derived from BRFSS data (500 Cities Project Data [CSV Data File], 2018). A description of the PLACES dataset is provided in the following paragraph.

PLACES (500 Cities), CDC

PLACES, formerly known as 500 Cities, is a collaboration between the CDC and Robert Wood Johnson Foundation that uses data from the CDC Behavioral Risk Factor Surveillance System (BRFSS) to derive census tract estimates of health outcomes. PLACES provides a model-based population level analysis by linking geocoded, county-level BRFSS data with block-level U.S. Census data to generate health outcome estimates at the census tract-level (Centers for Disease Control and Prevention, 2018). The CDC has conducted studies to validate the methodology used to generate PLACES data (Y. Wang, 2017; Zhang et al., 2014), and the PLACES dataset has been applied in a variety of studies that assess the association between built environment factors and health outcomes (Mullenbach, 2018). I accessed the archive of PLACES (500 Cities) data in March 2020 and downloaded the “500 Cities: Census Tract-level Data, 2018 release” from https://chronicdata.cdc.gov/browse?category=500+Cities+%26+Places.

Variables

A complete summary of the variables I used in my dissertation are provided in Table 3.3. The table lists the data source corresponding with each variable and provides a checklist to indicate which variables were used for each research question. The remainder of the section describes, in detail, the unit measures of each variable. The variable descriptions were organized based on their use as independent or dependent variables.

Table 3.3: Summary of variables by research question

Variable Name	RQ1	RQ2	RQ3	Source
Bike lane supply	✓	✓		League of American Bicyclists (LAB) Benchmarking Report (K. McLeod et al., 2019)
Bike path supply	✓	✓		
Bike lane quality	✓	✓		
Bike path quality	✓			
City employees working on bicycle projects		✓		
Cycling interest groups		✓		
Bike network supply	✓			Bicycle Network Analysis (Bicycle Network Analysis [Data], 2020)
Bike network density	✓	✓		
Bike network quality	✓			
Access to public transit	✓	✓		
Complete streets policy score		✓		Smart Growth America's Best Complete Streets Initiatives (Atherton et al., 2018)
Bicycle infrastructure budget		✓		LAB Lifting the Veil on Bicycle and Pedestrian Spending (K. McLeod, 2017)
Cyclist safety		✓		USDOT Fatality Analysis Reporting System 2015-2017, averages (Fatality Analysis Reporting System [CSV Data File], 2017)

Sprawl index	✓		Sprawl Indices (Ewing & Hamidi, 2013)
Annual precipitation	✓		NOAA 1981-2010 U.S. Climate Normals (NOAA National Centers for Environmental Information, 2010)
Bicycle commuting rate	✓	✓	ACS 2013-2017, averages (American Community Survey 5-Year Estimates [CSV Data File], 2017)
Vehicle ownership	✓		
College student population	✓		
Sex (Male)		✓	
Age (>45)		✓	
Household Income		✓	
Race (White)		✓	
Ethnicity (Hispanic)		✓	
Education (No College Education)		✓	
Occupation (Manual Labor)		✓	
Marital Status (Married)		✓	
Coronary Artery Disease Prevalence		✓	CDC PLACES, 2017 (500 Cities Project Data [CSV Data File], 2018)
Access to Healthy Food		✓	
Smoking Prevalence		✓	
Obesity Prevalence		✓	

Hypertension Prevalence	✓
Diabetes Prevalence	✓

Research Question 1

This section defines eight infrastructure variables that were evaluated for research question one.

Independent variables: Infrastructure

I explored eight infrastructure variables: bike lane supply, bike path supply, bike lane quality, bike path quality, bike network supply, bike network quality, bike network accessibility to transit, and bike network density. I gathered the data for these variables from the League of American Bicyclists Benchmarking Report (2019) and the PeopleForBikes (PfB) bicycle network analysis. I selected the eight infrastructure variables based on previous U.S. studies of bicycle commuting which found that, in large cities, both infrastructure supply and infrastructure networks were significantly associated with bicycle commuting rates (Buehler & Pucher, 2012; Dill & Carr, 2003; Schoner & Levinson, 2014).

I also used recommendations from these studies to explore measures of bike lane quality, bike path quality, and bike network accessibility to public transit. The methods I used when selecting these variables were similar to other cycling mode share studies that select previously tested variables in combination with new variables to improve the descriptive quality of the model (Arellana et al., 2020; Buehler & Pucher, 2012; Schoner & Levinson, 2014). A summary of the eight infrastructure variables I considered for model development are provided in Table 3.4. The following paragraphs will describe, in detail, how each variable was measured.

Table 3.4: Summary of infrastructure variables

Variable Name	Units	Source
Bike lane supply	Miles of bike lanes in city per 100,000 population	League of American Bicyclists (LAB) Benchmarking Report (K. McLeod et al., 2019)
Bike path supply	Miles of paved multi-use paths in city per 100,000 population	
Bike lane quality	Percent of bike lanes that are buffered or protected	
Bike path quality	Percent of paved multi-use paths out of the total supply of bike lanes and paths	
Bike network supply	Miles of low stress roadway for cyclists per 100,000 population	Bicycle Network Analysis (Bicycle Network Analysis [Data], 2020)
Bike network density	Miles of low stress roadway for cyclists per city area	
Bike network quality	Score of the bike network based on connectivity, directness, and fragmentation	
Access to public transit	Score of major transit hubs accessible on the low stress bike network	

The LAB Benchmarking Report provides the following data on bicycle infrastructure: the number of miles of paved public paths, number of miles of protected & buffered bike lanes, and number of miles of "other" bike lanes (i.e., non-protected, striped, etc.). I used data from the 2017 LAB Community Survey to measure the following infrastructure variables: bike lane supply, bike path supply, bike lane quality, and bike path quality.

Bike Lane Supply

Bike lane supply was defined as the miles of bike lanes per 100,000 population. Buehler and Pucher (2012) defined bike lane supply in the same manner. I calculated bike lane supply by summing the LAB data for miles of protected & buffered lanes with miles of "other" lanes and dividing by population size within city government jurisdiction.

Bike Path Supply

Bike path supply was defined as the miles of paved multi-use paths per 100,000 population. Buehler and Pucher (2012) defined bike lane supply in the same manner. I calculated bike path supply by dividing the LAB data for miles of paved public paths by population size within city government jurisdiction.

Bike Lane Quality

Bike lane quality was defined as the percentage of buffered and protected bike lane miles. I calculated bike lane quality by dividing the LAB data for miles of protected & buffered lanes with the total miles of bike lanes accounted for in the LAB report. I included a measure of bicycle lane quality based on recommendations from Buehler and Pucher (2012) and Buehler and Dill (2016).

Bike Path Quality

Bike path quality was defined as the percentage of paved multi-use path miles. I calculated bike path quality by dividing the LAB data for miles of paved public paths with the total miles of bike lanes and paths accounted for in the LAB report. I included a measure of bicycle path quality based on recommendations from Buehler and Pucher (2012) and Buehler and Dill (2016).

I used measures from the PfB bicycle network analysis to define the remaining infrastructure variables of: bike network supply, bike network quality, bike network accessibility to transit, and bike network density. PeopleForBikes conducted their bicycle network analysis through a procedural traffic stress analysis and destination analysis. The purpose of the traffic stress analysis was to categorize street segments and intersections as “low stress” or “high stress” for cyclists while the purpose of the destination analysis was to determine the variation of destinations accessible via a “low stress” bicycle network (PeopleForBikes, 2019). PeopleForBikes used the results from their traffic stress analysis and destination analysis to calculate city-level bicycle network scores.

PeopleForBikes’ traffic stress analysis categorized roadway miles as “low stress” or “high stress” based on methodology from the Mineta Transportation Institute (MTI). MTI’s categorization scheme considers the speed limit and number of lanes within any given street segment. As the number of lanes and speed limit increases, cyclists’ stress also increases (Mekuria & Nixon, 2012). However, PfB expanded upon MTI’s categorization scheme in their traffic stress analysis by including bike lane type and presence of on-street parking within their categorization scheme. For example, within PfB’s analysis, a buffered bike lane on a street with a speed limit greater than 35 mph would be considered “high stress” for cyclists. Whereas a shared residential roadway with a speed limit of 20 mph or less would be considered “low stress” for cyclists.

After categorizing low and high stress miles with the traffic stress analysis, PfB conducted a destination analysis to evaluate, by census block, the number of surrounding census blocks that would be accessible by bike on the “low stress” bicycle network. PfB defined a bikeable distance as the distance that an average bicyclist could cover in ten minutes at ten miles per hour, approximately 1.67 miles, measured along streets and paths. Destinations were only assumed to be accessible by bike on the “low stress” network if the route did not require a cyclist to go out of their way by more than 25% in comparison to a car trip. PeopleForBikes calculated city-level bicycle network scores based on the number and types of destinations accessible on the "low stress" bicycle network. PeopleForBikes has calculated bicycle network scores and destination scores for hundreds of U.S. cities.

Bike Network Supply

I used PfB's categorization of "low stress" bicycle network miles to measure bike network supply. I calculated bike network supply as the number of "low stress" miles per 100,000 population.

Bike Network Density

I also used PfB's categorization of "low stress" bicycle network miles to measure bike network density. I calculated bike network density as the number of "low stress" miles per city area.

Bike Network Quality

I used PfB's overall bicycle network score to measure bike network quality.

Bike Network Accessibility to Transit

I used PfB's transit destination score as a measure of bike network accessibility to transit. The PfB bicycle network score that I used to define bike network quality represents a city's bicycle network ranking based on PfB's traffic stress and destination analyses. The PfB measure that I used to define bike network accessibility to transit represents a transit accessibility ranking based on the number of major transit stops that would be accessible on the "low stress" bicycle network.

Research Question 2

This section defines the dependent variable, independent variables, and covariates that were used to answer research question two. The eight infrastructure variables, defined under research question one, were also used for research question two, in addition to policy-relevant variables and covariates. The dependent variable for research question two was bicycle commuting rate.

Dependent variable: Bicycle commuting rate

Bicycle commuting rate was the dependent variable in my model. I compiled the data for bicycle commuting rate from the American Community Survey (ACS), the only publicly available U.S. dataset that provides annual, cross-regional data on cycling to work (Buehler et al., 2020; Buehler & Pucher, 2012). The ACS asks the following question to collect data on bicycle commuting, “How did you usually get to work last week?” Therefore, the dependent variable represents the number of commuters cycling as their primary mode of transportation to work. I averaged bicycle commuting rates over a five-year period such that the dependent variable was defined as the average number of bicycle commuters per 10,000 population from 2013 to 2017.

Independent variables: Infrastructure

The same infrastructure variables described under research question one were used to model bicycle commuting rates for research question two.

Independent variables: Policy

I included four policy-relevant variables when modeling bicycle commuting rates. I also considered the eight infrastructure variables defined under RQ1 (see Table 3.4). In this section, I will provide details about the policy-relevant variables. A summary of the policy-relevant variables is provided in Table 3.5. The following paragraphs will describe, in detail, how each variable was measured.

Table 3.5: Summary of policy-relevant variables

Variable Name	Units	Source
Complete streets policy score	Score of complete streets policy based on ten ranking categories: vision and intent, diverse users, commitment, design, land use and context sensitivity, project selection criteria, jurisdiction, exceptions, performance measures, and implementation steps	Smart Growth America's Best Complete Streets Initiatives (Atherton et al., 2018)
Bicycle infrastructure budget	State data: percent of transportation spending budgeted for bicycle-only projects (FFY 2013-2016)	LAB Lifting the Veil on Bicycle and Pedestrian Spending (K. McLeod, 2017)
City employees working on bicycle projects	Number of full-time city employees working on bike or pedestrian issues per 100,000 population	LAB Benchmarking Report (K. McLeod et al., 2019)
Cycling interest groups	Number of League of American Bicyclist member organizations per 100,000 population	

Complete Streets Policy Score

The variable for complete streets policy score was provided by Smart Growth America who have scored complete streets policies for hundreds of U.S. cities using a ten-category ranking scheme (Riveron, 2019). The ten categories include: vision and intent, diverse users, commitment, design, land use and context sensitivity, project selection criteria, jurisdiction, exceptions, performance measures, and implementation steps. Their score was available for 48 of the 50 cities in my analysis. The main benefits of SGA's score for the purpose of modeling bicycle commuting rates were: 1) the implementation category was weighted most heavily, and 2) the score was reduced by four-points if a policy does not address cycling within their vision and intent statement. A measure of Complete Streets policy has not been evaluated in prior models of bicycle commuting even though Complete Streets is one of the most prevalent bicycle-related policies in the United States (Marleau Donais et al., 2019).

Bicycle Infrastructure Budget

Bicycle infrastructure budget was defined as the annual percent of state transportation budget allocated to bicycle projects. A LAB report on the intermodal distribution of state transportation budgets provided the data for this measure (K. McLeod, 2017). The LAB compiled the data provided in the report by analyzing the budgets provided by Statewide Transportation Improvement Programs. Dill and Carr (2003) found a statistically significant association between bicycle commuting rates and state spending on bicycle and pedestrian projects. Bicycle infrastructure budget in my study accounts for state spending on bicycle projects, not including pedestrian projects.

City Employees Working on Bicycle Projects

City employees working on bicycle projects was defined as the number of full-time equivalent (FTE) city employees working on bicycle or pedestrian issues per 100,000 population. The LAB Benchmarking Report provided the data for this measure (K. McLeod et al., 2019). When collecting this data, the LAB defined FTE employees as employees spending at least one tenth of their time on bicycle or pedestrian issues. In recent years, bicycle and pedestrian coordinators have cited the number of employees working on bicycle projects as a top barrier for implementation of bicycle projects (Dill et al., 2017).

Cycling Interest Groups

Cycling interest groups were defined as the number of LAB member organizations per 100,000 population. The LAB Benchmarking Report provided the data for this measure (K.

McLeod et al., 2019). The efforts of *cycling interest groups* can influence the adoption of bicycle policies (Aytur et al., 2013; Dill et al., 2017; Moreland-Russell et al., 2013).

Covariates: Safety, sprawl, precipitation, vehicle ownership & college population

I included five covariates when modeling bicycle commuting rates. The covariate measures were selected based on statistically significant determinants of bicycle commuting from four U.S studies (Buehler & Pucher, 2012; Dill & Carr, 2003; Nelson & Allen, 1997; Schoner & Levinson, 2014). Refer to the Literature Review section for a complete review of covariates. A summary of the covariates is provided in Table 3.6. The covariate measures were downloaded directly from the sources listed in Table 3.6 (no additional calculations were performed).

Table 3.6: Summary of covariates for research question two

Variable Name	**Units**	**Source**
Cyclist safety	State data: three-year average number of bicyclist fatalities per 10,000 bicycle commuters	USDOT Fatality Analysis Reporting System 2015-2017, averages (*Fatality Analysis Reporting System [CSV Data File]*, 2017)
Sprawl index	MSA data: index combining 22 variables measuring residential density, mix of land uses, strength of downtowns, and connectivity of street network (Note: higher scores=less sprawl)	Sprawl Indices (Ewing & Hamidi, 2013)
Annual precipitation	30-year average annual number of rainfall days with 0.01 inches or more	NOAA 1981-2010 U.S. Climate Normals (NOAA National Centers for Environmental Information, 2010)

Vehicle ownership	Percent of households without a motorized vehicle	ACS 2013-2017, averages (American Community Survey 5-Year Estimates [CSV Data File], 2017)
College student population	Percent of total population enrolled in college or university	

Research Question 3

This section defines the dependent variable, independent variables, and covariates that were used to answer research question three.

Dependent variable: Coronary artery disease prevalence

The dependent variable was coronary artery disease (CAD) prevalence. I used the CDC 500 Cities Project, census tract estimates from 2017 BRFSS data to measure CAD prevalence(*500 Cities Project Data [CSV Data File]*, 2018). BRFSS measures CAD prevalence based on responses to the following question, “Has a doctor, nurse, or other health professional ever told you that you had angina or coronary artery disease?” Therefore, CAD prevalence was defined in my study as the percentage of respondents age 18 years or older who reported ever having angina or CAD.

Independent variable: Bicycle commuting rate

The independent variable was bicycle commuting rate. The data for bicycle commuting was compiled from the American Community Survey and was the same data that was used to measure bicycle commuting rates for research question two (American Community Survey 5-Year Estimates [CSV Data File], 2017). The ACS asks the following question to collect data on bicycle commuting, “How did you usually get to work last week?” Therefore, the variable represents the number of commuters who cycle as their primary mode of transportation to work.

I averaged bicycle commuting rates for each census tract over a five-year period to determine the average number of bicycle commuters per 10,000 population from 2013 to 2017.

Covariates: Socioeconomic, demographic characteristics & health risk factors

I considered 13 covariates when evaluating the association between bicycle commuting and CAD prevalence. The covariates represent demographic, socioeconomic, and health risk factors with potential influence on the association between bicycle commuting and CAD prevalence (Celis-Morales et al., 2017; Hu et al., 2007; Wennberg et al., 2006). A summary of the covariates is provided in Table 3.7. The covariate measures were downloaded directly from the sources listed in Table 3.7 (no additional calculations were performed).

Table 3.7: Summary of covariates for research question three

Variable Name	Variable Units	Source
Sex (Male)	Percent of the population that is male	ACS 2013-2017, averages (American Community Survey 5-Year Estimates [CSV Data File], 2017)
Age (>45)	Percent of the population that is greater than 45 years of age	
Household Income	Median value of household income ($)	
Race (White)	Percent of the population that is white alone	
Ethnicity (Hispanic)	Percent of the population that is of Hispanic origin	
Education (No College Education)	Percent of the population that has not completed a college degree	

Occupation (Manual Labor)	Percent of the population with manual labor occupations (including Census categories of production, materials moving, natural resources, construction, and grounds/ building cleaning or maintenance)	
Marital Status (Married)	Percent of the population that is married	
Access to Healthy Food	Percent of the population living more than ½ mile from the nearest supermarket, supercenter, or large grocery store	(500 Cities Project Data [CSV Data File], 2018)
Smoking Prevalence	Percent of respondents aged 18 years or older who reported smoking at least 100 cigarettes in their lifetime	
Obesity Prevalence	Percent of respondents aged 18 years or older who are obese (BMI>30)	
Hypertension Prevalence	Percent of respondents aged 18 years or older who reported ever having high blood pressure	
Diabetes Prevalence	Percent of respondents aged 18 years or older who reported having diabetes	

Statistical Analysis

Various statistical analyses were conducted to answer the research questions. The analysis methodologies for each research question are described in this section. Bivariate correlations and principle component analysis were used for research question one; ordinary least squares regression models and k-fold cross validation were used for research question two; latent class cluster analysis and prevalence odds ratios were calculated for research question three.

Research Question 1

Bivariate correlations and principal component analysis were conducted to identify infrastructure variables that would be representative of bicycle commuting rates. Spearman's correlation coefficient was computed to evaluate bivariate correlations. The bivariate correlations were used to determine the relative strength of association between bicycle commuting and infrastructure variables. Infrastructure variables with correlation coefficients of at least 0.3 were assumed to be representative of bicycle commuting rates. Several studies of bicycle commuting rates have computed bivariate correlations in combination with a regression analysis, to bolster conclusions about the significance of independent variables (Braun et al., 2019; Buehler & Pucher, 2012). Previous studies found correlation coefficients as high as 0.5.

Infrastructure network variables, like bike network density, and infrastructure supply variables, like bike lane supply, have historically been evaluated in separate models of bicycle commuting rates (Buehler & Pucher, 2012; Dill & Carr, 2003; Schoner & Levinson, 2014). These models found that network and supply variables were significantly associated with bicycle commuting rates. However, the relative significance of network and supply variables has been inconclusive because their association with bicycle commuting has historically been evaluated in separate models. To evaluate the effects of network and supply variables in a combined model, I performed a principal components analysis (PCA) with varimax rotation. PCA was utilized to condense infrastructure variables for bicycle network and urban sprawl models in prior studies (Ewing & Hamidi, 2013; Schoner & Levinson, 2014). The PCA procedure ensured that all predictive factors were independent by calculating orthogonal single unit vectors, eigenvectors, for each dimension of the analysis (Shlens, 2014). The results of the PCA in combination with the bivariate correlations informed a reduction in the infrastructure variables used to represent bicycle commuting rates.

Research Question 2

Ordinary Least Squares (OLS) regression models were derived to model bicycle commuting rates. I performed a log transform on the dependent and independent variables in order to meet OLS assumptions. There are three assumptions of OLS regression models: 1) residuals are normally distributed, 2) residuals have constant variance, and 3) residuals are independent (Field et al., 2012). In addition to meeting assumptions for normally distributed residuals, a log transform of both the dependent and independent variables allowed for interpretation of the beta coefficients as percent changes in bicycle commuting rates. Similarly, Buehler and Pucher (2012) used a log-transformed OLS regression model when they evaluated the association between bike lanes and bicycle commuting rates.

Before deriving the regression models, I conducted a multiple imputation to generate data for missing points (Braun et al., 2016; Fitch et al., 2019; Moudon et al., 2005). I conducted multiple imputation using the predictive mean matching method and the MICE function in R version 3.6.3 (van Buuren, 2020). Bike lane and bike path data was missing from the LAB Community Survey for the cities of Detroit and Philadelphia. I attempted to request data directly from the municipalities, but I did not receive a response. In turn, multiple imputation accounted for 1% of the data used in the regression models, 10 data points in total. Appendix A contains a figure showing the value of each estimated data point over five imputations. The x-axis ranges from 0 to 5 representing data generated for each imputation, and the y-axis is the natural log of each variable. The figure in Appendix A demonstrates that all imputed data points (black) fall within the distribution of existing data (grey).

To evaluate the relative significance of infrastructure and policy variables, I derived six regression models using a stepwise procedure. The stepwise regression models were organized by the following variable groupings: 1) infrastructure variables alone, 2) policy variables alone, 3) infrastructure and policy variables, 4) infrastructure variables and covariates, 5) policy variables and covariates, and 6) a backward elimination regression where all variables were initially input to the model. I derived all regression models in R version 3.6.3 using the lm() function. Schoner and Levinson (2014) and Buehler and Pucher (2012) also used a stepwise regression procedure when modeling bicycle commuting rates.

