DEVELOPING AN ANALYTICAL TECHNIQUE FOR THE EXPLORATION OF SPATIOTEMPORAL PROBLEMS IN PUBLIC HEALTH

by

Michael Paul Morrish
Bachelor of Arts, Ryerson University, 2013

A Major Research Paper

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Spatial Analysis

in the Program of

Spatial Analysis

Toronto, Ontario, Canada, 2015

Author's Declaration

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A MRP

I hereby declare that I am the sole author of this MRP. This is a true copy of the MRP, including any required final revisions.

I authorize Ryerson University to lend this MRP to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this MRP by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my MRP may be made electronically available to the public.

Abstract

Income-related health inequalities can affect all members of society. Public health agencies wish to better understand relationships between social determinants and health outcomes to improve health policy. However, the relationship between social determinants and health is both spatial and temporal. Certain health outcomes may take years to develop and residential mobility in populations may be high. An important challenge to understanding of these relationships is the need to analyze individuals in their spatiotemporal context.

This study develops a technique to decompose health related spatiotemporal problems into objects, contexts, and relationships. Data from the 2001 and 2006 Statistics Canada Census are spatially analyzed to develop a composite measure of spatiotemporal context. To illustrate how these composite contexts could be used to understand an individual movement, hypothetical individual space-time paths were generated. Leveraging Structured Query Language (SQL) this study's technique connects moving individuals with shifting contexts in a relational database. SQL queries explore object movements and spatiotemporal contexts.

Future studies can build on this study's technique with more frequent data samples for both individual space-time paths and spatiotemporal contexts. This study represents an incremental addition to an emerging literature on understanding and managing the uncertain geographic context problem.

Acknowledgements

"Be not simply good, be good for something."

Henry David Thoreau

Author: Walden; Or Life in the Woods

I would like to acknowledge Toronto Public Health, Dr. Peter Kedron, and my parents for their support in this research. Toronto Public Health has provided access to data and important insight into the topic of income-related health inequalities. As my academic supervisor, Dr. Kedron has provided guidance in the completion of this research. Without my parents, this higher-level education would not have been possible at this time.

Table of Contents

List of Tables

List of Figures

List of Appendices

Introduction

Experts in fields such as urban planning, transportation, health, and epidemiology have been using the theories and tools provided by geographers for some time now. Simple problems such as the mapping of points of interest like buildings, roads, customers, patients, schools, and health facilities are well understood. More advanced techniques have been used to present and analyze the spatial pattern of phenomena including population counts, voting patterns, criminal activities, retail sales patterns, and the clustering of disease cases at a point in time. Success in these earlier challenges is prompting these same experts to ask if geographers can contribute to even more complex problems. Public health agencies are one of these groups of experts; and one of their areas of interest is the topic of income-related health inequalities.

Public health agencies wish to better understand the relationship between social determinants of health and health outcomes. Advancing our understanding of health depends on geographic questions such as: (i) “Where are high prevalence of disease and high prevalence of specific social determinants of health (e.g., low income status) co-located?”, (ii) “Do concentrated areas represent true statistical clusters of disease, and are those clusters co-located with clusters of specific determinants of health?”, and (iii) “How have these patterns changed over time?”. However, as proposed, these questions overlook the fact that their answers are not just spatial problems but spatiotemporal problems. It takes many years to develop the negative health outcomes associated with income-related health inequalities. During this time residential mobility and the impact of place will play a role in how an individual’s health status develops. The challenge is to better understand the relationship between social determinants of health and health outcomes and address the problem of analyzing interactions between objects (individuals) and the effect of the spatiotemporal context (dynamic places) on those objects.

Despite the prominence of the problem, spatiotemporal contexts are only beginning to be addressed within the health geography literature (Andrienko, Andrienko, and Heurich, 2011; Kwan, 2012a). Andrienko, Andrienko, and Heurich observe that most studies explore the movement of some objects, but have ignored the spatiotemporal context that movement takes place in. This approach provides only a partial picture of the phenomena being explored. Drawing from work in other fields, the authors acknowledge the particular contributions of Lundblad, Eurenius, and Heldring (2009) who develop a system to track ship movements on the ocean while accounting for the context of changing weather conditions. Transferred to a public health context, an analogous approach could inform the study of income-related health inequalities.

This study develops a simplified methodology to contextualize health problems in both space and time by assessing both object (e.g., person) movement and context (e.g., neighbourhood) change. The methodology presented leverages concepts developed by Andrienko et al. (2011) to decompose spatiotemporal problems into object and contexts, changes in space and time. Spatial statistics are applied to census variables to develop spatiotemporal contexts for a study area. A relative score is developed for spatiotemporal contexts so that cumulative effects of different space-time paths may be considered. Structured Query Language (SQL) is then used to analyze movement objects as space-time paths and to create spatiotemporal relationships that link the movement objects and spatiotemporal contexts. This approach is applied to hypothetical individual space-time path movements for the 2001 and 2006 Census years. The spatiotemporal context used in this set of analytical techniques is defined by census variables for income, education, visible minority, and immigration status as the literature reviewed demonstrates a connection between these variables and income-related health inequalities (Joyce and Bambra, 2010; Kawachi, Subramanian, and Almeida-Filho, 2002; and Toronto Public Health, 2008). This methodology offers the ability to estimate the relative impact on health for a given space-time path within the study area.

The Toronto Census Metropolitan Area (CMA) has been chosen as the study area for this research. The literature reviewed connects income, education, ethnicity, and immigration to health inequalities, and the Toronto CMA is ideal for exploring these variables. Toronto is not only diverse, but also dynamic (Statistics Canada, 2015e). High levels of residential mobility and immigration play a role in defining these characteristics of this area. Toronto is also suitable for technical reasons. The city has a population large enough that data is aggregated and made available at the census tract level. Census tracts make analysis at local levels possible while avoiding most issues of data suppression and spurious statistical results that can occur due to the "small numbers" problem.

Improving our understanding of relationships that exist between the social determinants of health and income-related health inequalities over time and space has implications for policy. That is, if it is clear that certain segments of the population are in a position of greater risk for undesirable health outcomes, actions may be taken to mediate the negative influences that could lead to the negative health outcome. At present, it is challenging to connect health outcomes to causal factors and this can result in incorrect responses or even no response at all. The presented methodology has the potential to offer additional benefit if data on the health outcomes of individuals is made available. This data, along with the location history, could be used to identify historical health influences in space and time.

Collectively, this research addresses spatiotemporal variants of the spatial questions presented earlier: (i) "Can we verify that those with poor health outcomes have lived in one or more places when negative influences on health were present?", (ii) "Do certain space-time paths show greater health risks to individuals?", (iii) "Can we identify those on poor trajectories and take action to reduce the negative impact of their path?", and (iv)"Are there any patterns of 'path bundling' that are of interest?".

The remainder of this study is organized into five sections. The following section presents the Toronto CMA study area. The third section ties the study's approach to existing literature on income-related health inequalities and identifies issues geographers face when addressing problems in the field of health. Two common issues are addressed: the challenges that spatiotemporal data create, and the statistical and spatial analytical techniques that are applicable to the geographic analysis of health. The methodology follows. This section develops and operationalizes an approach to work with spatiotemporal problems where a spatiotemporal context needs to be accounted for. The required spatiotemporal contexts are also developed as part of the methodology and contexts for the study areas for each of the 2001 and 2006 Census periods are produced here. Results present the development and operationalization of the analytical techniques and use of these techniques to answer the research questions that have been specified for this work. The paper concludes with a discussion of the approach developed in this study and opportunities for future research.

Study Area

The Toronto Census Metropolitan Area (CMA) serves as the study area for this research. Toronto is located on the north-west shore of Lake Ontario in the province of Ontario, Canada. This CMA consists of 23 census subdivisions including Brampton, Mississauga, Pickering, Richmond Hill, and Toronto itself. Figure 1 presents the Toronto CMA and features the boundaries for census tracts and subdivisions provided for reference.

The CMA offers a rich environment to explore movement objects (individuals) within a dynamic and varied spatiotemporal context. The Toronto CMA is Canada's largest census metropolitan area with 5,113,149 residents as of the 2006 Census (Statistics Canada, 2015e). Toronto's population has grown 9.2% since the 2001 Census while population density for the area is 866.1 persons per square kilometre.

Of the approximately 4.8 million people living in the study area over the age of five, 2.1 million have relocated at some point since the previous census in 2001 (Statistics Canada, 2015e). Some of this mobility may be accounted for by the CMA's immigration numbers. As of 2006, almost half of the CMA population are immigrants (Statistics Canada, 2015e). With such a diverse group of people, this area has also become home to many visible minorities. Approximately 42 percent of the Toronto CMA has been classified as a visible minority (Statistics Canada, 2015e). The census also indicates that some of this mobility is not related to immigration and this may be attributed to major life events such as changes in employment or family status. Lastly, the median income for all census families in the Toronto CMA was $69,321 (Statistics Canada, 2015e). The variability and dynamic nature of the Toronto CMA provides for a spatiotemporal context well suited to the study of movement objects and the impact that place has on those objects. More specifically, the Toronto CMA is ideal to study the income-related health inequalities of individuals within the changing environment of this area.

Technical criteria have also been included in the decision to select the Toronto CMA as the study area. This research will explore phenomena that could cross census subdivision boundaries. Setting the study area to include the whole CMA allows for observation of a larger portion of relevant behaviour. The second technical reason for selecting the Toronto CMA is to remove potential edge effects. This problem occurs when a boundary setting results in incomplete data (characteristics on the other side of the boundary are unknown), which can lead to inaccurate results just inside the boundary. Lastly, the Toronto CMA has been selected so that characteristics within the city of Toronto may be compared against characteristics across the CMA.

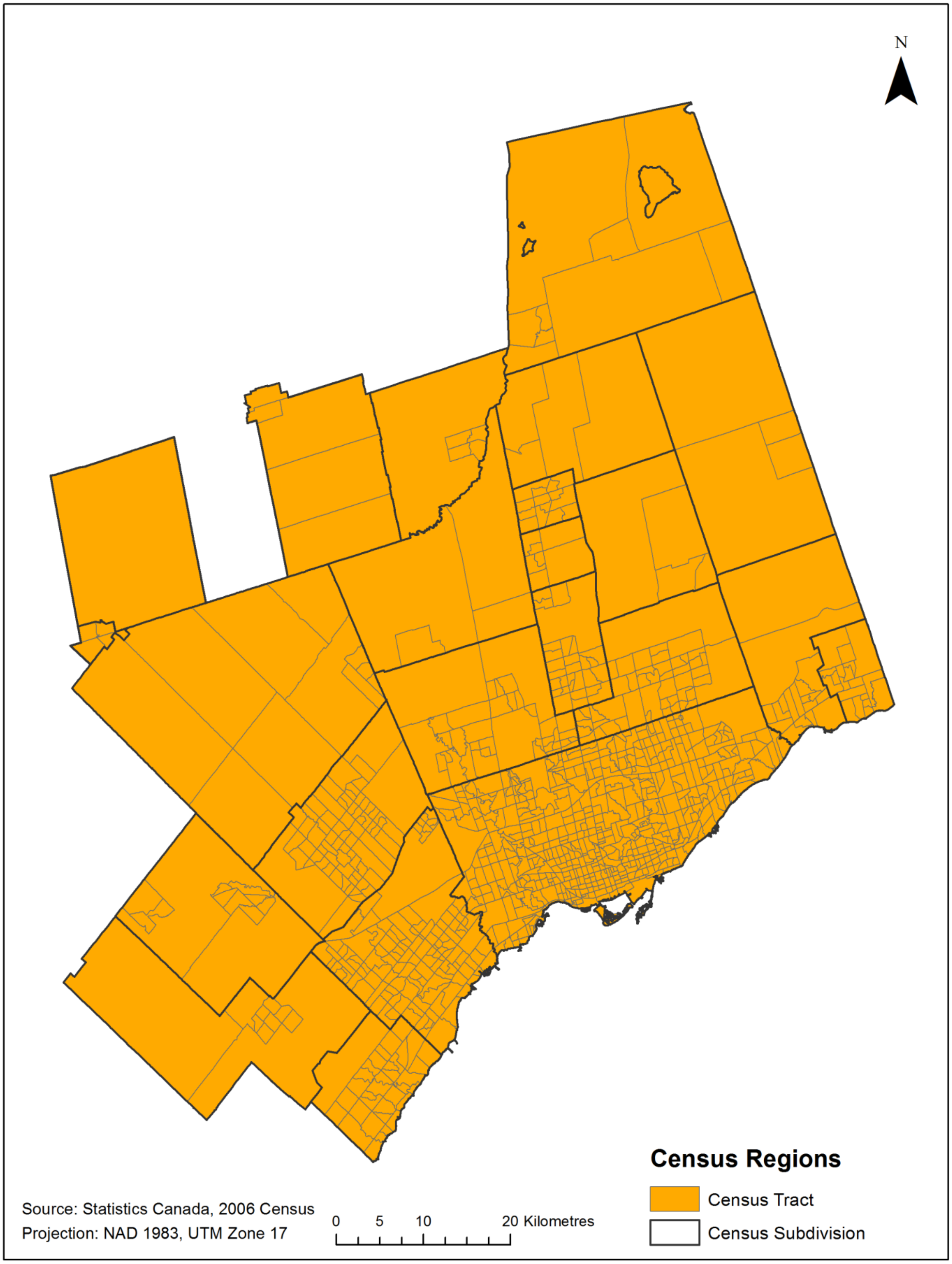

Figure 1: Study Area, Toronto Census Metropolitan Area (2006 Census Year Shown)

Literature Review

Developing a set of analytical techniques to explore social determinants of health and income-related inequalities is difficult. There are many theoretical and technical concepts that need to be applied in meeting this objective. Existing literature can provide guidance. Five sections review the concepts of "health inequalities", common problems found in health geography, approaches used for space-time analysis, spatial statistics, and map overlay techniques.

Health Inequalities

Existing literature provides a foundation of how to define "health inequalities", identification of those who are affected by "health inequalities", what variables may be used to develop a set of analytical techniques, and recognition that both space and time need to be accounted for.

Defining Health Inequalities

No single definition of health inequalities currently exists within the literature. Nonetheless, defining health inequalities is an important first step in the methodological development of this study. One problem is the common conflation of the closely related concept of "health inequities".

Kawachi et al. (2002, p. 647) provide a foundation for distinguishing between the two concepts. The authors define health inequalities as "the generic term used to designate differences, variations, and disparities in the health achievements of individuals and groups" and health inequities as "those inequalities in health that are deemed to be unfair or stemming from some form of injustice". As presented, health inequalities focus on the facts while health inequities attach opinions of fairness to these facts.

The City of Toronto's Toronto Public Health (2008, p. 27) recognizes health inequalities as "… differences in health status experienced by various individuals or groups in society. These can be the result of genetic and biological factors, choices made or by chance, but often they are because of unequal access to key factors that influence health like income, education, employment, and social supports". Health inequities are viewed as "differences which constitute a social injustice because the inequalities result from preventable causes – remediable, systematic barriers and forms of social exclusion" (The City of Toronto, Toronto Public Health, 2008, p. 3). Regardless of how a person or group arrives in their health situation, the first definition considers the facts alone, while the second definition extends this to consider whether the circumstance is just.

Other researchers have also commented on how to define health inequality and health inequity. Joyce and Bambra (2010) recognize Kawachi's definition and work within a context similar to Toronto Public Health's definition. Masseria, Hernández-Quevedo, and Allin (2010, p. 177) define health inequalities as "the systematic and avoidable differences in health outcomes such that poorer and/or more disadvantaged people are more likely to have illnesses and disabilities and shorter lives than those who are more affluent".

For the purposes of this research, and the requirements of quantitative analysis, this work will adopt the definition of health inequalities provided by Kawachi, Subramanian, and Almeida-Filho.

The Socioeconomic Gradient in Health

Marmot (2003) extends the discussion of health inequalities with the "social gradient of health". This idea suggests that everyone has a rank for their socioeconomic status and for their corresponding health status and that each of these are positively related. As there is always someone who is ranked better or worse than another individual, there is inequality across the range of the whole population. Problems of health inequalities are not reserved for those with the lowest socioeconomic rank.

The City of Toronto's Toronto Public Health (2008, p. 27) recognizes this gradient as "a step-wise distribution in health indicators that runs through society, where those that are poorest generally have the poorest health outcomes, and those in the middle generally experience poorer health outcomes than those in the higher social levels". Kawachi et al. (2002) suggest efforts should be focused to correct the inequalities of those at the lower end of the gradient. The Public Health Agency of Canada (2013) identifies the ability to have control over one's circumstances as a key factor in a person's health and they link this to socioeconomic status where those at the top of the gradient are better off than those who have a lessor status. Joyce and Bambra (2010) are also in agreement with Marmot.

The social gradient suggests that health inequalities do not divide people into groups of winners and losers, but rather into relative social positions (Marmot, 2003). Transferred to the task of developing a set of analytical techniques, this concept implies everyone has a position of social status that the techniques may account for. However, those positions must be linked to specific variables that measure health and socio-economic status.

Income and Health Inequalities

Existing literature supports the idea that there is a relationship between an individual's income and their health outcomes. Lofters, Slater, Shankardass, and Quinonez (2014) identify that health inequalities are income related. Reviewing 168 studies, Wilkinson (2006) finds populations that have better health are more equal in terms of income.

Joyce and Bambra (2010) have presented four explanations for health inequalities including "behavioural and cultural", "materialist", "psychosocial", and "lifecourse". Their second explanation, "materialist", connects income to a person's ability to acquire the goods and services needed to live in a modern society. The quality of these goods and services may also be connected to that same income; for example, a lower income may restrict a person to a poor diet and result in a lower standard of health. The ability to control circumstances identified as important to health by The Public Health Agency of Canada (2013) is also linked to income.

Kawachi et al. (2002) add to the discussion by distinguishing between absolute and relative incomes. While a person's income may be measured in a fixed amount of dollars, the status and purchasing power that this income provides is relative to others in the population and this determines rank. A certain absolute income may be relatively high in one area and low in another. Market forces may result in prices for items such as vehicles, property, or food that, relative to income, make one region more expensive to live in than others. The idea of relative income is important as it determines not just your rank compared to others in a region, but also the ability to satisfy the needs described in the Joyce and Bambra (2010) "materialist" explanation. Those with greater income resources will have the ability to better defend against negative health outcomes.