I evaluated the generalizability of the backward elimination regression model using a k-fold cross-validation procedure (Uyanık et al., 2020). A theoretical overview of the k-fold cross validation procedure can be found in Burman (1989). I conducted the k-fold cross validation (k = 4) in R version 3.6.3 using the *build_model* function from the *regclass* package (Petrie, 2016a). The first step of the k-fold cross validation was randomly dividing the data sample into training and holdout sets with an 80-20 split (Ton et al., 2020). The second step was deriving a series of regression models from the training set, where the number of regression models was equivalent to 2^k and where k was equal to the number of independent variables. A generalization error was computed with each regression model. The final step of the k-fold cross validation was computing an actual generalization error from the holdout sample and comparing it with the predicted generalization error of the seventh regression model. The seventh regression model met two criteria: 1) it had the smallest generalization error and 2) it had the fewest number of independent variables (Petrie, 2016a). The k-fold cross validation procedure is sometimes called a predictive modeling procedure because the resulting regression can be used to make predictions on new datasets (Petrie, 2016b). K-fold cross validation has been used in transportation studies to validate the determinants of bicycle crashes, bicycle commuting rates, walkability, and ride-share demand (Chen et al., 2020; Sabouri et al., 2020; Ton et al., 2020; Yencha, 2019).

Research Question 3

I conducted an ecologic study to evaluate the association between bicycle commuting rates and CAD prevalence. Ecologic studies have been used in the public health field to understand how geographic and environmental context affects population health (Susser, 1994). In the past, researchers have conducted ecologic studies to evaluate CAD prevalence in relation to groundwater quality, economics, and health resource distribution (Ferreira-Pinto et al., 2012; McLeod et al., 2018). In an ecologic study, the units of analysis are populations rather than individuals, and the exposure is a property of the population (Aschengrau & Seage, 2014c). In my study, bicycle commuting rate was the exposure and census tracts were the unit of analysis.

I considered 12,322 census tracts within the 50 most populous U.S. cities for my study sample. Approximately 60 census tracts were missing data on bicycle commuting or CAD prevalence. First, I imputed the missing data using the predictive mean matching method and the MICE function in R version 3.6.3 (van Buuren, 2020). Then, I compared the imputed sample (n=12,322) with the sample where missing data had been removed (n=12,261). I compared the statistical association between CAD prevalence and bicycle commuting for both samples by computing two sets of regression models. The beta coefficients for the association between CAD prevalence and bicycle commuting were equivalent across both models (β= -0.214), so I used the imputed data set for analysis. Appendix A contains the aforementioned regression models. During sample selection, I also determined that CAD prevalence had a statistically normal distribution after removing two outlier census tracts. The outlier census tracts were located in

New York City and contained populations where 90% of the residents were 45 years or older. Appendix A shows scatterplots of the distribution when outliers were included versus when outliers were removed. After removing the outlier tracts from the sample (*n*=12,320), the beta coefficient (β= -0.213) was still equivalent to the beta coefficient of the imputed sample (β= -0.214). Appendix A contains the regression model after outliers were removed. The associations between CAD prevalence and bicycle commuting were equivalent for both the imputed sample and the outlier sample, so I did not include the outlier census tracts in the analysis. Overall, the sample I used for analysis contained imputed data for approximately 60 census tracts and excluded two outlier census tracts (*n*=12,320).

The first steps of my statistical analyses were a descriptive and stratified analysis. I conducted a stratified analysis to evaluate the presence of effect modifiers and confounders. Within clinical practice, effect modification is assessed to identify whether the effect of an exposure significantly varies between populations with different characteristics (Corraini et al., 2017). In other words, effect modification is a change in the strength of association between an exposure and a disease according to the level of a third variable (Aschengrau & Seage, 2014b). Effect modification is sometimes called statistical interaction. In contrast, a confounder is sometimes called a control variable or covariate. A confounder must meet three criteria: 1) it must be associated with the exposure, 2) it must be an independent cause or predictor of the disease, and 3) it cannot be an intermediate step in the causal pathway between the exposure and disease (Aschengrau & Seage, 2014a). A decision tree for evaluating effect modification and confounding with stratified analyses can be found in Aschengrau and Seage (2014) on p. 359.

In accordance with the stratified analysis decision tree, I conducted a crude analysis of the association between CAD and bicycle commuting (Aschengrau & Seage, 2014b). Then, I conducted a stratified analysis that separated the crude data by level of potential effect modifier/confounder. To conduct the stratified analysis, I dichotomized the prevalence data for each potential effect modifier/confounder by splitting data at the median values of each variable. Effect modification was present if the percent difference between stratum-specific estimates was greater than 10%. (Aschengrau & Seage, 2014b). If the stratum-specific estimates were less than 10% different, I evaluated confounding by statistically comparing the pooled estimate with the crude estimate using the Mantel-Haenszel Test (Aschengrau & Seage, 2014a). After identifying effect modifiers, I input the dichotomized variables into a latent class cluster analysis (LCA) to group census tracts with similar characteristics. Similar methods of dichotomization and LCA have been used in a study that evaluated the effect of the food environment, food security, income, and education on food acquisition and healthy shopping habits (Ma et al., 2018).

The purpose of the LCA was to group census tracts by CAD health risk. This approach applies the precedent established by Wilson et al. (1998) who developed the Framingham risk score (FRS) to approximate CAD health risk. The FRS has been used by the National Cholesterol Education Program of the National Institute of Health as an "office-based approach to estimate and stratify an individual's absolute short term risk of a CAD event," (National Institute of Health, 2001). The FRS has also been applied in Finnish studies that evaluate the association between bicycle commuting and CAD (Hu et al., 2007).

The FRS has traditionally been calculated by accounting for smoking, diabetes, hypertension, and cholesterol levels (Wilson et al., 1998). My approach accounts for smoking, diabetes, obesity, and access to healthy food. As a population study, I did not have access to biomarker data on cholesterol levels or blood pressure measurements. I used obesity, in terms of body mass index (BMI > 30), as a proxy for these variables because greater obesity has been associated with high LDL-C, low HDL-C, and high blood pressure (Hu et al., 2007). I also used USDA's access to healthy food measurement as an indicator of hypertension. Proximity to supermarkets has been shown to increase healthy intake of fruits and vegetables and has been associated with less obesity and hypertension (Morland et al., 2002, 2006). Access to healthy food has been especially important for the health of low-income communities where motor vehicle access and transportation options are limited (Mackett & Thoreau, 2015; Wang et al., 2007).

The FRS does not account for socioeconomic status (SES) or racial identity, which has been tied to SES in U.S. active transportation studies (Antonakos et al., 2020; Braun et al., 2019). The Framingham study sample, which informed the calculation of the FRS, was taken from a Boston suburb where the majority of participants were white (Hu et al., 2007). However, my sample incorporates 12,320 census tracts across 50 U.S. cities, so the population represented by my sample was more diverse than the Framingham sample. For this reason, I considered variables of SES and racial identity, in addition to the health indicators traditionally used, when categorizing census tracts by CAD health risk. A study conducted in Evans County, Georgia found that racial identity, education, income, and occupation were associated with CAD prevalence (McDonough et al., 1965). I incorporated racial identity, education, income, and occupation as indicators of SES in my analysis.

After using LCA to group census tracts with similar health risks and SES, I computed measures of association for each LCA cluster. Researchers evaluating CAD incidence typically derive logistic regressions to calculate prevalence odds ratios of the outcome to exposure risk (Celis-Morales et al., 2017; Riiser et al., 2018; Wilson et al., 1987). However, the data available for my study was CAD prevalence, not incidence. To calculate prevalence odds ratios, I dichotomized CAD prevalence based on the median value (*med*=5.5). This approach was similar to a median regression, which was used to analyze health outcomes that would otherwise be highly skewed due to non-normal distributions (McGreevy et al., 2009).

I derived logistic regressions of median CAD prevalence for each LCA cluster using the glm function in R version 3.6.3. Then, I computed prevalence odds ratios based on the beta coefficient of the regression model. Bicycle commuting rate was input to the model as a continuous variable. The prevalence odds ratios presented in the results represent the odds that a census tract will have CAD prevalence greater than 5.5 percent when bicycle commuting is present.

CHAPTER FOUR: RESULTS & DISCUSSION

Research Question 1

Five variables should be used to represent bicycle infrastructure when evaluating the association with bicycle commuting rates: *access to public transit, bike lane supply, bike lane quality, bike network density, and bike path supply*. Results of the bivariate correlations and principal components analysis are discussed in the following section.

Bivariate Correlations

To identify measures of bicycle infrastructure that were representative of bicycle commuting rates, I calculated bivariate correlations between bicycle commuting rates and eight infrastructure variables. Table 4.1 presents the bivariate correlations.

Table 4.1: Descriptive statistics by infrastructure variable and their bivariate correlations with bicycle commuting rate

Variable Name	Mean	Median	SD	Min	Max	ρ
Bike lane supply	15.3	10.9	12.7	0.4	48.5	0.26
Bike path supply	10.0	8.4	6.8	0.0	29.6	-0.10
Bike lane quality	13.1	6.7	19.1	0.0	100	0.15
Bike path quality	43.5	39.6	25.3	0.0	96.4	-0.31
Bike network supply	392.9	402.3	260.7	13.1	986.6	-0.23
Bike network density	17.6	19.3	9.7	0.6	34.0	0.41
Bike network quality	23.2	22.9	11.5	3.5	57.9	0.51*
Access to public transit	13.1	8.3	13.2	0.0	57.8	0.67*

*Note: $p < 0.001$

The bivariate correlations demonstrated statistically significant, positive correlations ($p < 0.001$) between bicycle commuting rates and the following infrastructure variables: *bike network quality*, and *access to public transit*. The correlations between bicycle commuting rates, *bike lane supply* (0.3), and *bike path supply* were statistically consistent with Buehler and Pucher (2012) who found that bike lane supply was more strongly correlated with bicycle commuting rates than bike path supply. However, the correlation between *bike path supply* and bicycle commuting rate was not statistically significant.

This study is not the first to find a statistically insignificant relationship between bike paths and bicycle commuting, yet findings are mixed. Nelson and Allen (1997) and Buehler and Pucher (2012) found a statistically significant association while Dill and Carr (2003) found an insignificant association. There are a few reasons bike path supply may not be significantly correlated with bicycle commuting rates. The first being, bike path variables, in this study and in prior studies, have been defined as the number of miles of multi-use paths. Multi-use paths are created for multiple uses, not exclusively for bicycle commuters. They are often useful for increasing recreational activity and, in many cases, help support multi-modal transportation but they may not be the most efficient means for increasing bicycle commuting. In fact, some U.S. studies have demonstrated that bicycle commuters are more likely to use on-street facilities than off-street paths (Broach, Dill, and Gliebe 2012; Sener, Eluru, and Bhat 2009).

However, paths reserved exclusively for bicyclists, also known as cycle tracks, are prevalent in European cities and are widely used by bicycle commuters as an efficient means of transport (Pucher, Dill, and Handy 2010; Pucher, Buehler, and Seinen 2011). The mileage of cycle tracks in the U.S. is too small to evaluate their efficacy across multiple cities, but the number of bicycle commuters using cycle tracks likely differs from the number of bicycle commuters using multi-use paths. Unless cycle tracks become more prevalent in the U.S., measures of bike lane supply may be more representative of bicycle commuting rates than measures of bike path supply.

Similar to bike path supply, The correlation between bicycle commuting, bike path quality and bike lane quality were not statistically significant. The non-significant results may be related to the measure of bicycle commuting used in this study which only accounts for existing bicycle commuters, not persons who may be interested in adopting the behavior. Pre-existing bicycle commuters may not be as concerned about bike lane quality as those interested in adopting the behavior (Garrard, Rose, and Lo 2008; Pucher and Buehler 2017; Rossetti et al. 2018; Branion-Calles et al. 2019). Measuring a change in bicycle commuting rates over time may be more reflective of bike lane quality than the cross-sectional measure of bicycle commuting that was used in this study.

Overall, the statistical significance between bicycle commuting and bike network variables was stronger than the statistical significance with bike path and bike lane variables. In fact, bike network quality and access to public transit had the strongest correlation with bicycle commuting rates (0.51, 0.67) followed by bike network density (0.41). Connected bike networks are more likely to increase bicycle commuting rates than the addition of disconnected bike paths or bike lanes (Buehler & Dill, 2016). The measures I used for the bike network variables in this study were equivalent to the bike network score and the transit score provided by PeopleForBikes BNA. The strong correlation between bicycle commuting rates, bike network quality, and access to public transit suggests that the BNA is a valuable measure of bicycle networks. The BNA scores' statistically significant correlation with commuting rates shows promise for practitioners looking for a tool to measure the strength of bicycle networks.

I also used data from the BNA to measure bike network density and bike network supply, albeit these variables were calculated in conjunction with geographic and population data, not taken directly from the BNA. I found that the bivariate correlation between bicycle commuting rates and bike network density was strong, but not statistically significant. My results are consistent with Schoner and Levinson (2014) who found that the correlation between network density and bicycle commuting rates was statistically significant while network size was not. My results, in conjunction with the results of Schoner and Levinson (2014), suggest that a variable measuring the density of the bike network per city area was more representative of bicycle commuting rates than a variable measuring the mileage of the bicycle network per population.

Principal Components Analysis

Bicycle network and supply variables have historically been evaluated in separate models of bicycle commuting rates (Buehler & Pucher, 2012; Dill & Carr, 2003; Schoner & Levinson, 2014). These models have found that network and supply variables are both significantly associated with bicycle commuting rates. However, the relative significance is unclear because network and supply variables have historically been evaluated in separate models. In order to evaluate the effects of network and supply variables in a combined model, I performed a principal components analysis (PCA). I used PCA to assess multicollinearity between eight infrastructure variables. The factor results of the PCA helped determine which variables should be used to represent bicycle infrastructure.

Table 4.2 presents the factor results from a PCA of all eight infrastructure variables. Three of the bicycle network variables exhibited multicollinearity– *bike network supply, bike network density,* and *bike network quality*. These three network variables also accounted for the greatest proportion of variance (0.28) in bicycle commuting rates. However, one of these network variables, *bike network quality,* factored onto multiple components, indicating additional multicollinearity with *access to public transit*. The multicollinearity between bike lane and bike path variables varied. *Bike path supply* and *bike path quality* exhibited multicollinearity while *bike lane supply* and *bike lane quality* were independent.

Table 4.2: PCA of eight bicycle infrastructure variables[1]

Variables	F1	F2	F3	F4	F5	F6
Bike network supply	0.96					
Bike network density	0.91					
Bike network quality	0.70		0.50			0.50
Bike path supply		0.97				
Bike path quality		0.79				
Access to public transit			0.97			
Bike lane supply				1.00		
Bike lane quality					0.98	
Proportion Variance	**0.28**	**0.20**	**0.16**	**0.15**	**0.13**	**0.03**

[1] Loadings smaller than 0.5 are suppressed.

Table 4.3 presents results from a PCA where *bike network quality* was excluded from the analysis. I excluded bike network quality from the PCA for two reasons: 1) it factored onto the first and third principal component, and 2) it had the weakest correlation among the variables in both the first and third principal component.

Table 4.3: PCA of seven bicycle infrastructure variables[1]

Variables	F1	F2	F3	F4	F5
Bike network supply	0.96				
Bike network density	0.91				
Bike path supply		0.96			
Bike path quality		0.82			
Bike lane supply			0.99		
Access to public transit				0.98	
Bike lane quality					0.98
Proportion Variance	**0.25**	**0.23**	**0.17**	**0.16**	**0.15**

[1] Loadings smaller than 0.5 are suppressed.

After excluding *bike network quality*, the first principal component still accounted for the largest proportion of variance (0.25). Additionally, *access to public transit* factored independently and accounted for 16% of the variance. The three remaining network variables (bike network supply, bike network density, and access to public transit) accounted for a total of 41% variance in bicycle commuting rates, compared to 44% when bike network quality was included. However, multicollinearity still existed between some network variables and bike path variables. *Bike* *network supply* and *bike network density* exhibited multicollinearity within the first principal component; *bike path supply* and *bike path quality* exhibited multicollinearity in the second component.

Bike network supply and *bike network density* explained 25% of the variance, but multicollinearity prohibited simultaneous inclusion in a regression model. *Bike network density* and *bike network supply* were both calculated using the measure of low-stress miles provided by PeopleForBikes, which led to multicollinearity. When *bike network density* was evaluated independently from *bike network supply* in the bivariate correlations, the correlation between bicycle commuting rates and *bike network density* was statistically significant (see Table 4.1), while the correlation between bicycle commuting rates and bike network supply was not. The statistical significance of *bike network density* may make it a more representative measure of bicycle networks.

The variables in the second component, *bike path supply* and *bike path quality*, were both calculated from the mileage of bike paths provided by League of American Bicyclists, so they were highly dependent. The bike path variables explained a smaller proportion of variance than the bike network variables, and their correlations with bicycle commuting rates were not statistically significant (see Table 4.2). Although the correlation was not statistically significant in this study, *bike path supply* has historically been included in models of bicycle commuting rates (Buehler & Pucher, 2012; Dill & Carr, 2003; Nelson & Allen, 1997). The correlation between *bike path supply* and bicycle commuting rates was not statistically significant in this study, but the correlation was in the expected positive direction (see Table 4.1). Based on historical use and the positive correlation in this study, *bike path supply* may be a more representative measure of bike path infrastructure than *bike path quality*.

Overall, PCA revealed five independent principal components where each component accounted for at least 10% of the variance in bicycle commuting rates. Three variables were independent measures of bicycle infrastructure: *access to public transit, bike lane supply, and bike lane quality*; while four variables exhibited multicollinearity: *bike network supply, bike network density, bike path supply,* and *bike path quality*. *Access to public transit, bike lane supply*, and *bike lane quality* were independent measures because they factored onto their own components. Bike network variables (supply and density) and bike path variables were multicollinear because they factored onto the same component. To determine which of the multicollinear variables might best represent bicycle infrastructure in a model of bicycle commuting, I revisited the variables' bivariate correlation. The bivariate correlations suggested that *bike network density* and *bike path supply* may be more representative of bicycle commuting than *bike network supply* or *bike path quality* .

In summary, five infrastructure variables can be used as independent measures of bicycle infrastructure when evaluating bicycle commuting rates in U.S. cities: *access to public transit, bike lane supply, bike lane quality, bike network density, and bike path supply*. Access to public transit is a score provided by PeopleForBikes that indicates the number of major transit hubs accessible on the low stress network; bike lane supply is the number of miles of bike lanes per 100,000 population; bike lane quality is the percent of buffered or protected lanes; bike network density is the miles of low stress roadway for cyclists per city area; and bike path supply is the number of miles of multi-use paths per 100,000 population. These five variables factor independently from one another, but the network variables, in particular, accounted for 41% of the variance in bicycle commuting rates. Furthermore, the network variables: access to public transit (ρ= 0.67), bike network quality (ρ= 0.51), and bike network density (ρ= 0.41) had the strongest bivariate correlations with bicycle commuting rates. Evidence from the PCA in combination with the bivariate correlations suggest that network measures should be prioritized when using infrastructure variables to evaluate bicycle commuting rates.

Research Question 2

Three variables explained 70% of the variation in bicycle commuting rates. The three variables included cyclist safety, city employees working on bicycle projects, and access to public transit. When considering the relative significance of infrastructure and policy-related variables, neither stood out as more statistically significant. On the contrary, the results suggest that a combination of infrastructure and policy factors contribute to bicycle commuting rates. A detailed account of the results and a discussion of each variables' association with bicycle commuting rates is provided in the following section.

Stepwise Regression Models

The regression model results are presented in Table 4.4. The average variance inflation factor (VIF) for all models was less than 2 which is much lower than the recommended VIF of 5 (Sabouri et al. 2020). The low VIF of these models suggests that the procedure used to select infrastructure variables in research question one successfully reduced issues of multicollinearity. The remainder of this section will describe the detailed results of each model.

Table 4.4: Multiple regression models of bicycle commuting rate

	OLS regression of ln (bike commuters per 10,000 population)						
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**
ln (bike lane supply)	0.29**		0.07	0.16			
ln (bike path supply)	-0.11		-0.27	-0.24		-0.36**	
ln (bike network density)	0.06		0.10	-0.23		-0.16	
ln (access to public transit)	0.51**		0.33**	0.32**		0.21**	0.28**
ln (bike lane quality)	0.10		0.03	0.08			
ln (complete streets policy)		-0.20	-0.17		-0.16	-0.24	
ln (city employees)		1.03**	0.82**		0.87**	0.74**	0.66**

ln (cycling interest groups)		1.10*	0.85		0.10		
ln (infrastructure budget)		0.47	0.09		0.27		
ln (sprawl index)				0.08	0.34		
ln (annual precipitation)				0.23	0.06		
ln (vehicle ownership)				-0.01	0.46		
ln (college student pop.)				1.01	0.78		
ln (cyclist safety)				-1.02**	-0.78**	-1.04**	-0.80**
Constant	1.69	3.32	2.87	2.48	0.99	7.33	4.44
Observations	50	50	50	50	50	50	40
F-statistic	7.79**	7.54**	7.06**	9.17**	11.96**	25.56**	29.27**
Average VIF	1.2	1.1	1.6	1.9	1.5	1.6	-
Adjusted R^2	**0.41**	**0.35**	**0.53**	**0.63**	**0.67**	**0.75**	**0.69**

Note: * or **; coefficient significance at the p<0.05 or p<0.01

Model 1 evaluated the association between bicycle infrastructure variables and bicycle commuting rates. *Bike lane supply* and *access to public transit* had statistically significant associations with bicycle commuting in this model. As the mileage of bike lanes increased and accessibility to public transit improved, bicycle commuting increased. The association of bicycle commuting with *bike path supply, bike network density, and bike lane quality* were not statistically significant. Model 1 explained 41% of the variability in bicycle commuting rates (Adj R^2= 0.41).

Model 2 evaluated the association between policy-related variables and bicycle commuting rates. *City employees working on bicycle projects* and *cycling interest groups* had statistically significant associations with bicycle commuting. As the number of full-time city employees working on cycling issues increased and the number of LAB member organizations increased, bicycle commuting also increased. The association of bicycle commuting with *complete streets policy score* and *bicycle infrastructure budget* were not statistically significant. Model 2 explained 35% of the variability in bicycle commuting rates (Adj R^2= 0.35).