Each of these authors confirms the link between income and good health. What is also suggested by each contribution is that status can change. A person's income relative to others may change over time. The real economic value of a person's income may vary across time and space. This supports the need for spatiotemporal techniques that accommodate changing phenomena of interest along with the context the phenomena exist in.

Social Determinants of Health Inequalities

While income plays an important role in determining a person's level of health, it is not the only factor. Joyce and Bambra (2010) consider the role education plays in determining health. Their research finds that, like income, people with lower educations have poorer health. This should not be a surprise as education and income are also related; that is, a better education leads to better employment opportunities with a higher income and the result is better health. Masseria et al. (2010) also identify education status, along with measures of income and wealth, as being influential in determining a person's level of health.

Kawachi et al. (2002) point to gender, race, and culture as additional influences on a person's social status. Opportunities are not equally available to people of different genders, races, or cultures and this has an impact on health status. Toronto Public Health (2008) suggests race, immigration status, and education are also important health determinants. Evidence in the existing literature makes it clear that there is a complex system of social determinants of health and that these determinants should be carefully considered when researching health inequalities.

Some of the complexities that are present in the system of social determinants may be illustrated by the "healthy immigrant" and "ethnic density" effects. A recent study by Ng (2011) reveals the "healthy immigrant" effect to describe immigrants who often arrive to their new country healthier than the people born there. Over time the level of health of these immigrants can decrease. Ng notes that the initial level of health and the change in health can vary with an immigrant's country of origin. Pickett and Wilkinson (2008) describe the concept of the "ethnic density" effect and how it links to the Joyce and Bambra (2010) "psychosocial" explanation of health inequalities. Pickett and Wilkinson (2008) find that "ethnic density" effect relates to social capital and that there is non-materialistic value provided by a community of like people. Racism, discrimination, and stigmatization have been shown to have a negative effect on health. These authors present that when a minority group forms a community with like people a degree of protection may be achieved against these negative influences. This has the potential to improve health. The complexities found in social determinants of health indicate that relationships to health outcomes may be non-linear.

There is sufficient support for including education status, immigration status, and ethnicity related census variables (along with an income variable) in the development of spatiotemporal techniques.

Health Inequalities and Place

The literature also connects geographic location with health status. People do not exist in isolation from their surroundings. There is ongoing interest in how place impacts health. Kawachi et al. (2002) suggest that place provides compositional or contextual effects on health.

Compositional effects are the result of shared characteristics. For example, many women experience breast cancer. It does not matter where women are located, if women are present, there is good chance that breast cancer will also be present. In this way, patterns of health may be matched to the locations of where people with certain characteristics are located.

Alternatively, contextual effects are independent of the person or group. That is, contextual effects potentially impact everyone present. Examples of contextual effects include the impacts of pollution or policy decisions. Anyone who makes use of a local contaminated water source may be negatively impacted. Dr. John Snow's work to identify the Broad Street water pump as a transmitter of cholera (Longley, P. A., Goodchild, M. F., Magcuire, D. J., Rhind, D. W., 2011) is a famous example of the impact of a contextual effect. Similar people may also experience different health outcomes simply based on where they live and Kawachi et al. (2002) refer to this as contextual heterogeneity.

In a modern example, Yen and Kaplan (1998) consider the data provided by the Alameda County Study and consider the impact that place has on an individual's physical activity level over time. Their results show that those in areas of poverty had lower levels of physical activity than those who lived in more prosperous areas. These authors explain the difference in activity levels by the availability and quality of area resources such as parks, pools, playgrounds, and gardens. Levels of crime were also a factor.

The recognition of place as a factor in one's health outcomes is important and a geographic perspective should be included in the development of a set of analytical techniques to consider income-related health inequalities. Places also change over time and how people have been impacted by a place in the past can be very different than how they may be affected in the future.

Health Inequalities and Life Course

Marmot (2003) recognizes that a person's health status is not just dependent on one's present, but also on their past. Early life circumstances and decisions lead to the present situation. Joyce and Bambra (2010) expand on Marmot by noting that there is also a cumulative effect of positive or negative influences that occur over time. That is, someone who has been poor for their whole life and has a low quality diet can be expected to be in poorer health than someone who has had more favorable circumstances until recent times and is only now subject to a low quality diet. Even though these two people are both presently poor and subject to a poor diet, the first person has the greater burden as they have a greater accumulation of the negative effects of their poor diet. Joyce and Bambra (2010) continue on to express that health inequalities are "a result of inequalities in the accumulation of social, psychological, and biological advantages and disadvantages over time" (p. 22).

Kawachi et al. (2002) also discuss the life course concept and the related latency, pathway, and cumulative effects. Latency effects, as Marmot observed, are those impacts that materialize in the present due to circumstances or decisions that occurred in the past. Pathway effects reflect the impact that place and environment has on the individual. Lastly, cumulative effects are those that have been discussed already by Joyce and Bambra.

The literature demonstrates attributes associated with an individual and their health may change over time, but may also change as a result of residential mobility along with the fact that places are dynamic. A set of analytical techniques to better understand income-related health inequalities must take into account all of these potential changes and account for both the individual and their contextual environment.

Common Problems in Health Geography

Certain characteristics of spatiotemporal problems identified in the literature create potential challenge for studies of health geography. Geographers usually work with phenomena within one time period or compare phenomena between multiple periods. Creating analytical techniques to explore the relationship between social determinants of health and health outcomes requires that those techniques recognize activity that spans several time periods. These techniques need to account for latency, residential mobility, and accommodate the uncertain geographic context problem. Issues of ecological fallacy and the modifiable aerial unit problem also need to be managed. The literature explored in this section will provide a foundation for dealing with these problems.

Latency

Kawachi et al. (2002) recognize the role that latency plays in health problems. These authors present that latent effects occur when some causal event or action takes place and the outcome is delayed until some future point in time; results are not immediate. This creates a challenge in matching causal events to outcomes. Kwan (2012a) also points out that there may be multiple causal events and outcomes may be the result of cumulative effects of repeat exposures. Any research developing a set of analytical techniques to explore social determinants of health in an effort to estimate health outcomes will need to capture data at multiple points in time prior to the resulting outcome.

Cromley and McLafferty (2012) show that geographic information is usually only collected at a person's home address at the time that they are diagnosed with their particular health issue. This effectively ignores that latency exists. Gatrell and Löytönen (1998) argue that inferring anything about a causal relationship based on spatial patterns at the time of diagnosis, instead of spatial patterns at the time of exposure, can lead to inaccurate conclusions. It is acknowledged that part of this problem is related to a lack of data and that results would be better if data on a person were available for multiple points in time. Gatrell and Löytönen (1998) do recognize that there would be considerable effort and cost to collect and manage the relevant data at this level of detail.

Jacquez, Greiling, and Kaufmann (2005) provide a different perspective on this issue. While they agree that latency is an issue that is often ignored, they also find the tools of geographic information systems (gis) to be lacking in capabilities to manage and analyze information and problems across multiple time periods. Even if data was readily available, it would be a challenge to make use of them. Modern maps, whether paper-based or electronic, are two-dimensional and static. Efficient and effective ways to present space-time information are still being researched. A focus on locations, and not people, is also expressed by Jacquez, Greiling, and Kaufmann to be part of the problem. While latency is an issue, the data and tools required to incorporate the latency effect into analytical techniques is lacking.

Migration Bias and Demographic Cycle

Migration bias is a misclassification error that can occur due to the incorrect or incomplete recording of individual residential location(s) during a study. Beale, Abellan, Hodgson, and Jarup (2008) recognize that this error has the potential to associate an individual with either a causal effect or an outcome that they were not actually subject to. People may relocate during the study period. As many studies have only recorded one address for individuals, those studies that span a long period of time to allow for latency effects can be subject to this type of error. Cromley and McLafferty (2012) share this concern and note that movers are subject to cumulative effects as presented by Kawachi et al. (2002). When combined with the problem of latency and only recording a person's address at the time of diagnosis, migration and residential mobility has the potential to complicate the analysis of long-term health issues such as those related to health inequalities. Gatrell and Löytönen (1998) expand on the issue of only recording a person's address once by noting that people in industrialized countries relocate frequently and in large numbers. When only the address at the time of diagnosis is recorded, the historical context of place is lost.

Is mobility an issue that needs to be accounted for? Tong (2000) provides statistics on Australian mobility for the 1986 to 1991 period. At that time, approximately forty-one percent of Australians changed their home location. Australian males were expected to change their home address 11.1 times over their life, while Australian females were expected to move 11.5 times. In the city of Toronto, as of the 2006 Statistics Canada Census, immigrants from outside of Canada accounted for half of Toronto's population (City of Toronto, 2014). 2006 Census data also reveal a very mobile population with half of those over the age of five relocating from their 2001 location; of these, sixty-three percent moved within the city (City of Toronto, 2014). Given the numbers, ignoring residential mobility could have a significant impact on research where knowledge of an individual's location is important.

Why do people change their place of residence? Entwisle (2007) suggests that people move because they have the choice to do so and they may take advantage of that choice due to some motivation. People may move to be closer to certain ethnic or cultural environments, they may move to take advantage of schools that appear to be better than what is currently available to a child, or they may select a new location for better health outcomes. Entwisle presents these examples but also warns that not everyone has choice and while some can move out of less desirable areas, others may lack the means to do so and be forced to stay.