Model 3 was a stepwise regression that evaluated the association of bicycle commuting rates with both infrastructure and policy-related variables. *Access to public transit* and *city* *employees working on bicycle projects* remained statistically significant in this model. As *access to public transit* improved and the number of *city employees working on bicycle projects* increased, bicycle commuting also increased. The association of bicycle commuting with *bike path supply, bike lane supply, bike network density, bike lane quality, cycling interest groups, complete streets policy score,* and *bicycle infrastructure budget* were not statistically significant. Model 3 explained 53% of the variability in bicycle commuting (Adj R^2= 0.53).

Model 4 was a stepwise regression that evaluated the association of bicycle commuting rates with covariates and infrastructure variables, while Model 5 was a stepwise regression that evaluated the association with covariates and policy-related variables. *Access to public transit* and *city employees working on bicycle projects* remained statistically significant in models 4 and 5. The adjusted R^2 of both models increased after adding covariates to the models (Adj R^2= 0.63, Adj R^2= 0.67). However, *cyclist safety* was the only covariate that had a statistically significant association with bicycle commuting rates. *Sprawl index, annual precipitation, vehicle ownership,* and *college student population* were not statistically significant in model 4 or 5.

Model 6 was a backward elimination regression where all the variables listed in Table 4.5 were input to the model. Table 4.5 provides descriptive statistics of each variable input to model 6 and the corresponding correlation coefficients with bicycle commuting rates.

Table 4.5: Descriptive statistics of 50 U.S. cities and their bivariate correlations with bicycle commuting rate

Variable Name	Mean	SD	Min	Max	r
Bike lane supply	15.3	12.7	0.4	48.5	0.30*
Bike path supply	10.0	6.8	0.0	29.6	0.04
Bike lane quality	13.1	19.1	0.0	100	0.12
Bike network density	17.6	9.7	0.6	34.0	0.38*
Access to public transit	13.1	13.2	0.0	57.8	0.54*
Complete streets policy score	58.4	21.4	15.2	92.8	-0.22
Bicycle infrastructure budget	0.46	0.57	0.0	3.5	0.25

City employees	1.2	1.7	0.0	10.4	0.59*
Cycling interest groups	0.3	0.3	0.0	0.9	0.37*
Cyclist safety	11.6	6.4	0.8	23.4	-0.58*
Sprawl index	101.7	30.2	41.0	203.4	0.33*
Annual precipitation	56.2	22.1	11.2	91.0	0.15
Vehicle ownership	7.8	4.0	4.2	29.9	0.23
College student population	7.6	1.1	5.6	9.4	0.13

*Note: $p<0.05$

Six variables emerged from the backward elimination (see Table 4.4): *cyclist safety, city employees working on bicycle projects, bike path supply, complete streets policy, access to public transit,* and *bike network density*. The statistically significant variables included: *access to public transit, bike path supply, city employees,* and *cyclist safety*. *Access to public transit, city employees,* and *cyclist safety* had the strongest bivariate correlations with bicycle commuting (see Table 4.1) and were statistically significant in models 1-5. Although *bike path supply* was statistically significant in model 6, its bivariate correlation with bicycle commuting was equivalent to zero and it was not statistically significant in models 1-5. Model 6 explained 75% of the variability in bicycle commuting (Adj R^2= 0.75).

Regression Model Validation

I evaluated the robustness of model 6 by excluding outlier cities in a re-estimation of the model. I used Cook's distance (*D*) statistic to identify El Paso and Tucson as outlier cities (Stevens 1984). When El Paso and Tucson were excluded from the sample (n=48), *city employees working on bicycle projects* and *cyclist safety* remained statistically significant, but *access to public transit* was no longer statistically significant. The mixed significance of *access to public transit* after removing outlier cities raised questions of generalizability. A k-fold cross validation procedure was performed to assess generalizability.

Model 7 was derived using a repeated k-fold cross validation procedure. The k-fold procedure evaluated the generalization error of over 16,000 potential regressions. Figure 4.1

summarizes the k-fold cross validation results. In figure 4.1, the estimated generalization error (RMSE) is plotted against the number of regression variables. Figure 4.1 depicts the estimated generalization error for the model with the lowest generalization error at each number of regression variables. The dashed horizontal line distinguishes models within one standard deviation of the lowest generalization error. The “X” marks the model with the lowest generalization error, and the “O” marks the model with the least number of variables and a generalization error that is within one standard deviation of the lowest error. A five variable model had the lowest generalization error (marked by ‘X’); a three-variable model had a generalization error within one standard deviation of the five-variable model (marked by ‘O’ in Figure 4.1). Table 4.4 lists the beta coefficients for the three-variable model (model 7). I tested the predictive power of model 7 by comparing the estimated generalization error ($RMSE_{estimate}$) with the actual generalization error ($RMSE_{actual}$). The percent difference between $RMSE_{actual}$ (0.615) and $RMSE_{estimate}$(0.572) was less than 10 percent.

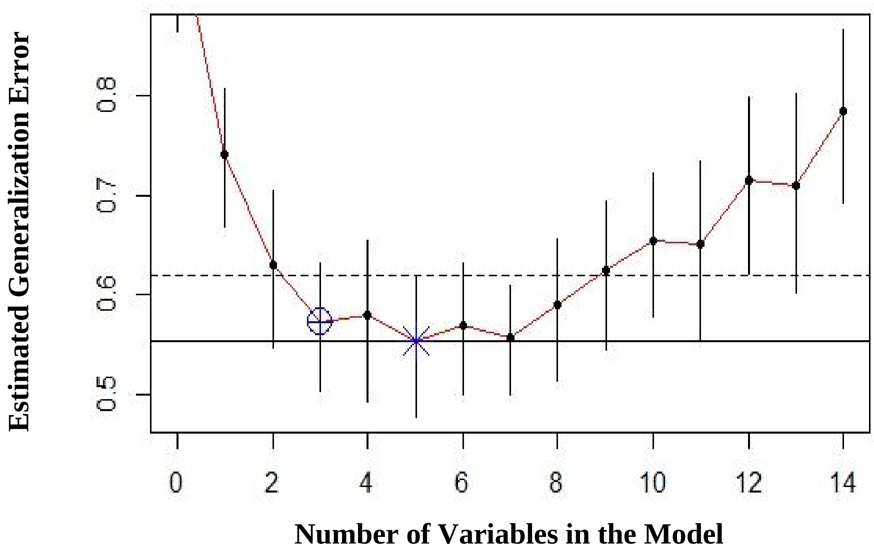
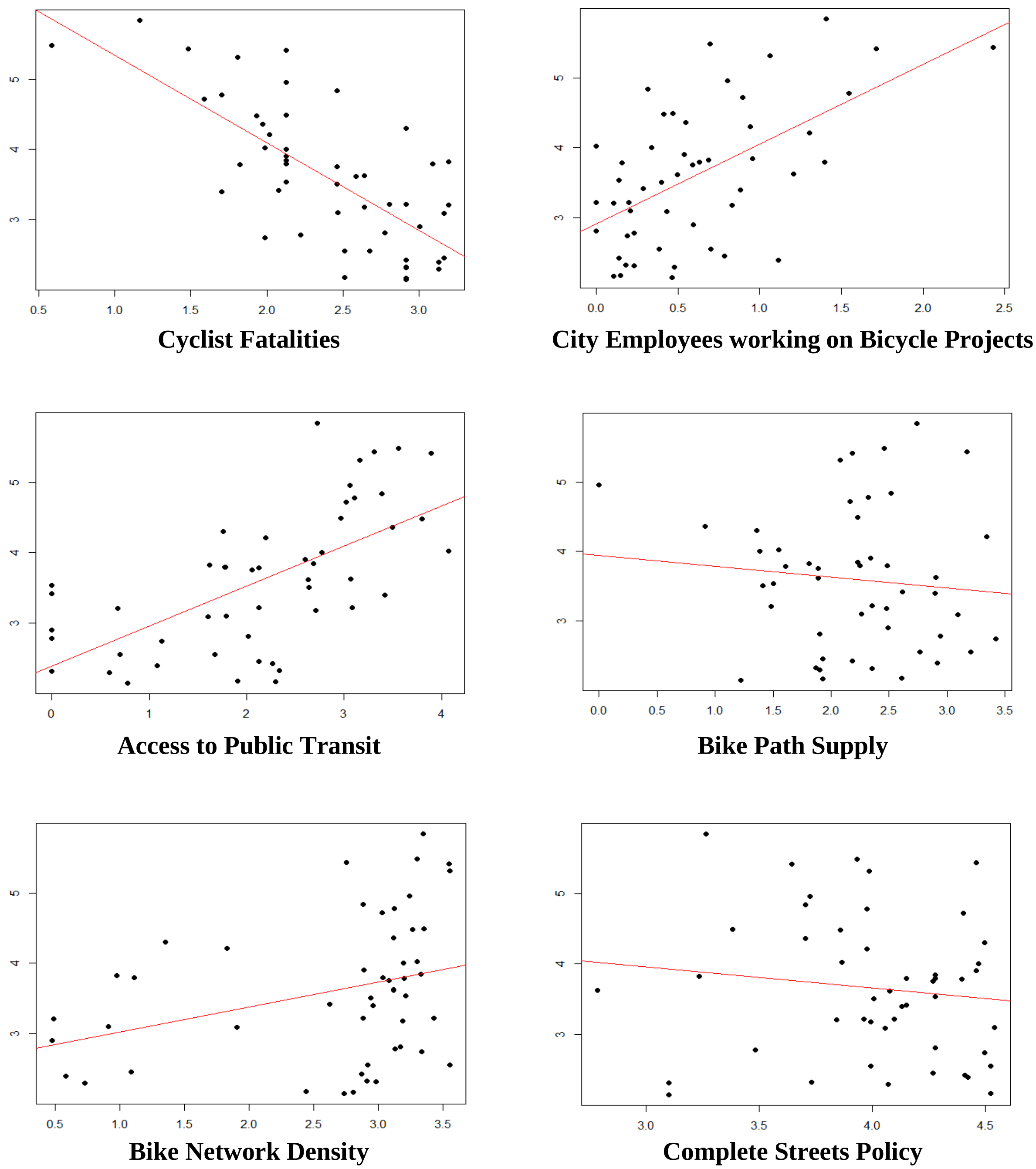


Figure 4.1: Estimated generalization error of k-fold cross validation regression models

The three variables retained in model 7 were: *access to public transit, city employees working on bicycle projects*, and *cyclist safety*. These three variables were consistent with the statistically significant variables in models 1-6 and had the strongest bivariate correlations with bicycle commuting (see Table 4.1). The results of model 7 were interpreted as follows: 1) a 10% higher access to transit score was associated with a 2.8% greater number of bike commuters per 10,000 population; 2) a 10% greater number of employees working on cyclist issues per 100,000 population was associated with a 6.6% greater number of bike commuters per 10,000 population, and 3) a 10% higher cyclist fatality rate per 10,000 commuter cyclists was associated with 8% fewer bike commuters per 100,000 population.

Model 7 explained 70% of the variance in bicycle commuting rates while model 6, the backward elimination regression, explained 75% of the variance. The minor decrease in adjusted R^2 values from model 6 to model 7 (0.75, 0.69) suggests that the three additional variables included in model 6 (bicycle path supply, complete streets policy, and bike network density) were minimally significant to the overall model of bicycle commuting rates. Scatterplots of the correlations between bicycle commuting and model 6 variables are shown in Figure 4.2.

Figure 4.2: Scatterplots of the correlation between bicycle commuting rates and model 6 variables



The scatterplots demonstrate the strength of association between bicycle commuting and the three variables that were retained in model 7 (access to public transit, cyclist safety, and city employees working on bicycle projects). In contrast, the associations between bicycle commuting and three of the variables included in model 6 (bicycle path supply, complete streets policy, and bike network density) were weak. The beta coefficient for bike path supply in model 6 was statistically significant. However, in a scatterplot of bike path supply, it was apparent that the association had been calculated from a skewed sample. Bike path supply for all but three of the cities in the study sample was greater than 3 miles per 100,000 population. Cities with bike path supply less than 3 miles would need to be included in the study sample in order to determine a reliable estimate of the association between bike path supply and bicycle commuting rate. Overall, the scatterplots demonstrated weak associations between bicycle commuting and the variables that were excluded from model 7 which reinforces the validity of the variables that were retained in model 7. In summary, three variables best predicted bicycle commuting rates in this study of the 50 largest U.S. cities: cyclist safety, city employees working on bicycle projects, and access to public transit.

The covariate, cyclist safety, had the strongest association with bicycle commuting (β= -0.80). Several transport models have cited safety as a primary barrier to the uptake of bicycle commuting (Aziz et al. 2018; Adam, Jones, and te Brömmelstroet 2020; Arellana et al. 2020). Policies that reduce roadway fatalities and support safe streets for cyclists, like Vision Zero and speed limit reductions, will help increase bicycle commuting rates (Prati et al. 2018).

The policy-related variable, city employees working on bicycle projects, had the second strongest association with bicycle commuting (β= 0.66). In 1999, the ISTEA required each state DOT to have a bicycle and pedestrian coordinator, but it did not require a cyclist workforce (Pucher, Komanoff, and Schimek 1999). In recent years, bicycle and pedestrian coordinators have cited the number of employees working on bicycle projects as a top barrier for implementation of bicycle projects (Dill, Smith, and Howe 2017). Support for bicycle projects is needed from more than just one enthusiastic champion to increase bicycle commuting rates (Weber 2017; Assunçao-Denis and Tomalty 2019).

The infrastructure variable, access to public transit, had the third strongest association with bicycle commuting (β= 0.28). A study conducted in the mid-Atlantic region of the U.S. supports the significance of *access to public transit.* The study found that public transit users reported less infrastructure barriers to bicycle commuting than non-transit users (Bopp, Gayah, and Campbell 2015). Transportation officials in large cities should focus on building robust bike networks with multi-modal transit connections.

Cyclist safety, city employees working on bicycle projects, and access to public transit accounted for 70% of the variance in bicycle commuting rates. When considering the relative significance of infrastructure and policy-related variables, neither stood out as more statistically significant than another. On the contrary, the results suggest that a combination of infrastructure and policy factors contribute to bicycle commuting rates in large U.S. cities.

Research Question 3

The odds of living in a census tract with above average CAD prevalence (>5.5%) decreased by 60% (OR= 0.40) when the bicycle commuting rate was greater than zero. After controlling for socioeconomic status, race/ethnicity, and health risk factors within a latent class cluster analysis, the association between bicycle commuting and CAD prevalence was only statistically significant in census tracts with high socioeconomic status, predominately white race, and above average health. In census tracts with high SES and above average health, the odds of living in a census tract with above average CAD prevalence (>5.5%) decreased by 20% (AOR=0.80) when the bicycle commuting rate was greater than zero. Obesity and diabetes prevalence were the strongest effect modifiers of the association between bicycle commuting and

CAD prevalence. The percent difference between the stratified prevalence odds ratios with obesity and diabetes were 58% and 42%, respectively.

Descriptive Analysis

Bivariate correlations with coronary artery disease

I investigated the association between bicycle commuting and coronary artery disease (CAD) alongside thirteen CAD risk factors. The thirteen CAD risk factors represented demographic, socioeconomic, and health risk factors. Table 4.6 lists descriptive characteristics of the census tract sample (n= 12,322) for each risk factor. The descriptive characteristics include the mean, median, standard deviation, and Pearson's *r* correlation coefficient. The correlation coefficient was calculated from the association with CAD prevalence.

Table 4.6: Descriptive characteristics by CAD risk factor and their bivariate correlations with CAD prevalence

Potential Risk Factor	Mean	Med	SD	r[1]
Bicycle Commuting (%)	1.2	0	2.2	−0.23
Sex (% Male)	48.9	48.8	5.0	−0.14
Age (% > 45)	36.6	36.3	10.5	0.44
Married (%)	6.6	5.9	4.0	0.09
White (%)	54.4	60.7	28.7	−0.28
Black (%)	24.1	9.4	29.8	0.37
Latinx (%)	26.8	16.3	26.2	0.03
Household Income ($1K)	77.9	66.3	45.7	−0.49
No College Education (%)	42.1	42.3	18.5	0.56
Manual Labor Job (%)	8.9	8.1	5.5	0.25
Healthy Food Access (%)	41.6	36.2	38.2	0.09
Smoking (%)	18.1	17.2	6.3	0.58
Obesity (%)	30.2	29.3	8.6	0.60
Diabetes (%)	11.1	10.5	4.5	0.85
Hypertension (%)	30.5	29.1	8.6	0.86

[1]All Pearson *r* correlation coefficients significant at $p<0.0001$

The average bicycle commuting rate across the 12,322 census tract sample was 1.2% with a standard deviation of 2.2%. The median bicycle commuting rate was zero percent. There was a negative correlation between bicycle commuting rate and CAD prevalence; as bicycle commuting rate increased, CAD prevalence decreased (r=-0.23). The Pearson's r correlation between CAD prevalence and bicycle commuting was greater than the Pearson's r correlation between CAD prevalence and male sex, marital status, Latinx ethnicity, or food access.

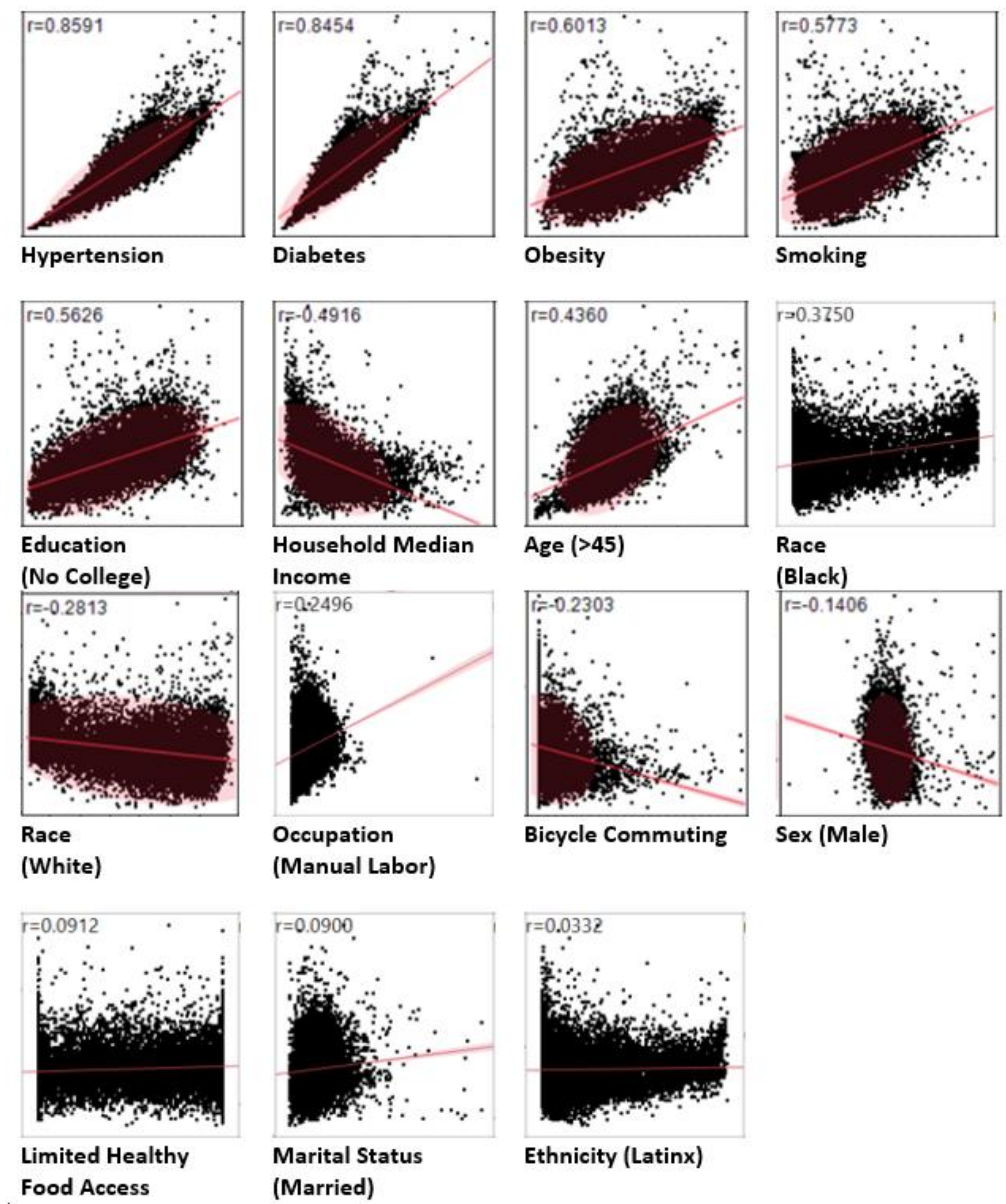
The demographic factors, age (% older than 45 years) and race (% white, % black), had greater Pearson's r correlation with CAD prevalence than bicycle commuting. There was a positive correlation between age, Black race, and CAD prevalence. In other words, as the percent of census tract population older than 45 years increased and the percent of Black population increased, CAD prevalence also increased (r = 0.44, r = 0.37). In contrast, CAD prevalence decreased as the percent of White population increased (r = -0.28).

The socioeconomic factors, household income, education, and occupation type, also had greater Pearson's r correlation with CAD prevalence than bicycle commuting. There was a negative correlation between household income and CAD prevalence; as household income increased, CAD prevalence decreased (r =-0.49). There was a positive correlation between the percent of population with no college education, the percent with manual labor occupations, and CAD prevalence. As the percent of the population without a college education increased and the percent of the population with manual labor occupations increased, CAD prevalence also increased (r = 0.56, r = 0.25).

The health risk factors, smoking, obesity, diabetes, and hypertension, had the strongest Pearson's r correlation with CAD prevalence. As smoking prevalence, obesity, diabetes, and hypertension prevalence increased, CAD prevalence also increased (r = 0.58, r = 0.60, r = 0.85, r

= 0.86). The scatterplots in Figure 4.3 show the bivariate correlations between CAD prevalence and each risk factor. The Pearson's *r* correlation coefficient is provided in the top left corner of each scatterplot and the scatterplots are ordered left to right from strongest correlation to weakest correlation.

Figure 4.3: Scatterplots of the bivariate correlations between CAD prevalence and demographic, socioeconomic, and health risk factors

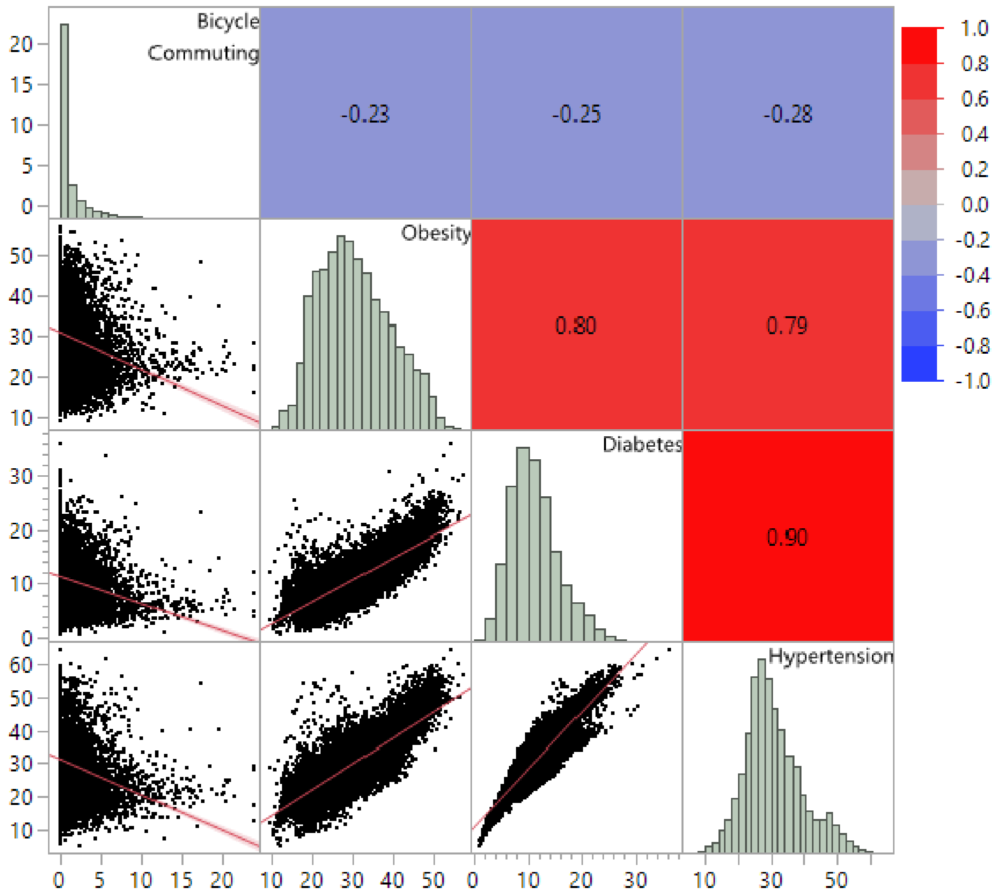


Hypertension, diabetes, obesity, and smoking had the highest Pearson's *r* correlation coefficients with CAD prevalence, and they are well known risk factors of CAD from the Framingham Study (Friedman, 1994). Age and male sex were also identified as risk factors of CAD during the Framingham study (Jekel et al., 1996). However, the bivariate correlation between age and CAD prevalence in my study was lower than the correlation with education, income, or chronic disease. Additionally, the correlation between male sex and CAD prevalence was low in my study. Measuring male sex at a population level likely led to a low correlation since there was little variation in the percent of males at a population level. Indicators of social class like education, income, occupation, and race have also been associated with CAD and were associated with CAD prevalence in my study (Cassel, 1971; Marmot & Theorell, 1988; McDonough et al., 1965). Male sex, food access, marital status, and ethnicity were not correlated with CAD prevalence at a population level.

Bivariate correlations with obesity, diabetes, and hypertension

The bivariate correlations graphed in Figure 4.4 explore obesity, diabetes, and hypertension as modifiers of the association between bicycle commuting and CAD prevalence. The first column of Figure 4.4 shows scatterplots of the bivariate correlations between bicycle commuting, obesity, diabetes, and hypertension prevalence. Bicycle commuting was negatively correlated with the three factors. The bivariate correlations between bicycle commuting and obesity (r = -0.23), bicycle commuting and diabetes (r= -0.25), and bicycle commuting and hypertension prevalence (r= -0.28) were similar to the correlation between bicycle commuting and CAD prevalence (r = -0.23). The remainder of the scatterplots in Figure 4.4 show the interactions between obesity and diabetes (r= 0.80), obesity and hypertension (r= 0.79), and diabetes and hypertension (r= 0.90). The strong correlations between obesity, diabetes, and hypertension indicated high multicollinearity. The correlation coefficients depicted in Figure 4.4 suggest that obesity, diabetes, and hypertension may modify the relationship between bicycle commuting and CAD prevalence.

Figure 4.4: Scatterplots of the bivariate correlations between bicycle commuting rate, diabetes, hypertension, and obesity prevalence



To further explore obesity, diabetes, and hypertension as mediating factors, I derived a regression model of CAD prevalence with bicycle commuting, obesity, diabetes, hypertension, and their interaction effects as the independent variables. The regression model output has been provided in Appendix B. The interaction effects between bicycle commuting, obesity, diabetes, and hypertension were statistically significant in the regression model, suggesting that there is a mediating relationship between bicycle commuting and the three variables. The following section compares the mediating effects of obesity, diabetes, and hypertension with other CAD risk factors.

Stratified Analysis

I conducted a stratified analysis to evaluate the effects of modifiers and confounders on the association between bicycle commuting and CAD prevalence. The results were calculated with bicycle commuting considered as a protective factor against above average CAD prevalence. Results of the crude analysis are provided in Table 4.7. According to the crude analysis, the odds of living in a census tract with above average CAD prevalence decreased by 60% (OR= 0.40) when bicycle commuting was present. The thirteen CAD risk factors included in my study were evaluated as effect modifiers and confounders in Tables 4.8- 4.20.

Table 4.7: Census tract CAD prevalence by bicycle commuting rate

Exposure	CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	2438	4071	6509
Bike Commuting	3468	2345	5813
Total	5906	6416	12322
Prevalence odds ratio	**0.40**		

Results from analyzing male sex as a potential modifier or confounder are provided in Table 4.8. Male sex was neither an effect modifier nor a confounder because the prevalence odds ratios differed by less than 10% and the Mantel-Haenszel p-value was not statistically significant. Male sex is a risk factor of CAD and is commonly associated with bicycle commuting, so in a cohort or case-control study it would likely be an effect modifier (Donaire-Gonzalez et al., 2015; Reis et al., 2013; Sallis et al., 2013). However, in an ecological study, such as this one, male sex was aggregated at a population level and was not an effect modifier.

Table 4.8: Census tract CAD prevalence by sex and bicycle commuting rate

Percent male < 49%				Percent male > 49%			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1194	2397	3591	No bike commuting	1244	1674	2918
Bike commuting	1437	1157	2594	Bike commuting	2031	1188	3219
Total	2631	3554	6185	Total	3275	2862	6137
Prevalence odds ratio	**0.40**			**Prevalence odds ratio**	**0.43**		**8% difference**

Results from analyzing age as a potential modifier or confounder are provided in Table 4.9. Age, and more specifically, the percent of the population older than 45 years of age, was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent (Δ = 22%). The odds of living in a census tract with above average CAD prevalence decreased by 53% (AOR=0.47) in census tracts where the majority of the population was older than 45 years and bicycle commuting was present. Age is a risk factor for CAD and is often associated with bicycle commuting rates (Donaire-Gonzalez et al., 2015; Reis et al., 2013; Sallis et al., 2013).

Table 4.9: Census tract CAD prevalence by age and bicycle commuting rate

Percent with age 45+ < 36%				Percent with age 45+ > 36%			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1366	1610	2976	No bike commuting	1072	2461	3533
Bike commuting	2220	989	3209	Bike commuting	1248	1356	2604
Total	3586	2599	6185	Total	2320	3817	6137
Prevalence odds ratio	**0.38**			**Prevalence odds ratio**	**0.47**		**22% difference**

Results from analyzing marital status as a potential modifier or confounder are provided in Table 4.10. Marital status was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent (Δ = 35%). The odds of living in a census tract with above average CAD prevalence decreased by 53% (AOR=0.47) in census tracts where a below average percentage of the population was married vs. 67% (AOR=0.33) in census tracts where an above average percentage of the population was married. Marriage can be an indicator of socioeconomic status and, in this case, appears to amplify the odds of CAD prevention from bicycle commuting (Donaire-Gonzalez et al., 2015; Kaplan & Keil, 1993; Orth-Gomér et al., 2000).

Table 4.10: Census tract CAD prevalence by marital status and bicycle commuting rate

Percent married < 6%				**Percent married > 6%**			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1438	1910	3348	No bike commuting	1000	2161	3161
Bike commuting	1716	1082	2798	Bike commuting	1752	1263	3015
Total	3154	2992	6146	Total	2752	3424	6176
Prevalence odds ratio 0.47				**Prevalence odds ratio 0.33**			**35% difference**

Results from analyzing education as a potential modifier or confounder are provided in Table 4.11. Education, and more specifically, the percent of the population without a college education, was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent (Δ = 36%). The odds of living in a census tract with above average CAD prevalence decreased by 38% (AOR=0.62) in census tracts where an above average percentage of the population was not college educated and bicycle commuting was present. Education is an indicator of socioeconomic status and college education has been positively associated with bicycle commuting(Stinson & Bhat, 2003; Tribby & Tharp, 2019; Wardman et al., 2007) (Cole-Hunter et al., 2015b; Donaire-Gonzalez et al., 2015; Reis et al., 2013; Sallis et al., 2013).

Table 4.11: Census tract CAD prevalence by education and bicycle commuting rate

Percent with no college education < 42%				Percent with no college education > 42%			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1534	1067	2601	No bike commuting	904	3004	3908
Bike commuting	2731	823	3554	Bike commuting	737	1522	2259
Total	4265	1890	6155	Total	1641	4526	6167
Prevalence odds ratio 0.43				**Prevalence odds ratio 0.62**			**36% difference**

Results from analyzing income as a potential modifier or confounder are provided in Table 4.12. Median household income was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent (Δ = 20%). The odds of living in a census tract with above average CAD prevalence decreased by 50% (AOR=0.50) in census tracts where the median household income was less than $66,000 and bicycle commuting was present. Income is an indicator of socioeconomic status and has been associated with bicycle commuting rates (Kaplan & Keil, 1993; Stinson & Bhat, 2003; Tribby & Tharp, 2019; Wardman et al., 2007).

Table 4.12: Census tract CAD prevalence by household income and bicycle commuting rate

Median household income < 66.3K				Median household income > 66.3K			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	795	2911	3706	No bike commuting	1643	1160	2803
Bike commuting	865	1590	2455	Bike commuting	2603	755	3358
Total	1660	4501	6161	Total	4246	1915	6161
Prevalence odds ratio 0.50				**Prevalence odds ratio 0.41**			**20% difference**

Results from analyzing occupation as a potential modifier or confounder are provided in Table 4.13. Occupation, and more specifically, the percent of the population working a manual labor occupation, was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent (Δ = 61%). The odds of living in a census tract with above average CAD prevalence decreased by 68% (AOR=0.32) in census tracts where a below average percentage of the population works manual jobs vs. 40% (AOR=0.60) in census tracts where an above average percentage of the population works manual jobs. Manual labor can be an indicator of chronic stress and, in this case, appears to dampen the odds of prevention from bicycle commuting (Kaplan & Keil, 1993; Karasek et al., 1988; Marmot & Theorell, 1988; Timio et al., 1997).

Table 4.13: Census tract CAD prevalence by occupation and bicycle commuting rate

Percent with manual labor occupation < 8%				**Percent with manual labor occupation > 8%**			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1268	1575	2843	No bike commuting	1170	2496	3666
Bike commuting	2359	931	3290	Bike commuting	1109	1414	2523
Total	3627	2506	6133	Total	2279	3910	6189
Prevalence odds ratio	**0.32**			**Prevalence odds ratio**	**0.60**		**61% difference**

Results from analyzing ethnicity as a potential modifier or confounder are provided in Table 4.14. Ethnicity, and more specifically, the percent of the population with Latinx ethnicity, was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent (Δ = 73%). The odds of living in a census tract with above average CAD prevalence decreased by 73% (AOR=0.27) in census tracts where a below average percentage of the population was Latinx vs. 41% (AOR=0.59) in census tracts where an above average percentage of the population was Latinx. Ethnicity can be an indicator of socioeconomic status and, in this case, appears to dampen the odds of prevention from bicycle commuting (Braun et al., 2019; MacMahon & Trichopoulos, 1996; Sallis et al., 2013).

Table 4.14: Census tract CAD prevalence by ethnicity and bicycle commuting rate

Percent Latinx < 16%				Percent Latinx > 16%			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1108	2153	3261	No bike commuting	1330	1918	3248
Bike commuting	1890	1005	2895	Bike commuting	1578	1340	2918
Total	2998	3158	6156	Total	2908	3258	6166
Prevalence odds ratio	**0.27**			**Prevalence odds ratio**	**0.59**		**73% difference**

Results from analyzing race as a potential modifier or confounder are provided in Table 4.15. Race, and more specifically, the percent of the population who are white, was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent ($\Delta =$ 13%). The odds of living in a census tract with above average CAD prevalence decreased by 55% (AOR=0.45) in census tracts where a below average percentage of the population was white and bicycle commuting was present. Race can be an indicator of socioeconomic status and, in this case, appears to dampen the odds of prevention from bicycle commuting (Braun, 2021, p. 202; McDonough et al., 1965; Sallis et al., 2013).

Table 4.15: Census tract CAD prevalence by race and bicycle commuting rate

Percent white < 61%				Percent white > 61%			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1149	2400	3549	No bike commuting	1289	1671	2960
Bike commuting	1348	1261	2609	Bike commuting	2120	1084	3204
Total	2497	3661	6158	Total	3409	2755	6164
Prevalence odds ratio	**0.45**			**Prevalence odds ratio**	**0.39**		**13% difference**

Results from analyzing food access as a potential modifier or confounder are provided in Table 4.16. Food access, and more specifically, the percent of the population with access to healthy food, was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent (Δ = 63%). The odds of living in a census tract with above average CAD prevalence decreased by 43% (AOR=0.57) in census tracts where the percentage of the population with healthy food access was below average vs. 71% (AOR=0.29) in census tracts where the percentage of the population with healthy food access was above average. Access to healthy food can be an indicator of obesity and appears to amplify the odds of prevention from bicycle commuting (Mackett & Thoreau, 2015; Morland et al., 2002, 2006; M. C. Wang et al., 2007).

Table 4.16. Census tract CAD prevalence by food access and bicycle commuting rate

Percent with access to healthy food < 36%				**Percent with access to healthy food > 36%**			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1529	2294	3823	No bike commuting	909	1777	2686
Bike commuting	1266	1076	2342	Bike commuting	2202	1269	3471
Total	2795	3370	6165	Total	3111	3046	6157
Prevalence odds ratio	**0.57**			**Prevalence odds ratio**	**0.29**		**63% difference**

Results from analyzing smoking as a potential modifier or confounder are provided in Table 4.17. Smoking prevalence was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent (Δ = 45%). The odds of living in a census tract with above average CAD prevalence decreased by 34% (AOR=0.66) in census tracts where smoking prevalence was above average and bicycle commuting was present. Smoking is a risk factor of CAD and dampens the odds of prevention from bicycle commuting (Donaire-Gonzalez et al., 2015; Kaczynski, 2008; Wilson et al., 1998).

Table 4.17: Census tract CAD prevalence by smoking and bicycle commuting rate

Smoking prevalence < 17%				**Smoking prevalence > 17%**			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1712	1297	3009	No bike commuting	726	2774	3500
Bike commuting	2904	923	3827	Bike commuting	564	1422	1986
Total	4616	2220	6836	Total	1290	4196	5486
Prevalence odds ratio 0.42				**Prevalence odds ratio 0.66**			**45% difference**

Results from analyzing obesity as a potential modifier or confounder are provided in Table 4.18. Obesity prevalence was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent ($\Delta = 58\%$). The odds of living in a census tract with above average CAD prevalence decreased by 21% (AOR=0.79) in census tracts where obesity prevalence was above average and bicycle commuting was present. Obesity is a risk factor of CAD and dampens the odds of prevention from bicycle commuting (Friedman, 1994; Huy et al., 2008; Lusk et al., 2010; Ming Wen & Rissel, 2008).

Table 4.18: Census tract CAD prevalence obesity and bicycle commuting rate

Obesity prevalence < 29%				**Obesity prevalence > 29%**			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1390	1010	2400	No bike commuting	1048	3061	4109
Bike commuting	2836	892	3728	Bike commuting	632	1453	2085
Total	4226	1902	6128	Total	1680	4514	6194
Prevalence odds ratio 0.43				**Prevalence odds ratio 0.79**			**58% difference**

Results from analyzing diabetes as a potential modifier or confounder are provided in Table 4.19. Diabetes prevalence was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent ($\Delta = 42\%$). The odds of living in a census tract with above average CAD prevalence decreased by 27% (AOR=0.73) in census tracts where diabetes prevalence was above average and bicycle commuting was present. Diabetes is a risk factor of CAD and dampens the odds of prevention from bicycle commuting (Friedman, 1994; Huy et al., 2008; Pucher, Buehler, et al., 2010a; Wilson et al., 1998).

Table 4.19: Census tract CAD prevalence by diabetes and bicycle commuting rate

Diabetes prevalence < 11%				**Diabetes prevalence > 11%**			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	1717	779	2496	No bike commuting	721	3292	4013
Bike commuting	2958	639	3597	Bike commuting	510	1706	2216
Total	4675	1418	6093	Total	1231	4998	6229
Prevalence odds ratio	**0.48**			**Prevalence odds ratio**	**0.73**		**42% difference**

Results from analyzing hypertension as a potential modifier or confounder are provided in Table 4.20. Hypertension prevalence was an effect modifier because the stratified prevalence odds ratios differed by more than ten percent (Δ = 17%). The odds of living in a census tract with above average CAD prevalence decreased by 56% (AOR=0.44) in census tracts where hypertension prevalence was above average and bicycle commuting was present. Hypertension is a risk factor of CAD and dampens the odds of prevention from bicycle commuting (Friedman, 1994; Hu et al., 2002; Huy et al., 2008; Wilson et al., 1998).

Table 4.20: Census tract CAD prevalence by hypertension and bicycle commuting rate

Hypertension prevalence < 29%				**Hypertension prevalence > 29%**			
	CAD < 5.5%	CAD > 5.5%	Total		CAD < 5.5%	CAD > 5.5%	Total
No bike commuting	673	558	1231	No bike commuting	1765	3513	5278
Bike commuting	1185	362	1547	Bike commuting	2283	1983	4266
Total	1858	920	2778	Total	4048	5496	9544

Prevalence odds ratio 0.37	Prevalence odds ratio 0.44	17% difference

Table 4.21 orders the stratified prevalence odds ratios from largest effect modifier to smallest effect modifier. The table also lists whether the effect modifier exhibits synergy or antagonism with bicycle commuting. Effect modifiers that exhibit synergy work together with bicycle commuting to decrease the odds of CAD prevalence, while effect modifiers that exhibit antagonism work against bicycle commuting to increase the odds of CAD prevalence (Aschengrau & Seage, 2014b).

Table 4.21: Effect modifiers of the association between census tract CAD prevalence and bicycle commuting rate

Effect Modifier	**Adjusted Prevalence odds ratios**		**Classification**
	Below median	Above median	
Obesity Prevalence (%)	0.43	0.79	antagonism
Diabetes Prevalence (%)	0.48	0.73	antagonism
Smoking Prevalence (%)	0.42	0.66	antagonism
Manual Labor Job (%)	0.32	0.60	antagonism
Latinx (%)	0.27	0.59	antagonism
Healthy Food Access (%)	0.57	0.29	synergy
No College Education (%)	0.43	0.62	antagonism
Median Income ($1K)	0.50	0.41	synergy
Older than 45 (%)	0.38	0.47	antagonism
Married (%)	0.47	0.33	synergy
White (%)	0.45	0.39	synergy
Hypertension Prevalence (%)	0.37	0.44	antagonism

Obesity, diabetes, and smoking prevalence were the greatest effect modifiers of the association between bicycle commuting and CAD prevalence. In general, the chronic disease effect modifiers (obesity, diabetes, smoking, and hypertension) exhibited antagonism which means they worked against bicycle commuting to increase the odds of CAD prevalence.

Likewise, the effect modifiers that were indicators of low socioeconomic status (manual labor occupation, Latinx, and no college education) exhibited antagonism. In contrast, the effect modifiers that were indicators of high socioeconomic status (income, married, and white) exhibited synergism which means they worked together with bicycle commuting to lower the odds of disease prevalence.

Latent Class Cluster Analysis

The stratified analysis revealed 12 effect modifiers: age, education, marital status, income, occupation type, race, ethnicity, food access, smoking prevalence, obesity prevalence, diabetes prevalence, and hypertension prevalence. When I performed a latent class cluster analysis from the 12 effect modifiers, the census tracts were split into five clusters. Appendix B provides a multi-dimensional scaling plot that demonstrates the distance between the clusters. Table 4.22 lists the average characteristics of each cluster.

Table 4.22: Average census tract demographics, socioeconomics, and health risk factors by LCA cluster

Census Tract Characteristics	**Cluster 1** (n=4146)	**Cluster 2** (n=2736)	**Cluster 3** (n=2015)	**Cluster 4** (n=1735)	**Cluster 5** (n=1690)
Bicycle Commuting (%)	1.8	0.7	0.6	1.0	1.2
Sex (% Male)	49.2	49.8	46.3	49.8	48.8
Age (% > 45)	38.8	32.2	37.4	35.2	38.8
Married (%)	5.5	6.6	8.0	5.5	8.6
White (%)	74.4	55.5	14.2	69.9	35.5
Black (%)	6.9	18.3	74.9	15.0	23.9
Latinx (%)	14.0	54.7	10.1	25.8	34.2
Median Income ($1K)	89.4	36.1	30.8	55.5	57.1
No College Education (%)	23.7	61.1	55.4	40.2	42.8
Manual Labor Job (%)	4.6	15.1	8.3	9.5	9.6
Healthy Food Access (%)	40.8	45.2	42.7	62.0	15.2
Smoking Prevalence (%)	12.3	22.5	25.9	18.2	15.7
Obesity Prevalence (%)	22.5	36.2	41.7	30.3	25.6

Diabetes Prevalence (%)	7.2	13.7	16.9	9.5	11.5
Hypertension Prevalence(%)	24.2	32.9	42.7	28.9	29.1

Figure 4.5 provides a visualization of cluster characteristics. In Figure 4.5, there is a sliding bar for each effect modifier. The position of the marker on the sliding bars represents the percentage of census tracts (0-100%), within each cluster, that has above median prevalence of the corresponding characteristic. Appendix B provides a table with the percentages represented in Figure 4.5. Effect modifiers are organized in Figure 4.5 by indicators of socioeconomic status, race/ethnicity, and CAD health status.