Geist and McManus (2008) link a person's mobility to their life course stage. The life course model is a more current version of the classic family life-cycle model. Under these models people are expected to proceed through certain stages of life at a certain age and in a certain order. Children are expected to be in school until their late teens. This may be followed by post-secondary education. During the middle twenties people are expected to start their careers and get married. Adults typically have children in the late twenties and early thirties. People retire from work at sixty-five and may require late-in-life health-related services. The life course model relaxes many of the assumptions of the family life-cycle model as people today complete the "tasks" of life at later times or even out of order – a more individual-focused model is required.

Geist and McManus (2008) present that certain residential requirements exist at various stages of life and that these requirements may motivate people to relocate. After university or college, young adults move out of their parent's home. Young adults will find a starter home when they get married or move-in together. Children require additional space to sleep and play, and as Entwisle described, the children may motivate locating near certain schools. As people age, less space is required while the need for medical services may increase until death, empty-nesting and retirement/nursing home requirements may be a factor for a change in location.

Entwisle (2007) has described that, places are not equal on all dimensions and that people may be attracted to certain features of individual places. Life course stage will bias what features are of interest to people and it can be expected that similar people may cluster together to leverage a public resource in satisfying a common need. This is the case with seeking better schools for children. It is also the case of locating the elderly into retirement or nursing homes. This bias may have interesting consequences. If the elderly are collected in certain areas to take advantage of nursing home services and then die there, how does this impact mortality rates? It can be expected that those areas that externalize the elderly will benefit by a lower mortality rate, while those that take in the elderly will suffer a higher mortality rate. When considering mortality rates and health inequalities, care needs to be taken with biases such as these. Biases in the regional characteristics of place may be a reflection of regional specialization to satisfy common group needs at certain points in the life course.

This literature provides important guidance on the design of spatiotemporal analytical techniques. Along with insight into why people relocate, and the significance of the behaviour, it is necessary to create techniques that can track individual movements. Failure to properly locate individuals in space and time has the potential to create issues of migration bias and the resulting misspecification errors.

Uncertain Geographic Context Problem

Assumptions about place are also a problem for researchers. Jacquez et al. (2005) present that people are mobile and there are usually differences in where a health-related exposure takes place, where it is recognized to have occurred, and where it is recorded as data to have occurred. To illustrate, assume a commuter is exposed to a common illness during their morning commute on public transit. The next day at work, the commuter doesn't feel well and leaves early to visit their doctor and the doctor records the event with the commuter's home address. This scenario illustrates how caution needs to be taken when using home addresses as the location for where health events occur. The way in which data are collected creates a built-in potential for error.

Kwan (2012a) formalizes and expands on this issue with what she describes as the "uncertain geographic context problem". In many studies, standard area-based units such as census tracts are used. These units are used to not only locate a person of interest but also to define the context of spatial influence. Kwan argues that it is unlikely that the standard administrative area will match the true and unknown geographic context that a person is actually subject to. As people are mobile, and their daily lives take them in different directions, activity space will also vary person to person – and so will the influences that various spaces have over people. How much time is spent outside of the area of residence? What other areas do people interact with, and for how long are they doing so? Collecting data and identifying true activity spaces across multiple areas can be challenging, especially if these areas are not contiguous.

Kwan (2012a) also notes that both place and people may change over time and that these changes impact the geographic context. Changes in the physical space people live in may occur and introduce positive or negative influences. The introduction of a landfill would be an example of a negative change in the context of physical space. People may also change. As populations evolve due to births, deaths, aging, and migration, the social context of place will also change. For example, certain behaviours that were unacceptable in the past may now be acceptable in the present, and these may have positive or negative impacts on health. In other work, Kwan (2012b) discusses the influence of relocation from the individual's perspective. If a person moves to a new residence then they now have a new geographic context defined by both the physical and social characteristics of their new activity space. It should be noted that this mover will also provide a new influence to their new area and that influences are bidirectional.

Kwan (2012b) also recognizes that part of this uncertain geographic context problem is data related and is optimistic about the opportunities offered by location-aware technologies that could support research allowing for the better definition of a person's true activity space. Along with the potential for greater insight into an individual's movements comes the need for the protection of the same individual's privacy. Kwan (2012a) has illustrated that assumptions researchers make about the individual's true geographical context have the potential to lead to incorrect results and work needs to be done to manage the uncertain geographic context problem.

The uncertain geographic context problem reveals a limitation to the results that might be provided by spatiotemporal analytical techniques. Results will only be as accurate as the assumptions of the data used within the techniques. Aggregate-level data likely won't match to any one person. Operational limitations are also evident and a balance needs to be achieved between the resolution of data collected and the population's tolerance for surveillance. That is, to improve the results of these techniques, data could be collected on individuals by the minute, hour, day, or month – but would people be willing to participate? Technology also has limitations and collecting large volumes of data, and making use of it, could provide additional challenges.

Ecological Fallacy and Modifiable Aerial Unit Problem (MAUP)

While the temporal dimension of health-related problems creates some challenges for the health geographer, there are some issues that geographers must be aware of regardless of the types of problems they are working on. Freisthler, Lery, Gruenewald, and Chow (2006) caution against "ecological fallacy" as do Tong (2000) and Beale et al. (2008). Ecological fallacy is the error of assuming that the characteristics of a group or aggregate can be applied to an individual within that group. It is possible that individuals may deviate from the trend of the group's aggregate indicator.

Kwan (2012a) addresses the modifiable aerial unit problem as another issue that geographers should be cautious of when working with area-based data. According to Beale et al. (2008) the scale that is chosen for an analysis can influence the results obtained in a study. At different scales, data are aggregated differently and results can often vary. Cromley and McLafferty (2012) add that specifying different aerial units at the same scale will also yield different results.

Approaches for Space-Time Analysis

It is clear that there are challenges when working with spatiotemporal data and related problems. Available literature may be leveraged to provide guidance for the development of spatiotemporal analytical techniques, the operationalization of those techniques, and options for the visualization of space-time information.

Developing Spatiotemporal Analytical Techniques

There is variety in the types of spatiotemporal problems that a geographer might be faced with. On the simpler side, it may be necessary to compare the same phenomena in two different time periods. This might be for the comparison of satellite imagery or census data in choropleth maps. In more complex situations, a model might need to represent phenomena or relationships across time. For situations like these, Hornsby and Egenhofer (2002) offer their "geospatial lifeline" model. In this model, movements are sampled and data for location, time, and an object identifier (a person's social insurance number, for example) are recorded. This sampling would be similar to the location sampling that current location-aware smart phones might do. With this information, location may be represented as x and y coordinates on a horizontal plane while time is represented on the vertical plane as z. As points are connected by lines, movement in space-time may be considered. A line parallel to the z-axis would represent an object stationary in space while a near horizontal line would represent an object moving across space at high velocity. Cromley and McLafferty (2012) describe a similar system and refer to it as “space-time-path”. Hägerstrand (1970) presented a much earlier version of this model to account for activities in space and time with his work providing significant guidance for those who followed. Each of these authors is presenting a variation of the space-time cube concept.

Jacquez et al. (2005) developed their own “Space-Time Intelligence System” (STIS) software based on a variation of the space-time cube. As part of their work, a linear “object chain” movement model has been created. This model is constructed by using stacked layers representing the space dimensions at a given point in time; that is, layer 1 (on the bottom) represents the physical space at time index 1, layer two (second from bottom) would represent time index 2, and so on. The object(s) in motion are then located on each layer based on what its location was at that time. The object chain is completed by connecting the object’s points through time. This is not unlike the Hornsby and Egenhofer “geospatial lifeline” approach. The advantage of the Jacquez, Greiling, and Kaufmann approach is that it has the potential to account for the spatiotemporal context that movements take place in; this is one of the shortcomings of the original model.

Andrienko et al. (2011) have also presented a methodology for dealing with spatiotemporal problems. Their offering expands on the classic space-time cube approach to account for the spatiotemporal context of objects in motion. The methodology offered first considers spatiotemporal phenomena as an event. An event includes a movement object (m), a spatiotemporal context (c), and a relationship (R) between the movement object and the spatiotemporal context. An example of an event could include an individual (m) walking on (R) a sidewalk (c) in summer. If a similar event is considered in a winter month, that same individual (m) might be walking on (R) the same icy sidewalk (c). While the location is the same, changes in time can change the context. This individual is exposed to a risk in winter that is not present during warmer months. If the individual is changed to two different people, the same path offers different risk levels to these people based on changes in spatiotemporal context. Once events are defined they may then be analyzed and the impact of the event can be assessed. This approach is important as it offers a method for including the impact of place (the spatiotemporal context) as individuals proceed along their space-time path.

Operationalizing Spatiotemporal Models

While there are now models that allow geographers to approach complex spatiotemporal problems, operationalizing these theories still has challenges. Data collection and data management, along with a lack of suitable software tools, has attracted the attention of many conducting research in this field. It is only recently that researchers have had the opportunity to take advantage of location-aware devices to support the collection of data. Kwan (2012b) presents an application of these devices in some of her recent research. Lundblad et al. (2009) have also used global positioning system (gps) technology to support their research. In their work, the traffic movement of ships on the ocean is tracked and recorded so that a space-time path for each ship may be generated.

On issues related to software, Jacquez et al. (2005) have expressed concern over the ability of geographic information systems (gis) to support the research, analysis, and management of spatiotemporal problems. These authors have found gis to be lacking in the ability to manage both spatial and temporal information at the same time. They note that, despite having a theoretical foundation to leverage, software tools have not been built to take advantage of the theoretical models. To resolve this issue, these authors have created their own software for research; the "Space-Time Intelligence System" (STIS) software identified earlier. New features from theory have been incorporated into the STIS tool and this avoids the lack of functionality associated with commercially-available software products. These features are not generally available in today's popular gis products and this leaves others with the challenge of finding, or authoring, suitable tools of their own to tackle space-time problems.

Visualizing Space-Time Information

Depending on the complexity of spatiotemporal models, creating effective visualizations can be challenging. Kraak and Ormeling (2003) discuss a variety of techniques for visualizing space-time phenomena. Many of these solutions are used to present change that has occurred between two time periods. Kraak and Ormeling use the classic Minard map displaying Napoleon's troop movements in Europe in 1812 to illustrate how flow patterns may be used as one method to show change on one map. The use of multiple maps is also presented, and these may be static, for the user to review in a series, or animated using current technologies.

One of the limitations to maps is their two-dimensional nature. This creates issues when dealing with more complex spatiotemporal models. Andrienko, Andrienko, Demsar, Dransch, Dykes, Fabrikant, Jern, Kraak, Shumann, and Tominski (2010) acknowledge limitations in visualizing space-time data and models. While the space-time cube is suitable for conceptualizing ideas or for low volumes of data, it doesn't print well and the three dimensional space may create representations that are difficult to interpret. When large amounts of data need to be mapped, such as space-time paths in a cube, the resulting visualizations may be illegible. The classic space-time cube also fails to account for changes in spatiotemporal context.

With the current state of visualization technology, along with the limitations in capabilities of geographic information systems to model time data, a set of analytical techniques that generates visual output might not be possible at this time. The priority for this research will be a set of techniques that can provide results, even if it is only in numeric or tabular form.

Spatial Statistics for Spatial Analysis

One of the challenges of spatiotemporal modelling that has previously been identified is the inclusion of spatiotemporal context. The use of spatial statistics offers the opportunity to detect spatial clusters of phenomena of interest at a particular point in time. Employing spatial statistics to identify areas of potential risk to one's health outcomes could be achieved using relevant variables with spatial statistic techniques. This may then be used to inform spatiotemporal context. The literature discussed in this section provides a foundation in the concepts and application of these statistical methods.

Spatial Autocorrelation

Freisthler et al. (2006) point to the tools of exploratory spatial data analysis to analyze data for potentially significant outliers, clusters, or correlations. Any pattern that may be identified can then be used to develop hypotheses and tested with confirmatory techniques. Patterns of "spatial autocorrelation" may be detected through the use of various spatial statistical techniques. Jerrett, Gale, and Kontgis (2010, p. 1307) connect the concept of spatial autocorrelation to Tobler's first law of geography that states that "everything is related to everything else, but near things are more related than distant things". Expanding on this foundation, they clarify that characteristics of objects or places in space would be expected to be more similar when they are closer together than those that are farther apart. The authors provide an example using air quality where one would expect the air to be more similar at locations close together compared to those that are not.

Jerrett et al. (2010) present that tests for spatial autocorrelation may be conducted to detect a global pattern of spatial autocorrelation across an entire study area or that techniques may be used to explore only localized clustering in a subset of the study area. Semple, Cudnik, Sayre, Keseg, Warden, and Sasson (2013) identify many of the spatial techniques available for both global and local spatial analysis. These include: the Moran's I, Geary C, and the General G for global tests and the Local Moran's I and the Local G_i* for local tests. Samarasundera, Walsh, Cheng, Koenig, Jattansingh, Dawe, and Soljak (2012) add that when statistically significant clusters are identified that they are often referred to as "hotspots". These hotspots have been identified using appropriate statistical procedures and may differ from what the human eye detects as a clustering pattern on other types of maps. Samarasundera, et al. (2012) caution that results from spatial analysis should be trusted over the unassisted human detection of clusters as humans often do a poor job of detecting spatial patterns on their own. Cromley and McLafferty (2012) expand on the concept of cluster detection by more specifically defining a cluster as "a region that has unusually high counts or rates of disease, that is, a local concentration of high values". Rogerson and Yamada (2009) provide technical insight into the global and local statistics. A review of the various techniques for testing for spatial autocorrelation, including equations and interpretation of results, is provided in Appendix A; spatial weights matrices are also included in this review.

Spatial Empirical Bayes and the Small Unit Problem

Rezaeian, Dunn, St. Leger, and Appleby (2004) identify the problem of small areas that can occur if the number of events in aerial units is too few. Spatial statistics, as with aspatial statistical methods, require a sufficiently large sample size to overcome the potential for spurious results. There is the possibility that a variable's value has occurred at random and may not be representative of the area with a low number of cases. Beale et al. (2008) suggest using a smoothing technique to correct for this issue. They offer three methods that are available in the GeoDa software application: empirical Bayes, a spatial window average, and spatial empirical Bayes. Semple et al. (2013) clarify the differences in the methods. The standard empirical Bayes method will calculate an adjustment based on the whole study area, while the spatial empirical Bayes method will calculate the correction using surrounding regions such as neighbours.

Expert Knowledge

A valuable step in the research process that can often be overlooked is the verification of results against the knowledge of domain experts. Semple et al. (2013) have demonstrated how they have leveraged the expert knowledge of emergency response teams to confirm that their estimation of spatial clusters match what is seen "on the ground". Where possible, and if available, the knowledge of such domain experts should be used to help ensure results of studies are as reliable as possible.

Map Overlay Techniques for Spatial Analysis

While the use of spatial statistics has become a widely accepted method for detecting spatial clusters, different techniques are required to identify cases of colocation for multiple variables. Semple et al. (2013) have leveraged overlay techniques to identify where "high – high" local Moran's I clusters of one map layer are collocated with "low – low" local Moran's I clusters of a second layer. Where "high – high" clusters of the first layer were found to overlap with “low – low” clusters on the second layer the desired colocation would be identified. Jerrett et al. (2010) also use map overlay techniques to create new composite layers to support their analysis in the area of environmental risk to human populations. By using the overlay technique, the colocations of areas with high populations and areas with high incidences of environmental risks, such as flooding, were detected.

Samarasundera et al. (2012) also discuss the use of overlay techniques and summarize many of the available overlay operations. As is typically employed with Venn diagrams, union and intersection operations are available to allow for the selective combination of input layers when creating new composite layers. A union composite would include data for phenomenon that appears in either the first or second layer, but do not necessarily coincide, while an intersection composite would only include areas where that phenomenon appears in both of the source layers and they do coincide. Map algebra, typically used for raster-based layers, may also be performed to algebraically calculate a new value for the cell in the new layer from the corresponding cells in the source layers. Every cell in the new raster layer is determined by this algebraic formula. Cromley and McLafferty (2012) add that Boolean operators of AND and OR may also be used; these should provide results similar to union and intersection functions respectively.

Heywood, Cornelius, and Carver (2006) discuss how raster cells may be recoded to reclassify information on a map layer; for example, features (cells) representing risk may be recoded with an integer value that reflects the relative risk to a local human population. In this example, cells on a layer consisting of lakes or rivers might be recoded with “0” where there is no water and given a value of “1” where lakes or rivers are present. On a second layer, a road network may be recoded with “0” where there are no roads and “2” where there are roads. Applying map addition as presented by Heywood et al. (2006), a new layer would be created where “0” represented where there were neither road or water, “1” represents where there is only water, “2” represents where there are only roads, and “3” represents where there are both roads and water.

Map multiplication may also be considered if a fourth data layer for population is added to the example. Assume areas (cells) of population counts over 100 are coded as “100” and cells with fewer than 100 people are coded as “0”. Multiplying the population layer by the previously developed risk layer will result in only areas that have both a population greater than 100 and risk remaining relevant. Any “0” value on either layer will result in a “0” on the new composite layer indicating either “no risk” and/or “no population over 100 people”. Any non-zero number will be a multiple of 100 and represent where there are either colocations of at least 100 people and water risk (“100”), colocations of at least 100 people and road risk (“200”), or colocations of at least 100 people and both water and road risk (“300”). It is important to recognize that Heywood et al. (2006) have simply provided a more sophisticated method of performing unions and intersections across map layers by using map algebra. As demonstrated by the literature, map overlay techniques offer very useful methods to identify areas of colocation based on some specified criteria. It should be recognized that this concept may also be adapted for use with aerial regions in vector data.

Synthesis of Knowledge for the Development of a Methodology

Developing analytical techniques for income-related health inequalities is not an easy task and ideas presented in the literature may be synthesized to inform this task. This section has provided a working definition for income-related health inequalities, has identified what variables may be used to develop analytical techniques focused on health inequalities, has confirmed that public health problems are often spatiotemporal in nature, and has identified approaches that may be leveraged to develop techniques for these types of problems. These contributing ideas are considered and applied to the design and implementation of a new analytical technique.

Kawachi et al. (2002, p. 647) provided the working definition of “health inequalities” as “the generic term used to designate differences, variations, and disparities in the health achievements of individuals and groups”. This definition is compatible with the development of a new analytical technique as its perspective is restricted to facts that may be represented by data supplied to a model. Other literature focusing on health inequalities has provided insight into the selection of variables that should be targeted as inputs for such the new technique. As suggested by Toronto Public Health (2008), these variables include those related to income, education, visible minority, and immigration status.