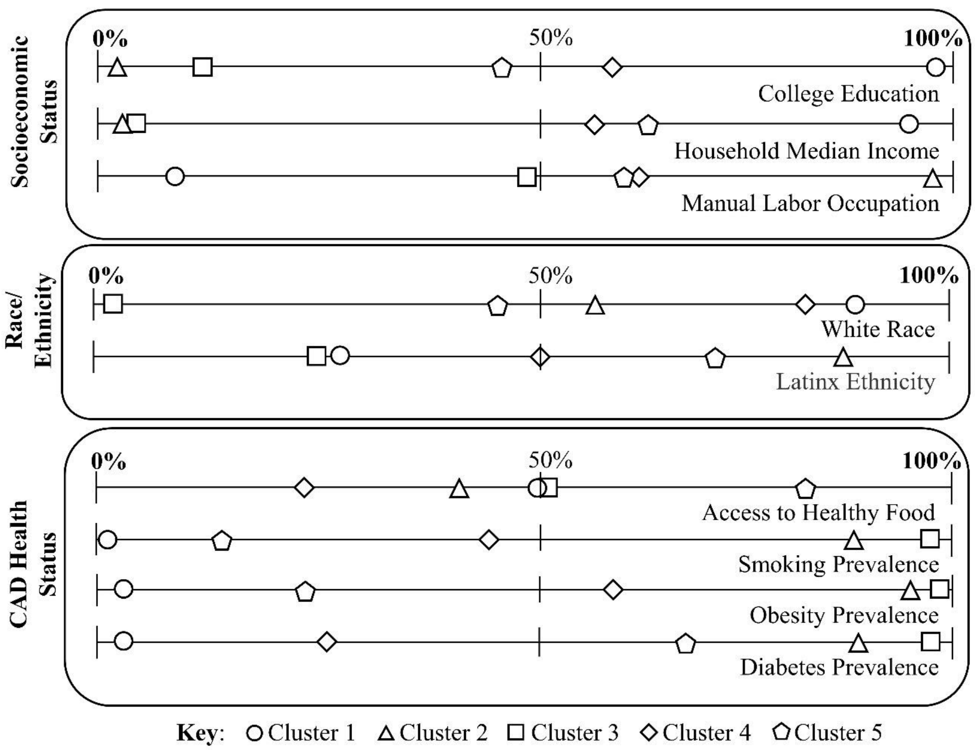


Figure 4.5: Percent of census tracts with above median effect modifier prevalence by LCA cluster

On average, the census tracts in cluster 1 (n=4146) had populations that were college educated ($\bar{x}$ = 76%), white ($\bar{x}$ = 74%), and high income (*Med* = $89,000); the average prevalence of smoking ($\bar{x}$ = 12%), obesity ($\bar{x}$ = 22%), and diabetes ($\bar{x}$ = 7%) were lower than other clusters.

Overall, cluster 1 includes census tracts with high SES, above average white race, and the best health indicators.

On average, the census tracts in cluster 2 (n=2736) had populations that were not college educated ($\bar{x}$ = 61%), Latinx ($\bar{x}$ = 55%), and low-income (*Med* = $36,000) with an above average percentage of manual labor jobs ($\bar{x}$ = 15%); the prevalence of smoking ($\bar{x}$ = 23%), obesity ($\bar{x}$ = 36%), and diabetes ($\bar{x}$ = 14%) were higher than average. Overall, cluster 2 includes census tracts with low SES, above average Latinx ethnicity, and below average health indicators.

On average, the census tracts in cluster 3 (n=2015) had populations that were not college educated ($\bar{x}$ = 55%), black ($\bar{x}$ = 75%), and low-income (*Med* = $31,000); the average prevalence of smoking ($\bar{x}$ = 26%), obesity ($\bar{x}$ = 42%), and diabetes ($\bar{x}$ = 17%) were the highest compared to other clusters. Overall, cluster 3 includes census tracts with low SES, above average black race, and the worst health indicators.

The average characteristics of cluster 4 were less extreme than clusters 1, 2, or 3. On average, the census tracts in cluster 4 (n=1735) had populations that were less college educated ($\bar{x}$ = 60%), less white ($\bar{x}$ = 70%), and lower income (*Med*= $55,000) than cluster 1; but more educated, more white, and higher income than clusters 2 and 3. The prevalence of obesity ($\bar{x}$ = 30%), smoking ($\bar{x}$ = 18%), and diabetes ($\bar{x}$ = 10%) were higher than cluster 1, but lower than clusters 2 and 3. Overall, cluster 4 includes census tracts with middle SES, above average white race, and average health indicators.

Like cluster 4, the average characteristics of cluster 5 (n=1690) were less extreme than those in clusters 1, 2, or 3. On average, the census tracts in cluster 5 had populations that were less college educated ($\bar{x}$ = 57%), less white ($\bar{x}$ = 36%), lower income (*Med*= $57,000) and more Latinx ($\bar{x}$ = 34%) than cluster 1, but less Latinx than cluster 2. The prevalence of obesity ($\bar{x}$ =

26%) and smoking ($\bar{x}$ = 16%) were higher than cluster 1, but lower than clusters 2, 3, and 4. The prevalence of diabetes ($\bar{x}$ = 12%) was higher than cluster 1, but lower than clusters 2 and 3. Overall, cluster 5 includes census tracts with middle SES and average health indicators.

Age was an effect modifier in addition to those presented in Figure 4.5. However, age did not correlate with the other effect modifiers in the LCA. For this reason, the five clusters were divided into two subgroups based on age. The five clusters were split into two subgroups by: census tracts where the percent of population older than 45 years was above the median, and census tracts where the percent of population older than 45 years was below the median. The age groups were split at 45 years in accordance with American College of Sports Medicine CAD risk factor thresholds where individuals older than 45 are at higher risk of developing CAD (ACSM, 2014). The five clusters and two subgroups within each cluster resulted in ten census tract groups. Appendix B provides a table with the average characteristics of each census tract group.

Measures of Association

Prevalence odds ratios measuring the association between bicycle commuting and CAD prevalence are provided in Table 4.23 by cluster and age group. The prevalence odds ratio (OR) represents the odds that a census tract will have CAD prevalence greater than 5.5% when bicycle commuting is not present.

Table 4.23: Prevalence odds ratios representing the odds that a census tract will have CAD prevalence greater than 5.5% when bicycle commuting rate is zero

ID #	Majority Characteristics: *Socioeconomic Status*	*Race/ Ethnicity*	*Health Status*	*Age Group*	Sample Size *(n)*	Odds Ratio *(95% CI)*
1	High (Income, *M*=89K)	White ($\bar{x}$ =74%)	above average	>45	2497	**1.24** (1.17, 1.31)
				<45	1649	**1.23** (1.07, 1.47)
2	Low (Income, *M*=36K)	Latinx ($\bar{x}$ =55%)	below average	>45	749	0.96 (0.71, 1.18)
				<45	1987	1.04 (0.97, 1.11)
3	Low (Income, *M*=31K)	Black ($\bar{x}$ =75%)	below average	>45	1064	1.01 (0.81, 1.17)
				<45	951	1.00 (0.90, 1.11)
4	Middle (Income, *M*=56K)	White ($\bar{x}$ =70%)	average	>45	804	1.02 (0.95, 1.10)
				<45	931	1.02 (0.95, 1.11)
5	Middle (Income, *M*=57K)	White, $\bar{x}$ =36% Latinx, $\bar{x}$ =34% Black, $\bar{x}$ =24%	average	>45	1023	0.98 (0.90, 1.06)
				<45	667	1.08 (1.00, 1.18)

Cluster 1 was the only cluster where the prevalence odds ratio was statistically significant. Cluster 1 includes census tracts with high SES, an above average percentage of white population, and above average health status. These results support the results from the stratified analysis. White race and high-income worked synergistically with bicycle commuting to reduce the odds of CAD prevalence while poor health worked antagonistically to increase the odds of CAD prevalence.

The strength of association between bicycle commuting rate and CAD prevalence varied by SES and health status. The association was only statistically significant in census tracts where: 1) SES was high, and 2) health indicators (e.g., obesity, smoking, diabetes) were above average. The census tracts with high SES and above average health indicators also had an above average white population. In other words, the census tract populations with statistically significant associations between bicycle commuting and CAD were generally affluent, healthy, and predominately white.

These results are consistent with European studies. In the United Kingdom (UK), bicycle commuting was associated with a lower risk of CAD (Celis-Morales et al., 2017), and the majority of bicycle commuters in the study sample were white (96%) with higher income, lower obesity, and lower diabetes than other commuter types (Celis-Morales et al., 2017). More specifically, bicycle commuters had lower prevalence of obesity and diabetes when compared to walking commuters and non-active commuters (Celis-Morales et al., 2017). Bicycle commuters also had the lowest risk of CAD mortality and CAD incidence (Celis-Morales et al., 2017). The results of Celis-Morales et al. (2017) are similar to my study because of the associations between bicycle commuting, income, health risk, and race. However, Celis-Morales et al. (2017) conducted a cohort study of individuals rather than an ecologic study of populations. Therefore, Celis-Morales et al. (2017) study provided evidence of causality, whereas the design of my study does not. A Swedish case-control study also provided evidence of causality because bicycle commuters had a lower risk of heart attack than non-active commuters (Wennberg et al., 2006). My study did not assess heart attack because it is a measure of CAD incidence not CAD prevalence, but it is important to note that European studies have found associations between bicycle commuting and both incidence and prevalence.

Another study, conducted in Norway, examined the association between bicycle commuting and CAD risk factors within a multi-ethnic, low SES population (Riiser et al., 2018).

This cross-sectional study differed from Celis-Morales et al. (2017) and Wennberg et al. (2006) because a low SES, multi-ethnic population was intentionally selected for the study sample. The Norwegian study found that bicycle commuters in the low SES population had lower odds of diabetes and lower risk factors for CAD than those who did not actively commute (Riiser et al., 2018). I selected my sample based on the most populous U.S. cities, not based on SES or ethnicity, so our results did not indicate a statistically significant association within low SES populations. If I had designed the study to evaluate associations within low SES populations, I may have seen different results.

In my study, the average percentage of bicycle commuters within the low SES cluster ($\bar{x}$ =0.6%) was half the average of the overall sample ($\bar{x}$ =1.2%), and one-third the average of the high SES cluster ($\bar{x}$ =1.8%). In other words, there was a significantly lower number of bicycle commuters in the low SES populations than in the high SES population. A study conducted across 22 U.S. cities also found that the bicycle commuting rate was low within "disadvantaged" census tracts (Braun, 2021). Because high income and white race were synergistic effect modifiers in my study, evaluating the association between bicycle commuting and CAD prevalence at a population level likely masked any positive associations within the low SES census tracts.

In my study, bicycle commuting rates were lower among low SES populations than among high SES populations. Furthermore, obesity and diabetes prevalence were above average in the low SES populations. Future research would benefit from examining the association between U.S. bicycle commuting rates and CAD, specifically in low SES populations. Likewise, future researchers should study the factors associated with bicycle commuting in low SES populations. A recent study suggested that bicycle infrastructure was not as strongly associated with bicycle commuting in “disadvantaged” census tracts as it was within “advantaged” census tracts (Braun, 2021). Non-infrastructure factors like equity goals in planning, planning representation, advocacy representation, public involvement and social norms may have a significant effect on bicycle commuting in low SES populations and should be evaluated in future research (Braun et al., 2019).

CHAPTER FIVE: CONCLUSIONS, IMPLICATIONS & CONTRIBUTIONS

Conclusions

This research aimed to understand the infrastructure and policy factors associated with bicycle commuting rates in U.S. cities and evaluated the association between bicycle commuting rates and heart disease prevalence by census tract. I would recommend five actions based on the results of this study: 1) decrease cyclist fatalities, 2) increase bike network density, 3) improve bike access to public transit, 4) investigate barriers to bike commuting in low-income communities, and 5) build partnerships and capacity for active transportation in low-income communities. Appendix C provides an infographic that summarizes these five actions. I recommend these actions based on a culmination of results from research question 1, 2, and 3. In the remainder of this section, I have summarized the results of each research question and explain how the results from each question support the five recommended actions.

Research Question 1

The first research question asked, "What variables should be used to quantitatively measure bicycle infrastructure when evaluating the association with U.S. bicycle commuting rates?" Based on the findings of this study, five variables can be used as independent measures of bicycle infrastructure: *access to public transit, bike lane supply, bike lane quality, bike network density, and bike path supply*. Table 5.1 summarizes the units of each variable and the corresponding data sources. More importantly, network measures should be prioritized when using infrastructure to evaluate bicycle commuting rates. Three network measures had the strongest correlations with bicycle commuting rates: *access to public transit, bike network* *quality, and bike network density.* Two of the recommended actions are supported by these findings- increase bike network density and improve bike access to public transit.

Table 5.1: Five variables used to quantitatively measure bicycle infrastructure when evaluating the association with U.S. bicycle commuting rates

Variable Name	Units	Source
Bike lane supply	Miles of bike lanes in city per 100,000 population	League of American Bicyclists (LAB) Benchmarking Report (K. McLeod et al., 2019)
Bike path supply	Miles of paved multi-use paths in city per 100,000 population	
Bike lane quality	Percent of bike lanes that are buffered or protected	
Bike network density	Miles of low stress roadway for cyclists per city area	Bicycle Network Analysis (Bicycle Network Analysis [Data], 2020)
Access to public transit	Score of major transit hubs accessible on the low stress bike network	

Research Question 2

The second research question asked, “How much of the variation in U.S. bicycle commuting rates can be explained by bicycle infrastructure and policy measures?” Research question two also included a sub-question asking, “What is the relative significance of bicycle infrastructure versus policy measures when evaluating their association with U.S. bicycle commuting rates?” Through a series of stepwise regression models and a k-fold cross validation procedure, I found that three variables explained 70% of the variation in bicycle commuting rates. The three variables included: *cyclist safety, city employees working on bicycle projects,* *and access to public transit*. Table 5.2 summarizes the units of the three variables and their corresponding data sources.

When considering the relative significance of infrastructure and policy-related variables, neither stood out as more statistically significant. On the contrary, the results suggest that a combination of infrastructure and policy factors contribute to U.S. bicycle commuting rates. For example, of the three variables explaining 70% of the variation in bicycle commuting rate, *city employees working on bicycle projects* was considered a policy-relevant variable, *access to public transit* was considered an infrastructure variable, and *cyclist safety* was considered a covariate. *Cyclist safety* had the strongest association with bicycle commuting rate (β= -0.80), followed by *city employees working on bicycle projects* (β= 0.66), and finally *access to public transit* (β= 0.28). Three of the recommended actions are supported by these findings- decrease cyclist fatalities, increase bike access to public transit, and build partnerships.

Table 5.2: Three variables explained 70% of the variation in U.S. bicycle commuting rates

Variable Name	**Units**	**Source**
Cyclist safety	State data: three-year average number of bicyclist fatalities per 10,000 bicycle commuters	USDOT Fatality Analysis Reporting System 2015-2017, averages (*Fatality Analysis Reporting System [CSV Data File]*, 2017)
City employees working on bicycle projects	Number of full-time city employees working on bike or pedestrian issues per 100,000 population	LAB Benchmarking Report (K. McLeod et al., 2019)
Access to public transit	Score of major transit hubs accessible on the low stress bike network	Bicycle Network Analysis (Bicycle Network Analysis [Data], 2020)

Research Question 3

The third research question asked, “What is the significance of association between U.S. bicycle commuting rates and heart disease prevalence, at the census tract level, when accounting for heart disease risk factors?” Research question three also included a sub-question asking, “How much do obesity, diabetes, and hypertension modify the association?” I calculated the crude prevalence odds ratio and adjusted prevalence odds ratios of the association between bicycle commuting rate and CAD prevalence to answer these questions. Prevalence odds ratios were calculated for census tracts where CAD prevalence was greater than 5.5% versus census tracts where CAD prevalence was less than 5.5%. The benchmark of 5.5% was chosen because the median CAD prevalence across all census tracts (n= 12,322) was 5.5%. Bicycle commuting rate was considered the risk factor of interest where census tracts with a bicycle commuting rate of zero percent were compared to census tracts with a bicycle commuting rate greater than zero percent. In other words, the prevalence odds ratios calculated in this study represent the odds that a census tract will have CAD prevalence greater than 5.5% when the bicycle commuting rate is zero percent.

There were 12 effect modifiers of the association between bicycle commuting and CAD prevalence. Obesity and diabetes prevalence were the strongest effect modifiers of the association. The percent difference between the stratified prevalence odds ratios for obesity and diabetes prevalence were 58% and 42%, respectively. In census tracts with above average obesity prevalence (> 29%), the odds of living in a census tract with above average CAD prevalence (> 5.5%) will decrease by 21% (OR= 0.79) when the bicycle commuting rate is greater than zero. In census tracts with above average diabetes prevalence (> 11%), the odds of living in a census tract with above average CAD prevalence (> 5.5%) will decrease by 27% (OR= 0.73) when the bicycle commuting rate is greater than zero. In other words, both obesity and diabetes work against bicycle commuting as effect modifiers and ultimately increase a census tract's odds of CAD prevalence. Overall, the statistical significance of the association between bicycle commuting rates and CAD prevalence (OR= 0.40) is reduced in census tracts with high prevalence of obesity (OR= 0.79) and diabetes (OR= 0.73).

In this study, the crude prevalence odds ratio was calculated as 0.40. The crude prevalence odds ratio can be interpreted as, the odds of living in a census tract with above average CAD prevalence (> 5.5%) will decrease by 60% (OR= 0.40) when the bicycle commuting rate is greater than zero. The adjusted prevalence odds ratio accounted for heart disease risk factors like socioeconomic status, race, and health status. In this study, the adjusted prevalence odds ratio was calculated as 0.80 within predominately affluent, healthy, and white census tracts. The adjusted prevalence odds ratio can be interpreted as, the odds of living in a census tract with above average CAD prevalence (> 5.5%) will decrease by 20% (AOR=0.80) when the bicycle commuting rate is greater than zero. The adjusted prevalence odds ratio was only statistically significant in census tracts with high socioeconomic status, above average health, and predominately white population. Furthermore, census tracts with high socioeconomic status had three times the number of bike commuters as census tracts with low socioeconomic status. Two recommended actions are supported by these findings- investigate barriers to bike commuting in low-income communities and build partnerships and capacity for active transportation in low-income communities.

Implications for Practitioners

Before discussing contributions to research and future research recommendations, it is important to summarize the broader implications of this research. I summarize implications with an emphasis on how practitioners might be able to use the results of this study in their work. Practitioners in this circumstance refer to transportation professionals like urban planners, transportation engineers, and bicycle and pedestrian coordinators, or public health professionals like health educators, program coordinators, or social epidemiologists.

Research Question 1

The first research question asked how to measure bicycle infrastructure when evaluating U.S. bicycle commuting rates. The findings and conclusions imply that bike network scores provided by PeopleForBikes are valuable measures of bicycle infrastructure. The correlation coefficients between PeopleForBikes bike network scores and bicycle commuting rates were greater than 0.40. In other words, practitioners looking to improve bicycle infrastructure in their municipality or town would benefit from referencing the bicycle network maps and bicycle network analysis tool provided PeopleForBikes. The tool can be accessed at https://bna.peopleforbikes.org/#/places/.

PeopleForBikes Bicycle Network Analysis tool provides bike network measures for hundreds of U.S. cities. Practitioners can use the tool to examine the bicycle networks in their town or in towns similar to their own. According to the findings of this study, the BNA score and transit score were the best infrastructure measures provided by PeopleForBikes. In other words, the correlations between BNA score, transit score, and bicycle commuting rate were more statistically significant than other infrastructure variables evaluated in this study. Practitioners interested in improving bicycle infrastructure in their municipality can use the tool to compare their municipality's BNA score and transit score with scores in similar cities or towns. The BNA tool also provides a map of bicycle routes which practitioners could use to strategize locations for network improvements. Strategizing locations to improve bike network access to public transit are especially important.

Research Question 2

The second research question asked how much of the variation in U.S. bicycle commuting rates could be explained by infrastructure or policy measures. The research question also evaluated the relative significance of infrastructure and policy measures. The findings and conclusions imply that three measures attributed to more than half of the variation in U.S. bicycle commuting rates. The three measures represent: 1) bicycle network connections to public transit, 2) the average number of cyclist fatalities, and 3) the number of FTE city employees spending at least one-tenth of their work hours on bicycle projects. In other words, as bicycle commuting rates increased, connections to public transit increased, the number of city employees working on bicycle projects increased, and the number of cyclist fatalities decreased. Practitioners aiming to improve bicycle-friendliness would benefit from tracking public transit connections, cyclist safety, and employee hours spent on bicycle projects.

There are a few ways practitioners may be able to incorporate these measures into their work. Strategies might include prioritizing projects that enhance cyclist safety. For example, to enhance cyclist safety, municipalities might assess the locations of cyclist crashes and prioritize roadway improvement projects at those locations. A recently introduced bill called the SAFE Streets Act (H.R. 508) would support these actions. The SAFE Streets Act was introduced in the U.S. House of Representatives in January of 2021 (SAFE Streets Act, 2021). The bill would require municipalities to conduct "vulnerable road user" safety assessments that evaluate locations and corridors where serious cyclist injuries and fatalities occur. The bill would also require municipalities to propose strategies reducing safety risks in those corridors. The implications of the H.R. 508 bill support the findings of my study which concluded that, out of 14 potential variables, cyclist safety had the most significant association with bicycle commuting rates.