The literature identifies that problems related to health inequalities are likely to be spatiotemporal in nature. Marmot (2003) identifies the issue of latency in such problems, as do Joyce and Bambra (2010) and Kwan (2012a). Any technique developed must be able to consider phenomena across multiple time periods. With latency there exists the opportunity for residential mobility. Tong (2000) confirms that people in industrialized regions are likely to relocate many times over the course of their lives. Entwisle (2007) and Geist and McManus (2008) also find that individuals are mobile and provide insight on why people relocate. Kawachi et al. (2002) recognize that place provides context and that this context can change over space and time. Questions related to health inequalities are indeed spatiotemporal in nature.

Each of the contributions made by these authors informs on the design requirements for any technique developed to explore income-related health inequalities. The analytical technique needs to have the capability to: 1) track the movements of individuals in space and time, 2) develop spatiotemporal context using census variables related to income, education, visible minority, and immigration status, 3) relate individual movements to spatiotemporal context, and 4) provide some evaluation of the impact on health of an individual's movements over space and time.

Additional literature has provided guidance on how the needs of a new analytical technique to explore health inequalities may be provided for. Andrienko et al. (2011) have provided an overall framework for spatiotemporal problems that need to account for both movement objects and spatiotemporal contexts. Their concept models all activities as spatiotemporal events that include the dimensions of a movement objects (m), a spatiotemporal context (c), and a relationship (R).

Within the framework provided by Andrienko et al. (2011), the movement object (m) may be represented by the Hornsby and Egenhofer (2002) "geospatial lifeline" or as a space-time path. This approach leverages the strengths of the space-time cube. Semple et al. (2013) offer the strategy of combining tests for spatial autocorrelation with map overlay techniques and this tactic may be employed to provide for the spatiotemporal context (c) at a particular point in time. Data reclassification techniques presented by Heywood et al. (2006) also support the development of this spatiotemporal context.

The literature reviewed has provided the conceptual ideas for both the design of a new analytical technique and the implementation of that technique. With these ideas in place, the focus of this study now shifts to operationalizing a new analytical technique to explore problems of health inequalities. The following methodology reviews the application of the design concepts and implementation ideas to operationalize the new technique and then apply the technique to answer the research questions specified in the "Introduction" section of this research.

Methodology: Application of a New Analytical Technique

This study's methodology decomposes spatiotemporal phenomena into three elements. The first element is spatiotemporal context (e.g., health-related variables associated with census tracts) and these are defined for a given place at a specified time. The spatiotemporal context defines what a movement object will be exposed to while in the presence of the spatiotemporal context. A census variable for income may be used to define an area as either a high or low income region at that point in time. The second element is a movement object (e.g., people) and this is typically the entity being studied. People in the study area of the Toronto CMA are the movement objects for this research; they are mobile and will be affected by their local context. Third, spatiotemporal relationships provide the last element used to define spatiotemporal phenomena. These relationships are used to identify if a given movement object should be subject to the contextual effects of place. If a person (a movement object) is present in a particular area then they will be subject to that area's context (positive effect of local wealth or negative effect of poverty) for that point in time; if a person is not present, then they cannot be affected by that local context.

Spatiotemporal context used in the new technique are developed using a range of socio-economic variables with map overlay and map algebra techniques presented by Semple et al. (2013) and Heywood et al. (2006) respectively. A composite value for each context is developed by combining these variables and is stored in a relational database. This database also stores the space-time path data for movement objects, and establishes the relationships between the two. Without available individual data, this study generates a sample set of movement data to illustrate the effectiveness of the analytical technique. SQL queries are developed that address this study's central research questions.

Developing the Spatiotemporal Context for the Analytical Technique

The Canadian census is used to establish contexts for each space-time object within the scope of this study. Tests for spatial autocorrelation identify local spatial clusters that contribute to the specification of spatiotemporal context. Individual contexts are then consolidated for a given space-time object and each consolidated context is associated with a score reflective of the impact of that spatiotemporal context on health.

Preparation of Data

Census profile data at the census tract level from the 2001 and 2006 Canadian Censuses for the Toronto CMA form the foundation of this study's methodology (Statistics Canada, 2015a and Statistics Canada, 2015b). Census data has been downloaded for education status, immigration status, income status, and visible minority status for each of the two census periods. These variables have been selected, in part, based on the guidance of the literature. In addition to the literature, Toronto Public Health has completed its own research (The City of Toronto, Toronto Public Health, 2008) on the City of Toronto's demographic profile. This has been presented in the 2008 "Unequal City" report and these results establish that there is a connection between low income and each of the variables of education status, immigration status, and visible minority status for Toronto. Each of these variables has a role in determining health outcomes for this region. The census tract level of geography has been selected for this study as this level is the compromise of data too aggregated at the census subdivision level to be useful and data suppression or small-area problems that occur at the census dissemination area level of data.

The "R" software package was used to consolidate the 2001 and 2006 Census data. Not all census tracts in a given year have complete data and those that are incomplete have been removed. Census data is also provided with values rounded to the nearest increment of 5 for most count data and the totals for these variables do not reconcile with the data provided. To correct for this difference, new "control" totals have been calculated and recorded as new variables.

One new variable each has been created for the education, immigration, and visible minority profiles. Statistics Canada separates income statistics for those who belong to a "family" and those who live on their own. To account for this separation, two variables are explored for income including one for the "family" statistics and one for the statistics for those who live on their own.

The new education variable is calculated as the percentage of the population in a census tract that has earned less than a bachelor degree or better and this variable covers all age ranges (Statistics Canada provides original 2006 variables by age range). For immigration, the percentage of residents in a census tract that are immigrants has been calculated. Visible Minorities are also represented by a new variable calculated as the percentage of visible minorities in a census tract. For the two income variables, Statistics Canada provides an existing variable that presents the percentage of population in the census tract that is below the low income cut-off (LICO). The result is five variables for the exploration of four census profiles for the Toronto CMA for a given census year.

Testing for Spatial Autocorrelation and Identifying Clustering

Calculation of global or local Moran's I depend on the specification of weights matrices that specify the relationship between locations of interest. The open source software product GeoDa has been used to create the required matrices. For this research, a first-order queen configuration was used. As the configuration of census tract areas changes over time, one spatial weights matrix file has been created for each of the 2001 and 2006 censuses. A global Moran's I analysis was conducted to determine if a non-random spatial pattern existed for each variable throughout the study area.

The coefficient and statistical significance of results of the global Moran's I tests are summarized in Table 1. This table presents the Moran's "I" coefficient for each of the five study variables in each of the two census years included for this study. The 10 scatterplots generated for these tests are available for review in Appendix B. Each of the variables tested with the global Moran's I test indicates some level of spatial autocorrelation. With a score of 0.7 and above, the "percentage of low education", "percentage of immigrants", and "percentage of visible minorities" indicate a high degree of spatial autocorrelation for the study area and localized clusters are likely to be found. While the coefficients for the low income variables are not as high as for the other variables, there is still an indication of spatial autocorrelation at the global level; some local clustering should be found for this variable too.

Table 1: Summary of Global Moran's I Tests – Moran's I Coefficients

Variable	2001 Moran's I Coefficient	2006 Moran's I Coefficient
Percentage of Low Education	0.743194*	0.760504*
Percentage of Immigrants	0.746764*	0.750246*
Percentage of Low Income (Family)	0.498248*	0.481791*
Percentage of Low Income (Individual)	0.373748*	0.440235*
Percentage of Visible Minorities	0.709936*	0.733352*
*Significant at 1 percent level		

An indication of spatial clustering was provided by the calculation of the global Moran's I for each contextual variable for 2001 and 2006 in the study area. The local Moran's I was calculated for each variable in each census tract in the Toronto CMA for each of the two study periods. The equation for the local Moran's I is presented below.

$$I_i = \frac{m(y_i - \bar{y})}{\sum_j (y_j - \bar{y})^2} \sum_j w_{ij}(y_j - \bar{y})$$

Results calculated for the local Moran's I are typically within a range of -1 to 1. Positive autocorrelation is indicated by positive values, while negative autocorrelation is indicated by negative values. Results around zero indicate an absence of pattern. Positive autocorrelation occurs when similar objects are clustered together while negative autocorrelation occurs when dissimilar objects cluster together.

The results calculated for the local Moran's I test have been used to classify each census tract in the study area with one of four types of clustering, if the results are significant. A "high – high" cluster exists where a census tract with a high variable value is surrounded by other census tracts with high values in the same variable. Likewise, a "low – low" cluster exists where a census tract with a low variable value is surrounded by other census tracts with low values in that same variable. These two types of clusters are examples of positive autocorrelation. Two additional clustering classifications of "low – high" and "high – low" occur when the census tract of interest is surrounded by dissimilar census tracts and these are examples of negative autocorrelation. Appendix C summarizes the results of the local Moran's I tests for each of the five variables tested in both 2001 and 2006. Ten maps are presented in this appendix that present the patterns of clustering that have been detected for each of the variables in each of their respective years. The patterns that occur are also considered.