Additionally, a practitioner looking to improve bicycle-friendliness should prioritize transportation projects that improve connections between public transit and bicycle networks. For example, the bicycle commuting rate within a municipality will likely benefit more from proposing a bicycle lane connection at a public transit corridor than a bicycle lane for the sole purpose of increasing bicycle lane mileage. In other words, strategic decision-making is crucial when planning projects to increase bicycle commuting rates. Practitioners can help advocate for these strategic decisions by prioritizing projects that improve multi-modal transit connections.

The Transportation Alternatives Enhancement Act (H.R. 463) was introduced in the U.S. House of Representatives in January of 2021 and would provide local practitioners with more decision-making power over local transportation priorities, like bicycle infrastructure (Transportation Alternatives Enhancements Act, 2021). The bill supports local decision making by preventing states from transferring Transportation Alternatives (TA) funds without a competitive application process, it also requires states to demonstrate there were no suitable applications before transferring funds. The TA bill would bolster employee hours spent on bicycle projects because it would require states to use a portion of TA funds to provide technical and engineering assistance to local municipalities. The implications of the H.R. 463 bill support the findings of my study because it would strengthen practitioners' decision-making power at a local level and would potentially bolster the number of employee hours spent on bicycle projects.

Overall, bicycle commuting rates in large U.S. cities will benefit most if TA funds are prioritized for: 1) projects that improve cyclists' safety and access to public transit, and for 2) increasing the number of hours spent working on bicycle projects at the local level.

Research Question 3

The third research question assessed the association between bicycle commuting rates and heart disease prevalence at the census tract level. The findings imply that bicycle commuting was significantly associated with heart disease in census tracts where the population was affluent, healthy, and predominately white. Obesity prevalence and diabetes prevalence were particularly relevant when assessing the health of these populations. I also found that bicycle commuting rates in high SES census tracts were three times greater than in low SES census tracts. Low bicycle commuting rates likely contributed to insignificant associations between bicycle commuting and heart disease prevalence in tracts with low socioeconomic status. Practitioners focused on the social determinants of health should focus more of their efforts on active transportation in low SES populations.

In particular, practitioners can make a difference by working to increase cycling in low SES populations. Barriers to cycling in low SES populations are likely different from the barriers in high SES populations. A first step would be working to understand those barriers. Some of those barriers may be related to access to resources, financial cost of a bike, and underrepresentation during planning decisions. Practitioners engaged in community-based programming are needed to help investigate these barriers, rather than external researchers, because they have already developed trusting relationships within their communities. A few ways practitioners may be able to help decrease barriers to cycling in low SES populations include: partnering with local bicycle organizations to improve service to low SES communities, organizing outreach within low SES communities to better understand their transportation assets and needs, and advocating for the public health benefits of cycling within low SES communities.

On the other hand, it is difficult to advocate for the health benefits of cycling within low SES populations without access to data on cyclists in these populations. The rarity of cyclists within low SES communities poses a challenge for practitioners looking to demonstrate positive health outcomes. From this perspective, community-based practitioners would benefit from federal government support. The CDC and USDA manage national chronic disease programs like the Diabetes Prevention Program and the Supplemental Nutrition Assistance Program (SNAP). These programs have protocols that are mandated at a federal level, yet they are implemented at a local level. For example, all Diabetes Prevention Programs are required to report the weight of their participants while SNAP-Ed programs require data collection on program participants. Mandating the collection of transportation mode through federal health programs reporting mechanisms may be one way to build a dataset on the association between cycling and health outcomes in low SES communities.

Contributions to Research

This study resulted in a number of distinct research contributions including:

1) An assessment of eight, publicly accessible, bicycle infrastructure measures and their association with U.S. bicycle commuting rates,
2) A relative comparison of fourteen variables effect on bicycle commuting rates in 50 U.S. cities,
3) An evaluation of the ecologic association between U.S. bicycle commuting rate and heart disease prevalence by socioeconomic status and population health risk.

These contributions have filled gaps in knowledge on the infrastructure and policy factors quantitatively associated with U.S. bicycle commuting rates and have created knowledge on the ecologic association between bicycle commuting rate and heart disease risk factors.

To my knowledge this study was the first to assess the association between PeopleForBikes network measures and bicycle commuting rates. The PeopleForBikes network measures were more advanced than prior measures of bike networks because the measures accounted for accessibility to destinations like jobs and transit. This study was also the first to include measures of *bike lane quality, complete streets policy, city employees working on bicycle projects,* and *access to public transit* within an aggregate model of U.S. bicycle commuting rates. The use of methods like principal components analysis made it possible for multiple infrastructure measures to be included in a single model. An assessment of infrastructure measures, through my study, fills gaps in the knowledge base on aggregate models of U.S. bicycle commuting rates.

Furthermore, the ecologic association between bicycle commuting rate and heart disease prevalence had not been evaluated across more than 12,000 census tracts prior to this study. Geographic population studies, while imperfect due to ecological fallacy, are valuable for hypothesis generation. The ecologic analysis I conducted was a first step towards exploring the association between U.S. bicycle commuting rates and heart disease at a national scale. The ecologic analysis sets a precedent for future longitudinal studies and clarifies a need to explore the association between bicycle commuting and heart disease within low-income communities.

Limitations and Research Recommendations

The quantitative, ecologic analysis presented through this study was limited by data availability. Data availability limited the sample size, the study design, and the measures used for each variable. The sample size was restricted to 50 cities based on data availability from the League of American Bicyclists; data availability dictated an ecologic study design because bicycle commuting and heart disease data were merged from two separate sources; and data availability influenced the measures used for infrastructure, policy, bicycle commuting, and heart disease variables. Curation of data from nine secondary sources was a limitation of this study, but it was also a strength because the results of this study provide a baseline for future researchers looking to examine the relationship between cycling and chronic disease outcomes.

Future researchers could help improve data availability by collecting bicycle infrastructure data beyond the 50 cities assessed in this study. Integrated bike maps containing network measures, cyclist fatalities, bicycle partners (i.e. bike shops, nonprofits etc.), health outcomes, and socioeconomic status are needed, all in one spot, to improve data accessibility and promote community involvement in bicycle advocacy. Additionally, cross-sectional or cohort data on transportation mode and health outcomes would allow for more advanced study designs in the future. A particularly salient outcome of this study is the need to research barriers to cycling in low SES populations and corresponding health outcomes. The ecologic design of this study likely masked any positive associations between cycling and heart disease in low SES populations, so future researchers should design studies intended for low SES communities. Researchers should aim to partner with community-based practitioners to complete this work. The results of this study are purely correlational, but conducting community-based research would present an opportunity to examine causal relationships over a longer period of time.

REFERENCES

500 Cities Project Data [CSV Data file]. (2018). Centers for Disease Control and Prevention. https://www.cityhealthdashboard.com

About Us. (2019). Active Living Research. https://www.activelivingresearch.org/aboutus

Adam, L., Jones, T., & te Brömmelstroet, M. (2020). Planning for cycling in the dispersed city: Establishing a hierarchy of effectiveness of municipal cycling policies. *Transportation*, *47*(2), 503–527. https://doi.org/10.1007/s11116-018-9878-3

Akar, G., & Clifton, K. J. (2009). Influence of Individual Perceptions and Bicycle Infrastructure on Decision to Bike. *Transportation Research Record*, *2140*(1), 165–172. https://doi.org/10.3141/2140-18

American Community Survey 5-Year Estimates [CSV Data file]. (2017). U.S. Census Bureau. https://factfinder.census.gov

Arellana, J., Saltarín, M., Larrañaga, A. M., González, V. I., & Henao, C. A. (2020). Developing an urban bikeability index for different types of cyclists as a tool to prioritise bicycle infrastructure investments. *Transportation Research Part A: Policy and Practice*, *139*, 310–334. https://doi.org/10.1016/j.tra.2020.07.010

Aschengrau, A., & Seage, G. (2014a). Confounding. In *Epidemiology in Public Health* (3rd ed., pp. 293–311). Jones & Bartlett Learning, LLC.

Aschengrau, A., & Seage, G. (2014b). Effect Measure Modification. In *Epidemiology in Public Health* (3rd ed., pp. 349–360). Jones & Bartlett Learning, LLC.

Aschengrau, A., & Seage, G. (2014c). Overview of Epidemiologic Study Designs. In *Epidemiology in Public Health* (3rd ed., pp. 143–171). Jones & Bartlett Learning, LLC.

Assunçao-Denis, M.-È., & Tomalty, R. (2019). Increasing cycling for transportation in Canadian communities: Understanding what works. *Transportation Research Part A: Policy and Practice*, *123*, 288–304. https://doi.org/10.1016/j.tra.2018.11.010

Atherton, A., Azeez, N., Lee Davia, S., Doyle, S., Hanzlik, M., Rosenberg, G., & Zaccaro, H. (2018). *The Best Complete Streets Initiatives of 2017*. Smart Growth America.

Aytur, S., Rodriguez, D., Kerr, Z., Ji, K., & Evenson, K. (2013). Spatial and Temporal Patterns of North Carolina Pedestrian and Bicycle Plans. *Journal of Public Health Management and Practice*, *19*, S83–S88. https://doi.org/10.1097/PHH.0b013e31828404a0

Barengo, N. C., Hu, G., Lakka, T. A., Pekkarinen, H., Nissinen, A., & Tuomilehto, J. (2004). Low physical activity as a predictor for total and cardiovascular disease mortality in middle-aged men and women in Finland. *European Heart Journal*, *25*(24), 2204–2211. https://doi.org/10.1016/j.ehj.2004.10.009

Bassett, D. R., Pucher, J., Buehler, R., Thompson, D. L., & Crouter, S. E. (2008). Walking, Cycling, and Obesity Rates in Europe, North America, and Australia. *Journal of Physical Activity and Health*, *5*(6), 795–814. https://doi.org/10.1123/jpah.5.6.795

Beenackers, M. A., Foster, S., Kamphuis, C. B. M., Titze, S., Divitini, M., Knuiman, M., van Lenthe, F. J., & Giles-Corti, B. (2012). Taking up cycling after residential relocation: Built environment factors. *American Journal of Preventive Medicine*, *42*(6), 610–615. https://doi.org/10.1016/j.amepre.2012.02.021

Benjamin, E., Muntner Paul, Alonso Alvaro, Bittencourt Marcio S., Callaway Clifton W., Carson April P., Chamberlain Alanna M., Chang Alexander R., Cheng Susan, Das Sandeep R., Delling Francesca N., Djousse Luc, Elkind Mitchell S.V., Ferguson Jane F., Fornage Myriam, Jordan Lori Chaffin, Khan Sadiya S., Kissela Brett M., Knutson Kristen L., …

null null. (2019). Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association. *Circulation*, *139*(10), e56–e528. https://doi.org/10.1161/CIR.0000000000000659

Bicycle Network Analysis [data]. (2020). PeopleForBikes. https://bna.peopleforbikes.org/#/places/

Biton, A., Daddio, D., & Andrew, J. (2014). *Statewide Pedestrian and Bicycle Planning Handbook* (p. 94). FHWA.

Braun, L. M. (2021). Disparities in Bicycle Commuting: Could Bike Lane Investment Widen the Gap? *Journal of Planning Education and Research*, 0739456X21993905. https://doi.org/10.1177/0739456X21993905

Braun, L. M., Rodriguez, D. A., Cole-Hunter, T., Ambros, A., Donaire-Gonzalez, D., Jerrett, M., Mendez, M. A., Nieuwenhuijsen, M. J., & de Nazelle, A. (2016). Short-term planning and policy interventions to promote cycling in urban centers: Findings from a commute mode choice analysis in Barcelona, Spain. *Transportation Research Part A: Policy and Practice*, *89*, 164–183. https://doi.org/10.1016/j.tra.2016.05.007

Braun, L. M., Rodriguez, D. A., & Gordon-Larsen, P. (2019). Social (in)equity in access to cycling infrastructure: Cross-sectional associations between bike lanes and area-level sociodemographic characteristics in 22 large U.S. cities. *Journal of Transport Geography*, *80*. https://doi.org/10.1016/j.jtrangeo.2019.102544

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, *46*(10), 1730–1740. https://doi.org/10.1016/j.tra.2012.07.005

Bronfenbrenner, U. (1981). *The Ecology of Human Development*.

Bronfenbrenner, U. (1992). Ecological systems theory. In *Six theories of child development: Revised formulations and current issues* (pp. 187–249). Jessica Kingsley Publishers.

Brunsing, J. (1997). Public transport and cycling: Experience of modal integration in Germany. *Wiley (John) and Sons, Limited*. https://trid.trb.org/view/501897

Buehler, R., & Dill, J. (2016). Bikeway Networks: A Review of Effects on Cycling. *Transport Reviews*, *36*(1), 9–27. https://doi.org/10.1080/01441647.2015.1069908

Buehler, R., & Pucher, J. (2012). Cycling to Work in 90 Large American Cities: New Evidence on the Role of Bike Paths and Lanes. *Transportation*, *39*(2), 409–432.

Buehler, R., Pucher, J., & Bauman, A. (2020). Physical activity from walking and cycling for daily travel in the United States, 2001–2017: Demographic, socioeconomic, and geographic variation. *Journal of Transport & Health*, *16*. https://doi.org/10.1016/j.jth.2019.100811

Cassel, J. C. (1971). Summary of Major Findings of the Evans County Cardiovascular Studies. *Archives of Internal Medicine*, *128*(6), 887–889. https://doi.org/10.1001/archinte.1971.00310240041003

Castro, C. M., King, A. C., & Brassington, G. S. (2001). Telephone versus mail interventions for maintenance of physical activity in older adults. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, *20*(6), 438–444.

Celis-Morales, C. A., Lyall, D. M., Welsh, P., Anderson, J., Steell, L., Guo, Y., Maldonado, R., Mackay, D. F., Pell, J. P., Sattar, N., & Gill, J. M. R. (2017). Association between active commuting and incident cardiovascular disease, cancer, and mortality: Prospective cohort study. *BMJ*, *357*. https://doi.org/10.1136/bmj.j1456

Centers for Disease Control and Prevention. (2014). *Behavioral risk factor surveillance system overview: BRFSS 2013*. http://www.cdc.gov/brfss/annual_data/2013/pdf/overview_2013.pdf

Centers for Disease Control and Prevention. (2018). *About the project*. https://www.cdc.gov/500cities/about.htm

Cervero, R., Sarmiento, O. L., Jacoby, E., Gomez, L. F., & Neiman, A. (2009). Influences of Built Environments on Walking and Cycling: Lessons from Bogotá. *International Journal of Sustainable Transportation*, *3*(4), 203–226. https://doi.org/10.1080/15568310802178314

Chen, C., Wang, H., Roll, J., Nordback, K., & Wang, Y. (2020). Using bicycle app data to develop Safety Performance Functions (SPFs) for bicyclists at intersections: A generic framework. *Transportation Research Part A: Policy and Practice*, *132*, 1034–1052. https://doi.org/10.1016/j.tra.2019.12.034

Clifton, K. J., & Dill, J. (2005). Women's Travel Behavior and Land Use: Will New Styles of Neighborhoods Lead to More Women Walking? *Transportation Research Board Conference Proceedings*. Conference on Research on Women's Issues in Transportation. https://trid.trb.org/view/773072

Cole-Hunter, T., Donaire-Gonzalez, D., Curto, A., Ambros, A., Valentin, A., Garcia-Aymerich, J., Martínez, D., Braun, L. M., Mendez, M., Jerrett, M., Rodriguez, D., de Nazelle, A., & Nieuwenhuijsen, M. (2015a). Objective correlates and determinants of bicycle commuting propensity in an urban environment. *Transportation Research Part D: Transport and Environment*, *40*, 132–143. https://doi.org/10.1016/j.trd.2015.07.004

Cole-Hunter, T., Donaire-Gonzalez, D., Curto, A., Ambros, A., Valentin, A., Garcia-Aymerich, J., Martínez, D., Braun, L. M., Mendez, M., Jerrett, M., Rodriguez, D., de Nazelle, A., & Nieuwenhuijsen, M. (2015b). Objective correlates and determinants of bicycle commuting propensity in an urban environment. *Transportation Research Part D: Transport and Environment*, *40*, 132–143. https://doi.org/10.1016/j.trd.2015.07.004

Complete Streets policies nationwide. (2019). Smart Growth America. https://smartgrowthamerica.org/program/national-complete-streets-coalition/publications/policy-development/policy-atlas/

Corraini, P., Olsen, M., Pedersen, L., Dekkers, O. M., & Vandenbroucke, J. P. (2017). Effect modification, interaction and mediation: An overview of theoretical insights for clinical investigators. *Clinical Epidemiology*, *9*, 331–338. https://doi.org/10.2147/CLEP.S129728

Coughenour, C., Paz, A., de la Fuente-Mella, H., & Singh, A. (2016). Multinomial logistic regression to estimate and predict perceptions of bicycle and transportation infrastructure in a sprawling metropolitan area. *Journal of Public Health*, *38*(4), e401–e408. https://doi.org/10.1093/pubmed/fdv179

Cradock, A. L., Troped, P. J., Fields, B., Melly, S. J., Simms, S. V., Gimmler, F., & Fowler, M. (2009). Factors associated with federal transportation funding for local pedestrian and bicycle programming and facilities. *Journal of Public Health Policy*, *30 Suppl 1*, S38-72. https://doi.org/10.1057/jphp.2008.60

Dalen, J. E., Alpert, J. S., Goldberg, R. J., & Weinstein, R. S. (2014). The Epidemic of the 20th Century: Coronary Heart Disease. *The American Journal of Medicine*, *127*(9), 807–812. https://doi.org/10.1016/j.amjmed.2014.04.015

Dill, J. (2009). Bicycling for transportation and health: The role of infrastructure. *Journal of Public Health Policy*, *30 Suppl 1*, S95-110. https://doi.org/10.1057/jphp.2008.56

Dill, J., & Carr, T. (2003). Bicycle Commuting and Facilities in Major U.S. Cities. *Transportation Research Board Annual Meeting*, 9.

Dill, J., Mohr, C., & Ma, L. (2014). How Can Psychological Theory Help Cities Increase Walking and Bicycling? *Journal of the American Planning Association*, *80*(1), 36–51. https://doi.org/10.1080/01944363.2014.934651

Dill, J., Smith, O., & Howe, D. (2017). Promotion of active transportation among state departments of transportation in the U.S. *Journal of Transport & Health*, *5*, 163–171. https://doi.org/10.1016/j.jth.2016.10.003

Dill, J., & Voros, K. (2007). Factors Affecting Bicycling Demand: Initial Survey Findings from the Portland, Oregon, Region. *Transportation Research Record*, *2031*(1), 9–17. https://doi.org/10.3141/2031-02

Ding, D., & Gebel, K. (2012). Built environment, physical activity, and obesity: What have we learned from reviewing the literature? *Health & Place*, *18*(1), 100–105. https://doi.org/10.1016/j.healthplace.2011.08.021

Dinu, M., Pagliai, G., Macchi, C., & Sofi, F. (2019). Active Commuting and Multiple Health Outcomes: A Systematic Review and Meta-Analysis. *Sports Medicine*, *49*(3), 437–452. https://doi.org/10.1007/s40279-018-1023-0

Donaire-Gonzalez, D., de Nazelle, A., Cole-Hunter, T., Curto, A., Rodriguez, D. A., Mendez, M. A., Garcia-Aymerich, J., Basagaña, X., Ambros, A., Jerrett, M., & Nieuwenhuijsen, M. J. (2015). The Added Benefit of Bicycle Commuting on the Regular Amount of Physical

Activity Performed. *American Journal of Preventive Medicine*, *49*(6), 842–849. https://doi.org/10.1016/j.amepre.2015.03.036

Dons, E., Rojas-Rueda, D., Anaya-Boig, E., Avila-Palencia, I., Brand, C., Cole-Hunter, T., de Nazelle, A., Eriksson, U., Gaupp-Berghausen, M., Gerike, R., Kahlmeier, S., Laeremans, M., Mueller, N., Nawrot, T., Nieuwenhuijsen, M. J., Orjuela, J. P., Racioppi, F., Raser, E., Standaert, A., … Götschi, T. (2018). Transport mode choice and body mass index: Cross-sectional and longitudinal evidence from a European-wide study. *Environment International*, *119*, 109–116. https://doi.org/10.1016/j.envint.2018.06.023

Duchosal, D., Groen, J., Hilleboe, H., Morris, J., Rojas Villegas, F., Rutstein, D., Speransky, J., & Torgersen, O. (1958). *Hypertension and Coronary Heart Disease: Classification and Criteria for Epidemiological Studies* (No. 168; Technical Report). World Health Organization.

Durand, C. P., Andalib, M., Dunton, G. F., Wolch, J., & Pentz, M. A. (2011). A systematic review of built environment factors related to physical activity and obesity risk: Implications for smart growth urban planning: Smart growth urban planning and obesity risk. *Obesity Reviews*, *12*(5), e173–e182. https://doi.org/10.1111/j.1467-789X.2010.00826.x

Ewing, R., & Hamidi, S. (2013). *Measuring Urban Sprawl and Validating Sprawl Measures*. Metropolitan Research Center. https://gis.cancer.gov/tools/urban-sprawl/

Fatality Analysis Reporting System [CSV data file]. (2017). USDOT National Highway Traffic Safety Administration. https://www.nhtsa.gov/node/97996/251

Federal Highway Administration. (2013). *Memorandum: Bicycle and Pedestrian Facility Design Flexibility*. USDOT Federal Highway Administration.

https://www.fhwa.dot.gov/environment/bicycle_pedestrian/guidance/design_flexibility.cfm

Ferreira-Pinto, L. M., Rocha-Gonçalves, F., & Teixeira-Pinto, A. (2012). An ecological study on the geographic patterns of ischaemic heart disease in Portugal and its association with demography, economic factors and health resources distribution. *BMJ Open*, *2*(4), e000595. https://doi.org/10.1136/bmjopen-2011-000595

Field, A., Miles, J., & Field, Z. (2012). Regression. In *Discovering Statistics using R* (pp. 245–311). SAGE Publications.

Fishman, E., Washington, S., Haworth, N., & Watson, A. (2015). Factors influencing bike share membership: An analysis of Melbourne and Brisbane. *Transportation Research Part A: Policy and Practice*, *71*, 17–30. https://doi.org/10.1016/j.tra.2014.10.021

Fitch, D. T., Rhemtulla, M., & Handy, S. L. (2019). The relation of the road environment and bicycling attitudes to usual travel mode to school in teenagers. *Transportation Research Part A: Policy and Practice*, *123*, 35–53. https://doi.org/10.1016/j.tra.2018.06.013

Frank, L. D., Saelens, B. E., Leary, L., Cain, K., Conway, T. L., Hess, P. M., & Sallis, J. F. (2009). *The Development of a Walkability Index: Application to the Neighborhood Quality of Life Study*. British Journal of Sports Medicine.

Friedman, G. (1994). *Primer of Epidemiology* (4th ed.). McGraw-Hill Inc.

Garrard, J., Rissel, C., & Bauman, A. (2012). Health Benefits of Cycling. In *City Cycling* (pp. 31–55). Massachusetts Institute of Technology.

Garrard, J., Rose, G., & Lo, S. K. (2008). Promoting transportation cycling for women: The role of bicycle infrastructure. *Preventive Medicine*, *46*(1), 55–59. https://doi.org/10.1016/j.ypmed.2007.07.010

Gauvin, L., Richard, L., Craig, C. L., Spivock, M., Riva, M., Forster, M., Laforest, S., Laberge, S., Fournel, M.-C., Gagnon, H., Gagné, S., & Potvin, L. (2005). From walkability to active living potential: An “ecometric” validation study. *American Journal of Preventive Medicine*, *28*(2, Supplement 2), 126–133. https://doi.org/10.1016/j.amepre.2004.10.029

Goldsmith, S. (1992). *National bicycling and walking study, case study no. 1: Reasons why bicycling and walking are not being used more extensively as travel modes* (FHWA-PD-92-041). FHWA, U.S. Department of Transportation. https://safety.fhwa.dot.gov/ped_bike/docs/case1.pdf

Goodno, M., McNeil, N., Parks, J., & Dock, S. (2013). Evaluation of Innovative Bicycle Facilities in Washington, D.C.: Pennsylvania Avenue Median Lanes and 15th Street Cycle Track. *Transportation Research Record*, *2387*(1), 139–148. https://doi.org/10.3141/2387-16

Gotschi, T. (2011). Costs and benefits of bicycling investments in Portland, Oregon. *Journal of Physical Activity & Health*, *8 Suppl 1*, S49-58. https://doi.org/10.1123/jpah.8.s1.s49

Guidance for Data Users. (2020, May 4). United States Census Bureau. https://www.census.gov/programs-surveys/acs/guidance.html

Hamer, M., & Chida, Y. (2008). Active commuting and cardiovascular risk: A meta-analytic review. *Preventive Medicine*, *46*(1), 9–13. https://doi.org/10.1016/j.ypmed.2007.03.006

Hames, C. G. (1971). Evans County cardiovascular and cerebrovascular epidemiologic study. Introduction. *Archives of Internal Medicine*, *128*(6), 883–886. https://doi.org/10.1001/archinte.128.6.883

Handy, S. L., McCann, B., Bailey, L., Ernst, M., McRee, L., Meharg, E., Ewing, R., & Wright, K. (2009). *The Regional Response to Federal Funding for Bicycle and Pedestrian Projects*. https://escholarship.org/uc/item/26j7x815

Handy, S. L., & Xing, Y. (2011). Factors Correlated with Bicycle Commuting: A Study in Six Small U.S. Cities. *International Journal of Sustainable Transportation*, *5*(2), 91–110. https://doi.org/10.1080/15568310903514789

Handy, S, & Transportation Research Board. (2005). *Does the Built Environment Influence Physical Activity?: Examining the Evidence—Special Report 282*. Transportation Research Board. https://doi.org/10.17226/11203

Handy, S., van Wee, B., & Kroesen, M. (2014). Promoting Cycling for Transport: Research Needs and Challenges. *Transport Reviews*, *34*(1), 4–24. https://doi.org/10.1080/01441647.2013.860204

Hankey, S., Marshall, J. D., & Brauer, M. (2012). Health Impacts of the Built Environment: Within-Urban Variability in Physical Inactivity, Air Pollution, and Ischemic Heart Disease Mortality. *Environmental Health Perspectives*, *120*(2), 247–253. JSTOR.

Hegger, R. (2007). Public Transport and Cycling: Living Apart or Together? *Public Transport International*, *56*(2). https://trid.trb.org/view/808957

Heron, M. (2019). *Deaths: Leading Causes for 2017* (p. 77) [National Vital Statistics]. U.S. Department of Health and Human Services.

Hoehner, C. M., Brennan Ramirez, L. K., Elliott, M. B., Handy, S. L., & Brownson, R. C. (2005). Perceived and objective environmental measures and physical activity among urban adults. *American Journal of Preventive Medicine*, *28*(2 Suppl 2), 105–116. https://doi.org/10.1016/j.amepre.2004.10.023

Hruschka, D. J., & Hadley, C. (2008). A glossary of culture in epidemiology. *Journal of Epidemiology and Community Health*, *62*(11), 947–951. https://doi.org/10.1136/jech.2008.076729

Hu, G., Pekkarinen, H., Hanninen, O., Yu, Z., Guo, Z., & Tian, H. (2002). Commuting, leisure-time physical activity, and cardiovascular risk factors in China: *Medicine and Science in Sports and Exercise*, *34*(2), 234–238. https://doi.org/10.1097/00005768-200202000-00009

Hu, G., Qiao, Q., Silventoinen, K., Eriksson, J. G., Jousilahti, P., Lindström, J., Valle, T. T., Nissinen, A., & Tuomilehto, J. (2003). Occupational, commuting, and leisure-time physical activity in relation to risk for Type 2 diabetes in middle-aged Finnish men and women. *Diabetologia*, *46*(3), 322–329. https://doi.org/10.1007/s00125-003-1031-x

Hu, Tuomilehto, J., Borodulin, K., & Jousilahti, P. (2007). The joint associations of occupational, commuting, and leisure-time physical activity, and the Framingham risk score on the 10-year risk of coronary heart disease. *European Heart Journal*, *28*(4), 492–498. https://doi.org/10.1093/eurheartj/ehl475

Hunt, J. D., & Abraham, J. E. (2007). Influences on bicycle use. *Transportation*, *34*(4), 453–470. https://doi.org/10.1007/s11116-006-9109-1

Huy, C., Becker, S., Gomolinsky, U., Klein, T., & Thiel, A. (2008). Health, medical risk factors, and bicycle use in everyday life in the over-50 population. *Journal of Aging and Physical Activity*, *16*(4), 454–464. https://doi.org/10.1123/japa.16.4.454

Jekel, J., Elmore, J., & Katz, D. (1996). Methods of Tertiary Prevention. In *Epidemiology, Biostatistics, and Preventative Medicine* (1st ed., pp. 225–232). W.B. Saunders Company.

Kaczynski, A. (2008). Smoking and Physical Activity: A Systematic Review. *American Journal of Health Behavior*, *32*(1). https://doi.org/10.5993/AJHB.32.1.9

Kannel, W., Kagan, A., & Stokes, J. (1961). Factors of Risk in the Development of Coronary Heart Disease- Six Year Follow-up Experience. *Annals of Internal Medicine*, *55*(1). http://www.medicine.mcgill.ca/epidemiology/courses/EPIB591/Fall%202010/Class%208%20-%2024%20Sept/FraminghamSixYear.pdf

Kaplan, G. A., & Keil, J. E. (1993). Socioeconomic factors and cardiovascular disease: A review of the literature. *Circulation*, *88*(4 Pt 1), 1973–1998. https://doi.org/10.1161/01.cir.88.4.1973

Karasek, R. A., Theorell, T., Schwartz, J. E., Schnall, P. L., Pieper, C. F., & Michela, J. L. (1988). Job characteristics in relation to the prevalence of myocardial infarction in the US Health Examination Survey (HES) and the Health and Nutrition Examination Survey (HANES). *American Journal of Public Health*, *78*(8), 910–918. https://doi.org/10.2105/ajph.78.8.910

Khare, M. M., Koch, A., Zimmermann, K., Moehring, P. A., & Geller, S. E. (2014). Heart Smart for Women: A Community-Based Lifestyle Change Intervention to Reduce Cardiovascular Risk in Rural Women. *The Journal of Rural Health*, *30*(4), 359–368. https://doi.org/10.1111/jrh.12066

Krieger, N. (2001). Theories for social epidemiology in the 21st century: An ecosocial perspective. *International Journal of Epidemiology*, *30*(4), 668–677. https://doi.org/10.1093/ije/30.4.668

Lee, C., & Moudon, A. V. (2006). Correlates of Walking for Transportation or Recreation Purposes. *Journal of Physical Activity & Health*, *3*(s1), S77–S98. https://doi.org/10.1123/jpah.3.s1.s77

Lovallo, W. R., & Gerin, W. (2003). Psychophysiological Reactivity: Mechanisms and Pathways to Cardiovascular Disease: *Psychosomatic Medicine*, *65*(1), 36–45. https://doi.org/10.1097/01.PSY.0000033128.44101.C1

Lusk, A. C., Mekary, R. A., Feskanich, D., & Willett, W. C. (2010). Bicycle Riding, Walking, and Weight Gain in Premenopausal Women. *Archives of Internal Medicine*, *170*(12), 1050–1056. https://doi.org/10.1001/archinternmed.2010.171

Ma, X., Sharpe, P. A., Bell, B. A., Liu, J., White, K., & Liese, A. D. (2018). Food Acquisition and Shopping Patterns among Residents of Low-Income and Low-Access Communities in South Carolina. *Journal of the Academy of Nutrition and Dietetics*, *118*(10), 1844–1854. https://doi.org/10.1016/j.jand.2018.04.017

Mackett, R. L., & Thoreau, R. (2015). Transport, social exclusion and health. *Journal of Transport & Health*, *2*(4), 610–617. https://doi.org/10.1016/j.jth.2015.07.006

MacMahon, B., & Trichopoulos, D. (1996). *Epidemiology Principles and Methods* (2nd ed.). LWW.

Marcus, B. H., Dubbert, P. M., Forsyth, L. H., McKenzie, T. L., Stone, E. J., Dunn, A. L., & Blair, S. N. (2000). Physical activity behavior change: Issues in adoption and maintenance. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, *19*(1S), 32–41.

Marleau Donais, F., Abi-Zeid, I., Waygood, E. O. D., & Lavoie, R. (2019). Assessing and ranking the potential of a street to be redesigned as a Complete Street: A multi-criteria decision aiding approach. *Transportation Research Part A: Policy and Practice*, *124,* 1–19. https://doi.org/10.1016/j.tra.2019.02.006

Marmot, M., & Theorell, T. (1988). Social class and cardiovascular disease: The contribution of work. *International Journal of Health Services*, *18*(4), 659–674. https://doi.org/10.2190/KTC1-N5LK-J1PM-9GRQ

Marsden, G., & Stead, D. (2011). Policy transfer and learning in the field of transport: A review of concepts and evidence. *Transport Policy*, *18*(3), 492–500. https://doi.org/10.1016/j.tranpol.2010.10.007

Martens, K. (2007). Promoting bike-and-ride: The Dutch experience. *Transportation Research Part A: Policy and Practice*, *41*(4), 326–338. https://doi.org/10.1016/j.tra.2006.09.010

McDonough, J. R., Hames, C. G., Stulb, S. C., & Garrison, G. E. (1965). Coronary heart disease among Negroes and Whites in Evans County, Georgia. *Journal of Chronic Diseases*, *18*(5), 443–468. https://doi.org/10.1016/0021-9681(65)90027-5

McLeod, K. (2017). *Lifting the Veil on Bicycle & Pedestrian Spending*. Alliance for Biking & Walking, The League of American Bicyclists. https://bikeleague.org/sites/default/files/LiftingTheVeil_ReportScoreCards.pdf

McLeod, K., Herpolsheimer, S., & Clarke, K. (2019). *Bicycling and Walking in the United States: 2018 Benchmarking Report*. League of American Bicyclists.

McLeod, L., Bharadwaj, L., Epp, T. Y., & Waldner, C. L. (2018). Ecological analysis of associations between groundwater quality and hypertension and cardiovascular disease in rural Saskatchewan, Canada using Bayesian hierarchical models and administrative health data. *Environmental Research*, *167*, 329–340. https://doi.org/10.1016/j.envres.2018.07.038

McLeroy, K. R., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological perspective on health promotion programs. *Health Education Quarterly*, *15*(4), 351–377.

McLeroy, K. R., Norton, B. L., Kegler, M. C., Burdine, J. N., & Sumaya, C. V. (2003). Community-based interventions. *American Journal of Public Health*, *93*(4), 529–533. https://doi.org/10.2105/ajph.93.4.529

Mekuria, M. C., & Nixon, H. (2012). *Low-Stress Bicycling and Network Connectivity* (p. 2). Mineta Transportation Institute.

Ming Wen, L., & Rissel, C. (2008). Inverse associations between cycling to work, public transport, and overweight and obesity: Findings from a population based study in Australia. *Preventive Medicine*, *46*(1), 29–32. https://doi.org/10.1016/j.ypmed.2007.08.009

Minkler, M., & Wallerstein, N. (Eds.). (2003). *Community based participatory research for health*. Jossey-Bass.

Moreland-Russell, S., Eyler, A., Barbero, C., Hipp, J., & Walsh, H. (2013). Diffusion of Complete Streets Policies Across US Communities. *Journal of Public Health Management and Practice*, *19*. https://doi.org/10.1097/PHH.0b013e3182849ec2

Morland, K., Diez Roux, A. V., & Wing, S. (2006). Supermarkets, other food stores, and obesity: The atherosclerosis risk in communities study. *American Journal of Preventive Medicine*, *30*(4), 333–339. https://doi.org/10.1016/j.amepre.2005.11.003

Morland, K., Wing, S., & Roux, A. D. (2002). The Contextual Effect of the Local Food Environment on Residents' Diets: The Atherosclerosis Risk in Communities Study. *American Journal of Public Health*, *92*(11), 1761–1768. https://doi.org/10.2105/ajph.92.11.1761

Moudon, A. V., Lee, C., Cheadle, A. D., Collier, C. W., Johnson, D., Schmid, T. L., & Weather, R. D. (2005). Cycling and the built environment, a US perspective. *Transportation Research Part D: Transport and Environment*, *10*(3), 245–261. https://doi.org/10.1016/j.trd.2005.04.001

Mullenbach, L. E. (2018). Assessing the Relationship Between a Composite Score of Urban Park Quality and Health. *Preventing Chronic Disease*, *15*. https://doi.org/10.5888/pcd15.180033

Murphy, S., Xu, J., Kochanek, K., Arias, E., & Tejada-Vera, B. (2021). Deaths: Final Data for 2018. *National Vital Statistics Reports*, *69*(13).

Nelson, A. C., & Allen, D. (1997). If You Build Them, Commuters Will Use Them: Association Between Bicycle Facilities and Bicycle Commuting: *Transportation Research Record*. https://doi.org/10.3141/1578-10

NOAA National Centers for Environmental Information. (2010). *1981-2010 U.S. Climate Normals [TXT data file]*. https://www1.ncdc.noaa.gov/pub/data/normals/1981-2010/

Olson, R., Piercy, K., Troiano, R., Ballard, R., Fulton, J., Galuska, D., & Pfohl, S. (2018). *Physical Activity Guidelines for Americans, 2nd edition* (p. 118). U.S Dept. HHS.

Orth-Gomér, K., Wamala, S. P., Horsten, M., Schenck-Gustafsson, K., Schneiderman, N., & Mittleman, M. A. (2000). Marital stress worsens prognosis in women with coronary heart disease: The Stockholm Female Coronary Risk Study. *JAMA*, *284*(23), 3008–3014. https://doi.org/10.1001/jama.284.23.3008

Parkin, J., Wardman, M., & Page, M. (2008). Estimation of the determinants of bicycle mode share for the journey to work using census data. *Transportation*, *35*(1), 93–109. https://doi.org/10.1007/s11116-007-9137-5

PeopleForBikes. (2019). *Bicycle Network Analysis Methodology*. People for Bikes. https://bna.peopleforbikes.org/#/methodology

Petrie, A. (2016a). Building a Descriptive Model. In *Introduction to Regression and Modeling with R* (1st ed., pp. 253–268). Cognella, Inc.

Petrie, A. (2016b). Predictive Modeling with Regression. In *Introduction to Regression and Modeling with R* (1st ed., pp. 271–289). Cognella, Inc.

Plaut, P. (2005). Non-motorized commuting in the United States. *Transportation Research Part D: Transport and Environment*, *10*(5), 347–356.

Pollack, K. M., Kercher, C., Frattaroli, S., Peek-Asa, C., Sleet, D., & Rivara, F. P. (2012). Toward environments and policies that promote injury-free active living—It wouldn't hurt. *Health & Place*, *18*(1), 106–114. https://doi.org/10.1016/j.healthplace.2011.07.010

Pucher, J., & Buehler, R. (2009). Integrating Bicycling and Public Transport in North America. *Journal of Public Transportation, 12*. https://doi.org/10.5038/2375-0901.12.3.5

Pucher, J., & Buehler, R. (2012). International Overview: Cycling Trends in Western Europe, North America, and Australia. In *City Cycling* (pp. 9–29). Massachusetts Institute of Technology.

Pucher, J., Buehler, R., Bassett, D. R., & Dannenberg, A. L. (2010a). Walking and cycling to health: A comparative analysis of city, state, and international data. *American Journal of Public Health*, *100*(10), 1986–1992. https://doi.org/10.2105/AJPH.2009.189324

Pucher, J., Buehler, R., Bassett, D. R., & Dannenberg, A. L. (2010b). Walking and Cycling to Health: A Comparative Analysis of City, State, and International Data. *American Journal of Public Health*, *100*(10), 1986. https://doi.org/10.2105/AJPH.2009.189324

Pucher, J., Buehler, R., & Seinen, M. (2011). Bicycling renaissance in North America? An update and re-appraisal of cycling trends and policies. *Transportation Research Part A: Policy and Practice*, *45*(6), 451–475. https://doi.org/10.1016/j.tra.2011.03.001

Pucher, J., Dill, J., & Handy, S. (2010). Infrastructure, programs, and policies to increase bicycling: An international review. *Preventive Medicine*, *50*, S106–S125. https://doi.org/10.1016/j.ypmed.2009.07.028

Pucher, J., Komanoff, C., & Schimek, P. (1999). Bicycling renaissance in North America? Recent trends and alternative policies to promote bicycling. *Transportation Research Part A: Policy and Practice*, *33*, 625–654.

Raser, E., Gaupp-Berghausen, M., Dons, E., Anaya-Boig, E., Avila-Palencia, I., Brand, C., Castro, A., Clark, A., Eriksson, U., Götschi, T., Int Panis, L., Kahlmeier, S., Laeremans, M., Mueller, N., Nieuwenhuijsen, M., Orjuela, J. P., Rojas-Rueda, D., Standaert, A., Stigell, E., & Gerike, R. (2018). European cyclists' travel behavior: Differences and similarities between seven European (PASTA) cities. *Journal of Transport & Health*, *9*, 244–252. https://doi.org/10.1016/j.jth.2018.02.006

Reis, R. S., Hino, A. A. F., Parra, D. C., Hallal, P. C., & Brownson, R. C. (2013). Bicycling and Walking for Transportation in Three Brazilian Cities. *American Journal of Preventive Medicine*, *44*(2), e9–e17. https://doi.org/10.1016/j.amepre.2012.10.014

Rérat, P. (2019). Cycling to work: Meanings and experiences of a sustainable practice. *Transportation Research Part A: Policy and Practice*, *123*, 91–104. https://doi.org/10.1016/j.tra.2018.10.017

Rietveld, P. (2000). The accessibility of railway stations: The role of the bicycle in The Netherlands. *Transportation Research Part D: Transport and Environment*, *5*(1), 71–75. https://doi.org/10.1016/S1361-9209(99)00019-X

Riiser, A., Solbraa, A., Jenum, A. K., Birkeland, K. I., & Andersen, L. B. (2018). Cycling and walking for transport and their associations with diabetes and risk factors for cardiovascular disease. *Journal of Transport & Health*, *11*, 193–201. https://doi.org/10.1016/j.jth.2018.09.002

Riveron, N. (2019). *The Best Complete Streets Policies of 2018*. Smart Growth America, National Complete Streets Coalition.

Ross, N., & Hermann, T. (2019). *Meeting Moderate-to-Vigorous Physical Activity 'Automatically' in Supportive Active Living Environments.* Active Living Conference, Charleston.

Sabouri, S., Park, K., Smith, A., Tian, G., & Ewing, R. (2020). Exploring the influence of built environment on Uber demand. *Transportation Research Part D: Transport and Environment*, *81*. https://doi.org/10.1016/j.trd.2020.102296

Saelens, B. E., & Handy, S. L. (2008). Built Environment Correlates of Walking: A Review. *Medicine & Science in Sports & Exercise*, *40*(7). https://doi.org/10.1249/MSS.0b013e31817c67a4

SAFE Streets Act, Pub. L. No. H.R. 508 (2021). https://www.congress.gov/bill/117th-congress/house-bill/508?s=1&r=6

Sallis, J. F., Conway, T. L., Dillon, L. I., Frank, L. D., Adams, M. A., Cain, K. L., & Saelens, B. E. (2013). Environmental and demographic correlates of bicycling. *Preventive Medicine*, *57*(5), 456–460. https://doi.org/10.1016/j.ypmed.2013.06.014

Sallis, J. F., Spoon, C., Cavill, N., Engelberg, J. K., Gebel, K., Parker, M., Thornton, C. M., Lou, D., Wilson, A. L., Cutter, C. L., & Ding, D. (2015). Co-benefits of designing communities for active living: An exploration of literature. *International Journal of Behavioral Nutrition and Physical Activity*, *12*(1), 30. https://doi.org/10.1186/s12966-015-0188-2

Salvo, G., Lashewicz, B., Doyle-Baker, P., & McCormack, G. (2018). Neighbourhood Built Environment Influences on Physical Activity among Adults: A Systematized Review of Qualitative Evidence. *International Journal of Environmental Research and Public Health*, *15*(5), 897. https://doi.org/10.3390/ijerph15050897

Sanders, R. L., & Cooper, J. F. (2013). Do All Roadway Users Want the Same Things?: Results from Roadway Design Survey of San Francisco Bay area Pedestrians, Drivers, Bicyclists, and Transit Users. *Transportation Research Record*, *2393*(1), 155–163. https://doi.org/10.3141/2393-18

Schoner, J. E., & Levinson, D. M. (2014). The missing link: Bicycle infrastructure networks and ridership in 74 US cities. *Transportation*, *41*(6), 1187–1204. https://doi.org/10.1007/s11116-014-9538-1

Schwartz, A. R., Gerin, W., Davidson, K. W., Pickering, T. G., Brosschot, J. F., Thayer, J. F., Christenfeld, N., & Linden, W. (2003). Toward a Causal Model of Cardiovascular Responses to Stress and the Development of Cardiovascular Disease: *Psychosomatic Medicine*, *65*(1), 22–35. https://doi.org/10.1097/01.PSY.0000046075.79922.61

Shafizadeh, K., & Niemeier, D. (1997). Bicycle Journey-to-Work: Travel Behavior Characteristics and Spatial Attributes. *Transportation Research Record*, *1578*(1), 84–90. https://doi.org/10.3141/1578-11

Shea, S., & Basch, C. E. (1990). A review of five major community-based cardiovascular disease prevention programs. Part II: Intervention strategies, evaluation methods, and results. *American Journal of Health Promotion: AJHP*, *4*(4), 279–287. https://doi.org/10.4278/0890-1171-4.4.279

Shlens, J. (2014). *A Tutorial on Principal Component Analysis*. arXiv. http://arxiv.org/abs/1404.1100

Simons-Morton, B., McLeroy, K., & Wendel, M. (2012a). A Social Ecological Perspective. In *Behavior Theory in Health Promotion Practice and Research* (pp. 41–68). Hones & Bartlett Learning, LLC.