Consolidating Clustering Patterns across Variables

Local Moran's I results were consolidated using the map overlay techniques described by Semple et al. (2013) and the map algebra approaches presented by Heywood et al. (2006). Instead of working with map layers as presented in the literature, variables in the attribute data were used to algebraically achieve similar results. That is, while the literature presents values being added, subtracted, multiplied or otherwise manipulated across map layers, the approach used here performs those same operations across variables (cells) in a spreadsheet.

Before the consolidation of the clustering patterns of each of the variables could take place, the two low income variables for individuals and "families" also needed to be consolidated. This consolidation has been completed based on the specifications defined in the income variable consolidation matrix presented in Table 2.

Table 2: Consolidation Matrix for Income Variables

Individual Income \ Family Income	1 High – High	2 Low – Low	3 Low – High	4 High – Low	0 Not Significant
1: High – High	1 (High – High)	1 (High – High)	1 (High – High)	1 (High – High)	1 (High – High)
2: Low – Low	2 (Low – Low)	2 (Low – Low)	2 (Low – Low)	2 (Low – Low)	2 (Low – Low)
3: Low – High	1 (High – High)	2 (Low – Low)	0 (Not Significant)	0 (Not Significant)	0 (Not Significant)
4: High – Low	1 (High – High)	2 (Low – Low)	0 (Not Significant)	0 (Not Significant)	0 (Not Significant)
0: Not Significant	1 (High – High)	2 (Low – Low)	0 (Not Significant)	0 (Not Significant)	0 (Not Significant)

The consolidation matrix presented in Table 2 has been constructed to give preference to the "High – High", "Low – Low", and "Not Significant" classes. When a tie exists in the matrix between either a "High – High" class and a "Not Significant" class or a "Low – Low" class and a "Not Significant" class, the "Not Significant" class loses. This gives preference to clusters that have been detected through tests for spatial autocorrelation. The values of the consolidation matrix have been applied to a new variable in the spreadsheet through the use of a complex conditional "if" statement. The result is one income variable that is now ready to be consolidated with the other three remaining variables. This has been done next.

To combine the clustering patterns of the four variables of interest, a new variable has been assigned the concatenated text values of the numbers GeoDa used to classify clustering characteristics (0, 1, 2, 3, and 4) in each of the percentage of low education, percentage of immigrants, and percentage of visible minority variables. The new consolidated percentage of low income variable provides the fourth value for concatenation.

The new variable developed to consolidate clustering information ultimately retains the information from the original variables and this is used to identify a specific spatiotemporal context for the analytical technique. It should be noted that position and value of the four digits have significance.

The value of "1110" indicates "high – high" clustering in each of the percentage of low education, percentage of immigrants, and percentage of visible minority variables while there is no significant clustering for the percentage of low income variable. For this reason, care has been taken to concatenate the values in a predetermined order, this order is: percentage of low education, percentage of immigrants, percentage of visible minority, and percentage of low income. This new variable associates the areas in the study area with specific colocation of clustering patterns. This variable will be defined as a "colocation identifier" and it is used, as other "key" values are, to link to other information that can define what a specific spatiotemporal context implies.

Associating Spatiotemporal Context with a Relative Score

Spatiotemporal context may be identified for a given space-time event, but that context is not useful without meaning or interpretation. Ideally, if context is known, some measure of the impact on short-term and/or long-term health could be indicated by the new analytical technique. For example, what does it mean to live in an area where clustering of low educational attainment exists? The technique developed here is built on methods that restrict the potential spatiotemporal contexts to a finite number. Each of these possibilities has been identified and meaning has been associated with each context. A "simple score" has been calculated (developed in Appendix D) to provide an indicator for the impact that a given spatiotemporal context will have on health.

Combinations of clustering pattern colocations may be defined by any of the classifications that are made during tests for local spatial autocorrelation. That is, a census tract may be classified as "high – high", "low – low", "low – high", "high – low", or "not significant" in a test for local spatial autocorrelation on a particular variable; this results in five possibilities. For this research, four variables are being considered: percentage of low education, percentage of visible minority, percentage of immigrants, and percentage of low income leading to 625 possible combinations and each combination may have a unique impact on an object in motion.

Appendix D identifies each of the 625 unique colocation possibilities and develops a simple score for each one. To calculate a score for a census tract for a given year, the number of "high – high" clusters is tallied and represented as a positive number. The number of "low – low" clusters are also tallied and represented as a negative number. The two values are then summed to provide the simple score. To illustrate, a colocation identifier of "1102" indicates two "high – high" clusters and one "low – low" cluster. This results in 2 + (-1) or a score of 1. It should be noted that the larger positive numbers are reflective of more negative health contexts, while smaller negative numbers indicate less negative influences. While a simple score can be used for the development of a methodology to consider spatiotemporal context on movement objects, an expert in a health-related field would be more qualified to assign alternative magnitudes for particular colocations combinations in a production environment.

Simple scores have been associated with each census tract in the study area for the 2001 and 2006 study years. Collectively, these simple scores represent the spatiotemporal context for the study area at each of these points in time. Cartographically, this approach has consolidated the previously reviewed individual LISA map results (see Appendix C) into one indicator, the “simple score”, and mapped that score for the appropriate census year. The resulting simple score map for 2001 appears in Figure 2 while the results for 2006 are in Figure 3. For these maps, blue regions are considered to have less of a negative impact on health outcomes, while red regions have a greater negative impact on those outcomes. An individual can add the scores for the locations along their space-time path and then divide by the number of scores to get the score for their complete space-time path.

One of the most notable characteristics across the two maps is that the 2001 map has no scores for either positive or negative 4. This result occurs as no region has either all four variables with “high – high” clusters or “low – low” clusters. Scores of positive or negative 3 are the limits of the possible scores for 2001. In 2006, scores for both positive and negative 4 are possible and this could suggest a polarization of the social classes. As identified in the discussion on the LISA maps in Appendix C, the Toronto “U” pattern is evident here and the city of Toronto appears to be relatively worse off than the rest of the Toronto CMA.