Simons-Morton, B., McLeroy, K., & Wendel, M. (2012b). Expectancy Value Models. In *Behavior Theory in Health Promotion Practice and Research* (pp. 97–126). Hones & Bartlett Learning, LLC.

Simons-Morton, B., McLeroy, K., & Wendel, M. (2012c). Social Influence Theory. In *Behavior Theory in Health Promotion Practice and Research* (pp. 155–179). Hones & Bartlett Learning, LLC.

Simons-Morton, D. G., Simons-Morton, B. G., Parcel, G. S., & Bunker, J. F. (1988). Influencing personal and environmental conditions for community health: A multilevel intervention model. *Family & Community Health*, *11*(2), 25–35.

Snizek, B., Sick Nielsen, T. A., & Skov-Petersen, H. (2013). Mapping bicyclists' experiences in Copenhagen. *Journal of Transport Geography*, *30*, 227–233. https://doi.org/10.1016/j.jtrangeo.2013.02.001

Stewart, O. T., Vernez Moudon, A., Saelens, B. E., Lee, C., Kang, B., & Doescher, M. P. (2016). Comparing Associations Between the Built Environment and Walking in Rural Small

Towns and a Large Metropolitan Area. *Environment and Behavior*, *48*(1), 13–36. https://doi.org/10.1177/0013916515612253

Stinson, M. A., & Bhat, C. R. (2003). Commuter Bicyclist Route Choice: Analysis Using a Stated Preference Survey. *Transportation Research Record*, *1828*(1), 107–115. https://doi.org/10.3141/1828-13

Suminski, R. R., Wasserman, J. A., Mayfield, C. A., Freeman, E., & Brandl, R. (2014). Bicycling Policy Indirectly Associated with Overweight/Obesity. *American Journal of Preventive Medicine*, *47*(6), 715–721. https://doi.org/10.1016/j.amepre.2014.07.048

Susser, M. (1994). The logic in ecological: I. The logic of analysis. *American Journal of Public Health*, *84*(5), 825–829.

The National FINRISK Study. (2017). Finnish Institute for Health and Welfare. https://thl.fi/en/web/thlfi-en/research-and-expertwork/population-studies/the-national-finrisk-study

Tilahun, N. Y., Levinson, D. M., & Krizek, K. J. (2007). Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A: Policy and Practice*, *41*(4), 287–301. https://doi.org/10.1016/j.tra.2006.09.007

Timio, M., Lippi, G., Venanzi, S., Gentili, S., Quintaliani, G., Verdura, C., Monarca, C., Saronio, P., & Timio, F. (1997). Blood pressure trend and cardiovascular events in nuns in a secluded order: A 30-year follow-up study. *Blood Pressure*, *6*(2), 81–87.

Ton, D., Bekhor, S., Cats, O., Duives, D. C., Hoogendoorn-Lanser, S., & Hoogendoorn, S. P. (2020). The experienced mode choice set and its determinants: Commuting trips in the Netherlands. *Transportation Research Part A: Policy and Practice*, *132*, 744–758. https://doi.org/10.1016/j.tra.2019.12.027

Transportation Alternatives Enhancements Act, Pub. L. No. H.R. 463 (2021). https://www.congress.gov/bill/117th-congress/house-bill/463

Tribby, C. P., & Tharp, D. S. (2019). Examining urban and rural bicycling in the United States: Early findings from the 2017 National Household Travel Survey. *Journal of Transport & Health*, *13*, 143–149. https://doi.org/10.1016/j.jth.2019.03.015

U.S. Census Bureau. (2020). *Understanding and using American Community Survey data: What all data users need to know*. U.S. Government Publishing Office. https://www.census.gov/content/dam/Census/library/publications/2020/acs/acs_general_handbook_2020.pdf

USDOT. (2010). *Policy Statement on Bicycle and Pedestrian Accommodation Regulations and Recommendations*. FHWA. https://www.fhwa.dot.gov/environment/bicycle_pedestrian/guidance/policy_accom.cfm

USDOT Federal Highway Administration. (2018). *NHTS Data User Guide*. https://nhts.ornl.gov/assets/NHTS2017_UsersGuide_04232019_1.pdf

Uyanık, T., Karatuğ, Ç., & Arslanoğlu, Y. (2020). Machine learning approach to ship fuel consumption: A case of container vessel. *Transportation Research Part D: Transport and Environment*, *84*. https://doi.org/10.1016/j.trd.2020.102389

van Buuren, S. (2020). *Package “mice.”* https://cran.r-project.org/web/packages/mice/mice.pdf

Walsh, S. M., Umstattd Meyer, M. R., Gamble, A., Patterson, M. S., & Moore, J. B. (2017). A Systematic Review of Rural, Theory-based Physical Activity Interventions. *American Journal of Health Behavior*, *41*(3), 248–258. https://doi.org/10.5993/AJHB.41.3.4

Wang, M. C., Kim, S., Gonzalez, A. A., MacLeod, K. E., & Winkleby, M. A. (2007). Socioeconomic and food-related physical characteristics of the neighbourhood environment are associated with body mass index. *Journal of Epidemiology and Community Health*, *61*(6), 491–498. https://doi.org/10.1136/jech.2006.051680

Wang, Y. (2017). Comparison of methods for estimating prevalence of chronic diseases and health behaviors for small geographic areas: Boston validation study, 2013. *Preventing Chronic Disease*, *14*. https://doi.org/10.5888/pcd14.170281

Wardman, M., Tight, M., & Page, M. (2007). Factors influencing the propensity to cycle to work. *Transportation Research Part A: Policy and Practice, 41*(4), 339–350. https://doi.org/10.1016/j.tra.2006.09.011

Waygood, E., Benard, M., Ishimo, Y., Michaud-Champagne, E., & Weinbuch, J. (2019). *Perceptions of Bicycle Safety in a Small City*. Active Living Conference, Charleston.

Wennberg, P., Lindahl, B., Hallmans, G., Messner, T., Weinehall, L., Johansson, L., Boman, K., & Jansson, J.-H. (2006). The effects of commuting activity and occupational and leisure time physical activity on risk of myocardial infarction. *European Journal of Cardiovascular Prevention and Rehabilitation: Official Journal of the European Society of Cardiology, Working Groups on Epidemiology & Prevention and Cardiac Rehabilitation and Exercise Physiology*, *13*(6), 924–930. https://doi.org/10.1097/01.hjr.0000239470.49003.c3

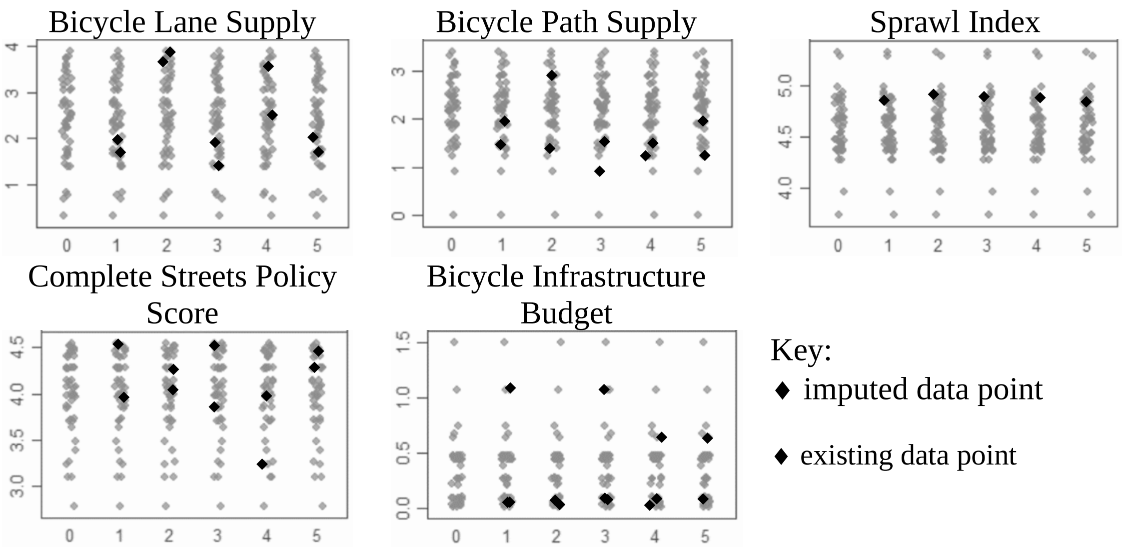
Williams, J., & Larson, J. (1996). Promoting bicycle commuting: Understanding the customer. *Transportation Quarterly, 50*(3), 67–78.

Wilson, P. W., D’Agostino, R. B., Levy, D., Belanger, A. M., Silbershatz, H., & Kannel, W. B. (1998). Prediction of coronary heart disease using risk factor categories. *Circulation*, *97*(18), 1837–1847. https://doi.org/10.1161/01.cir.97.18.1837

Yencha, C. (2019). Valuing walkability: New evidence from computer vision methods. *Transportation Research Part A: Policy and Practice*, *130*, 689–709. https://doi.org/10.1016/j.tra.2019.09.053

Zhang, X., Holt, J. B., Lu, H., Wheaton, A. G., Ford, E. S., Greenlund, K. J., & Croft, J. B. (2014). Multilevel Regression and Poststratification for Small-Area Estimation of Population Health Outcomes: A Case Study of Chronic Obstructive Pulmonary Disease Prevalence Using the Behavioral Risk Factor Surveillance System. *American Journal of Epidemiology*, *179*(8), 1025–1033. https://doi.org/10.1093/aje/kwu018

Zieff, S. G., Hipp, A., Eyler, A. A., & Kim, M.-S. (2013). Ciclovía initiatives: Engaging communities, partners and policymakers along the route to success. *Journal of Public Health Management and Practice : JPHMP*, *19*(3 0 1), S74–S82. https://doi.org/10.1097/PHH.0b013e3182841982



[1]All variables were transformed before imputation so that the regression models would meet assumptions of normally distributed residuals.

Figure A1: Strip plots of imputed data points over five imputations[1]

Model A1: The association between CAD prevalence and bicycle commuting rate when missing data is removed from the sample

```
Residuals:
    Min      1Q  Median      3Q     Max 
-5.4902 -1.2902 -0.2137  1.0758 28.4098 

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)    
(Intercept)             5.990195   0.020370  294.07   <2e-16 ***
Percent.Bike.Commuters -0.214250   0.008207  -26.11   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.997 on 12261 degrees of freedom
  (16 observations deleted due to missingness)
Multiple R-squared:  0.05266,	Adjusted R-squared:  0.05258 
F-statistic: 681.5 on 1 and 12261 DF,  p-value: < 2.2e-16
```

Model A2: The association between CAD prevalence and bicycle commuting rate when missing data is imputed

```
Residuals:
    Min      1Q  Median      3Q     Max 
-5.4885 -1.2885 -0.2199  1.0729 28.4115 

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)    
(Intercept)            5.98855    0.02036  294.14   <2e-16 ***
Percent.Bike.Commuters -0.21406    0.00821  -26.07   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.001 on 12322 degrees of freedom
Multiple R-squared:  0.05228,	Adjusted R-squared:  0.05221 
F-statistic: 679.8 on 1 and 12322 DF,  p-value: < 2.2e-16
```

Model A3: The association between CAD prevalence and bicycle commuting rate when missing data is imputed and two outlier census tracts are excluded from the sample

```
Residuals:
    Min      1Q  Median      3Q     Max 
-5.4837 -1.2837 -0.2178  1.0736 13.4495 

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)    
(Intercept)             5.983741   0.020128  297.28   <2e-16 ***
Percent.Bike.Commuters -0.213160   0.008116  -26.26   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.978 on 12320 degrees of freedom
Multiple R-squared:  0.05302,	Adjusted R-squared:  0.05294 
F-statistic: 689.7 on 1 and 12320 DF,  p-value: < 2.2e-16
```

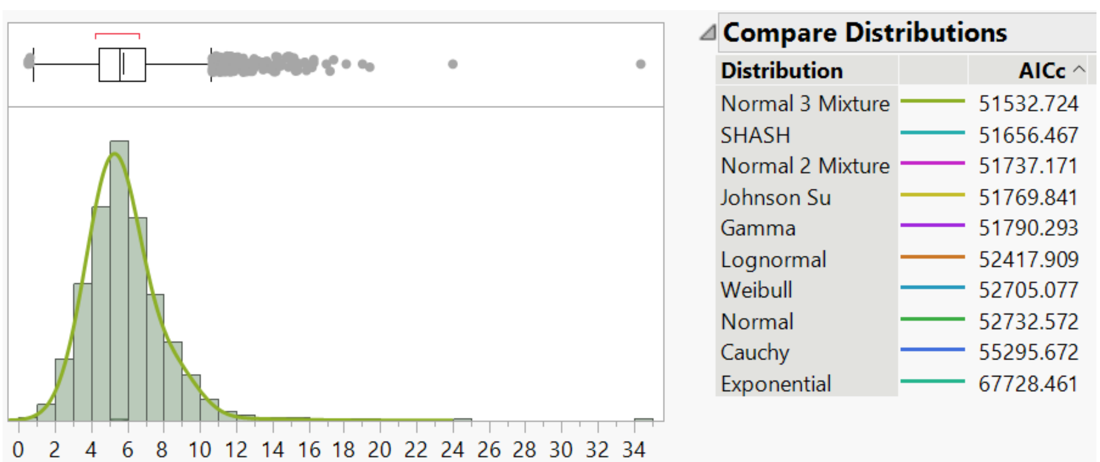


Figure A2: Distribution of CAD prevalence across 12,322 census tracts

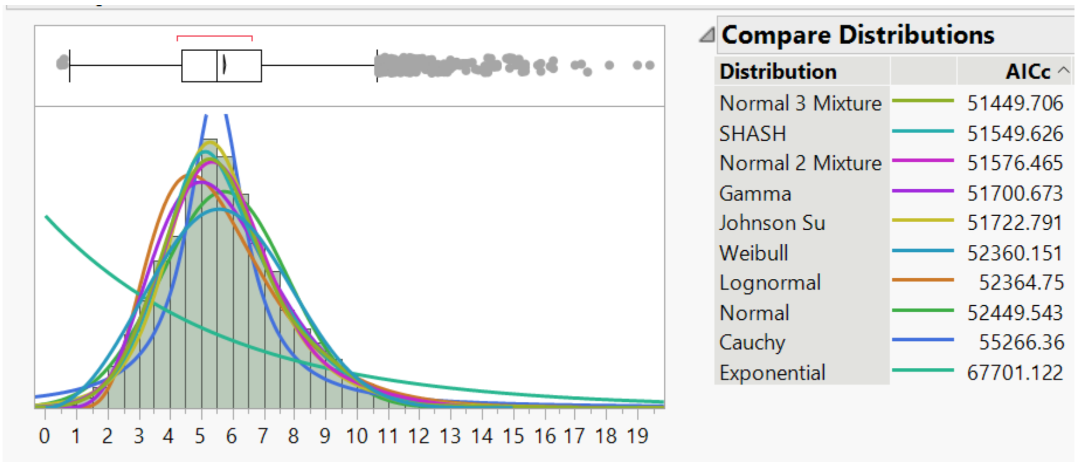


Figure A3: Distribution of CAD prevalence across 12,320 census tracts

Model B1: Exploratory regression model of obesity, diabetes, and hypertension as effect modifiers of bicycle commuting and CAD prevalence

```
Residuals:
    Min      1Q  Median      3Q     Max 
-3.5677 -0.5157 -0.0433  0.4536  7.3393 

Coefficients:
                                                              Estimate Std. Error t value Pr(>|t|)    
(Intercept)                                                 -3.6125499  0.1217694 -29.667  < 2e-16 ***
Percent.Bike.Commuters                                       0.0229306  0.0179560   1.277 0.201612    
Estimate.Hypertension.Prevalence                             0.4713918  0.0085921  54.863  < 2e-16 ***
Estimate.Diabetes.Prevalence                                -0.3248170  0.0171607 -18.928  < 2e-16 ***
Estimate.Obesity.Prevalence                                  0.0939585  0.0062085  15.134  < 2e-16 ***
Percent.Bike.Commuters:Estimate.Hypertension.Prevalence     -0.0090410  0.0012683  -7.128 1.07e-12 ***
Percent.Bike.Commuters:Estimate.Diabetes.Prevalence          0.0152081  0.0025327   6.005 1.97e-09 ***
Percent.Bike.Commuters:Estimate.Obesity.Prevalence           0.0034470  0.0008897   3.874 0.000107 ***
Estimate.Hypertension.Prevalence:Estimate.Diabetes.Prevalence  0.0032447  0.0003272   9.916  < 2e-16 ***
Estimate.Hypertension.Prevalence:Estimate.Obesity.Prevalence  -0.0110351  0.0002691 -41.006  < 2e-16 ***
Estimate.Diabetes.Prevalence:Estimate.Obesity.Prevalence     0.0145338  0.0005284  27.505  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8431 on 12311 degrees of freedom
Multiple R-squared:  0.828,	Adjusted R-squared:  0.8279 
F-statistic:  5928 on 10 and 12311 DF,  p-value: < 2.2e-16
```

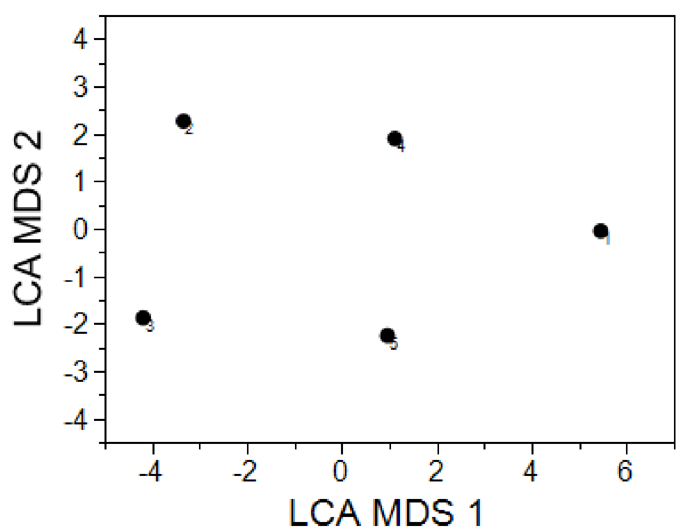


Figure B1: Multi-dimensional scaling plot of the distance between census tract clusters

Table B1: Percentage of census tracts with above median prevalence by effect modifier characteristic and cluster number

#	No College Education	White Race	Latinx Ethnicity	Median Income	Manual Occupation	Food Access	Smoking Prevalence	Obesity Prevalence	Diabetes Prevalence
1	2	89	29	95	9	51	1	3	3
2	98	59	87	3	98	42	88	95	89
3	88	2	26	5	50	52	98	99	98
4	40	83	52	58	62	24	46	60	27
5	53	19	72	64	61	83	14	24	69

Table B2: Average census tract demographics, socioeconomics, and health risk factors by cluster number

Census Tract Characteristics	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5	
	<45	*>45*	*<45*	*>45*	*<45*	*>45*	*<45*	*>45*	*<45*	*>45*
Bicycle Commuting (%)	2.7	1.2	0.8	0.7	0.7	0.6	1.1	0.8	1.7	0.8
Sex (% Male)	49.8	48.5	49.7	49.2	46.1	46.1	50.5	49.1	49.9	48.1
Age (% > 45)	29.4	45.0	29.2	40.4	31.3	42.4	27.6	44.0	30.4	44.3
Married (%)	6.3	5.0	6.6	6.7	7.5	8.3	5.9	5.2	8.8	8.5
White (%)	73.9	80.6	54.4	69.2	11.0	6.4	65.9	74.5	38.7	33.4
Black (%)	5.5	2.9	12.0	10.4	75.1	88.3	17.1	12.6	18.1	27.8
Latinx (%)	12.1	9.5	63.6	33.7	6.1	2.4	28.6	22.6	45.4	26.9
Median Income ($1K)	76.6	87.6	35.6	36.3	28.1	30.0	53.4	58.0	52.0	60.0
No College Education (%)	19.4	26.2	62.3	57.2	56.8	54.5	39.6	40.9	45.2	41.2
Manual Labor Job (%)	4.1	4.4	15.3	13.1	7.6	8.0	9.6	9.3	11.1	8.7
Healthy Food Access	9.0	46.8	38.6	53.1	27.3	42.4	60.6	63.6	14.6	15.6
Smoking Prevalence (%)	12.7	11.8	21.8	21.3	26.3	25.1	18.7	17.6	16.3	15.3
Obesity Prevalence (%)	22.1	22.7	36	35.3	41.4	43.3	30.4	30.2	26.4	25.1
Diabetes Prevalence (%)	5.7	8.1	13	14	15	18.1	8.6	10.7	10.5	12.1
Hypertension Prevalence(%)	20.2	26.6	31	35.7	39.3	46.4	26.1	32.2	26.0	31.0

APPENDIX C: RECOMMENDED ACTIONS INFOGRAPHIC