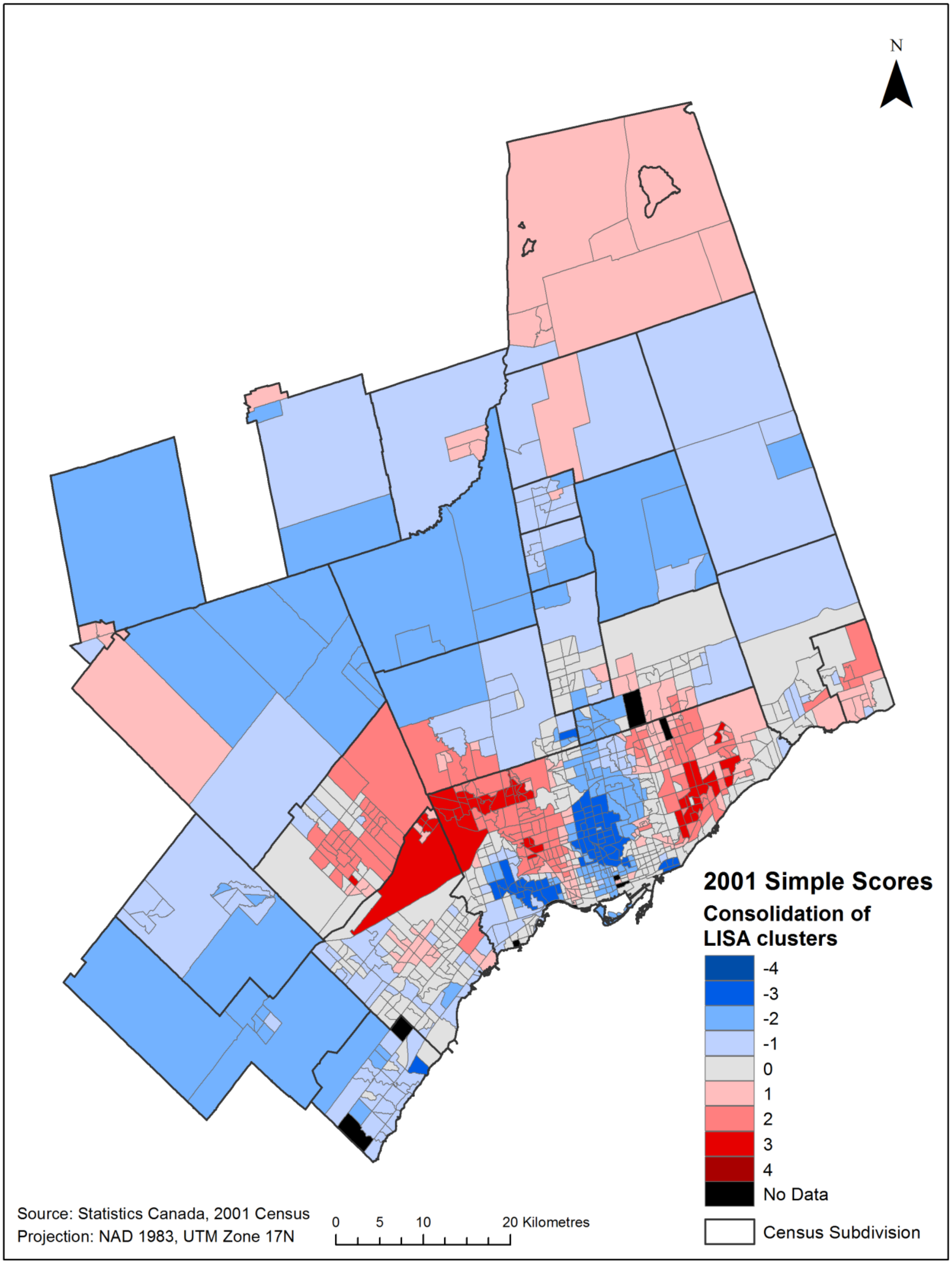

Figure 2: 2001 Toronto CMA Simple Scores

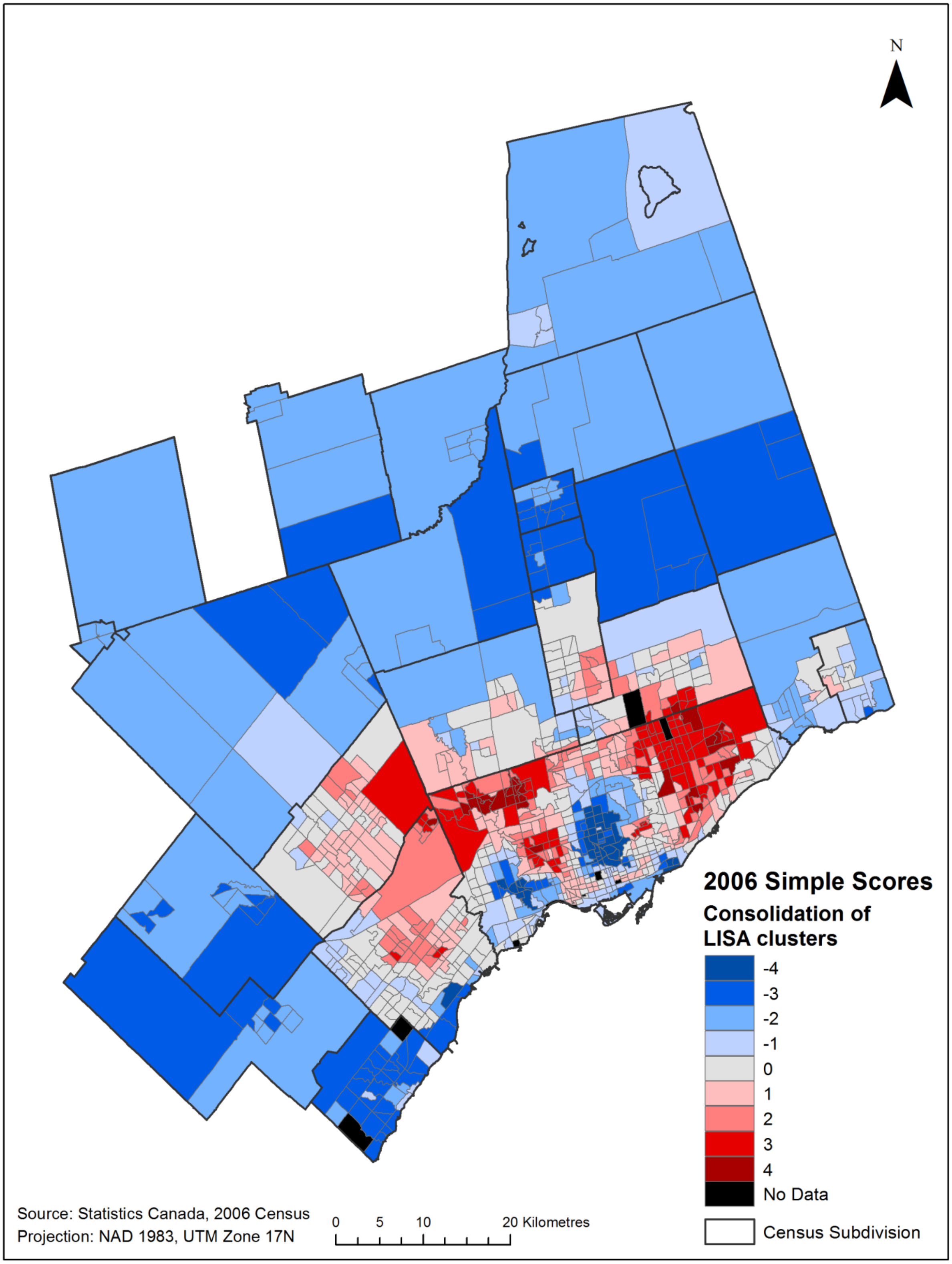

Figure 3: 2006 Toronto CMA Simple Scores

Development of a SQL Relational Database for the Analytical Technique

To connect individuals to context scores developed using spatial analysis at different points in time, this study develops three primary tables within a larger SQL database. A fourth database table to store and manage the "simple scores" is also developed. The database structure is presented in Figure 4.

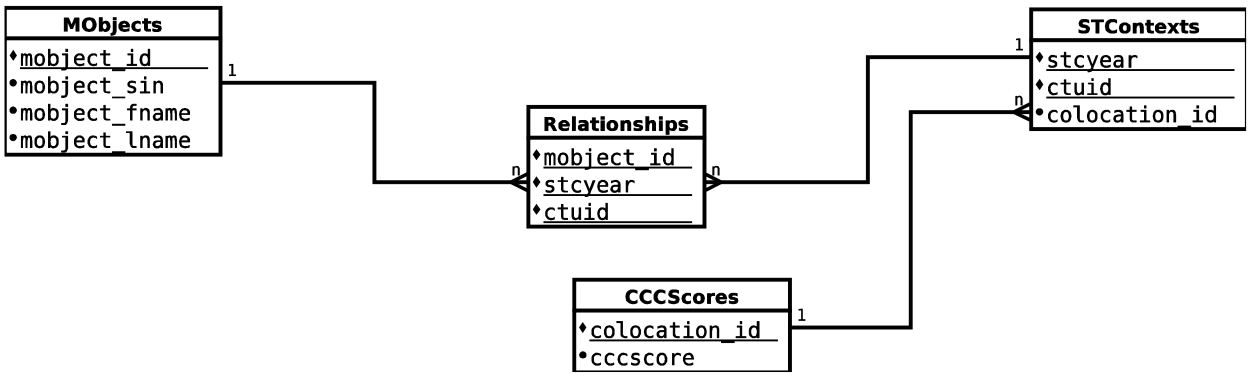

Figure 4: Relational Database for New Analytical Technique

Creation of the Movement Objects Table

Movement objects are individuals in the context of this study. The only significant requirement is that individuals are identified with a unique value, such as a social insurance number or a health insurance number; this permits them to be uniquely identified within the data and used in combination with other database tables. The diagram presented in Figure 4 presents the movement objects table as "MObjects" with the fields of "mobject_id", "mobject_sin", "mobject_fname", and mobject_lname". These fields represent a unique and independent identifying field (mobject_id) along with the social insurance number, first name, and last name for an individual. Using fictitious individuals, Table 3 illustrates how data in this table might appear.

Table 3: Fictitious Data for Movement Object Data

Movement Object Identifier	Social Insurance Number	First Name	Last Name
1	201500001	James	Smith
2	201500002	Mary	Johnson

Creation of the Spatiotemporal Context Table

The spatiotemporal contexts for possible space-time events for this study were developed using the methodology outlined in the previous section. Three components need to be identified to establish that context. A table within the database is required to track the time and location for the event; this is provided as a year and a census tract. A colocation identifier is used to specify the context. This table has been created as "STContexts" and has the fields "stcyear", "ctuid", and "colocation_id" to record the year, census tract identifier, and colocation identifier respectively. This table appears in the database diagram presented in Figure 4 and the data developed earlier in this methodology is used to illustrate in Table 4 how the data appear in that database table.

Table 4: Data for Spatiotemporal Context Data

Year	Census Tract ID	Colocation Identifier
2006	5350215.00	4022
2006	5350312.02	1111

Creating the Spatiotemporal Relationships Table

A spatiotemporal relationships table captures an individual's movement and associates that movement event with a spatiotemporal context. This table requires that the identifier for the movement object (mobject_id) and space-time identifiers (stcyear and ctuid) for the context of the movement be recorded. The "Relationships" table appears in Figure 4 as part of the database concept. Table 5 presents the table design with hypothetical data.

Table 5: Data for Spatiotemporal Relationship Data

Movement Object Identifier	Year	Census Tract ID
1	2001	5350175.02
1	2006	5350275.00
2	2001	5350311.05

Database relationships are often classified as "one to one", "one to many", or "many to many" relationships and the "one to many" relationship type is typically used. To illustrate, one individual may have many points on their space-time path. It is expected that each individual will have multiple entries in the "Relationships" table for the one entry in each of the "MObjects" and STConexts tables. Table 5 illustrates the multiple entries in the "Relationships" table: one entry for each location an individual has lived within for each time period they have lived in the Toronto CMA is recorded.

Creating the Contextual Simple Scores Table

Figure 4 presents the "CCCScores" table and it is created with two fields: "colocation_id" and "cccscore". "colocation_id" is the colocation identifier that has previously been developed and stored in the STContexts table while the "cccscore" field stores the "simple score" developed using the procedure outlined in Appendix D. A "one to many" relationship exists between the "STContext" and "CCCScores" table where one colocation identifier, linking to its associated score, may appear on many spatiotemporal contexts. Table 6 illustrates with the data from Appendix D how this table might look.

Table 6: Data Table for Simple Scores

Colocation Identifier	Simple Score
1211	2
1220	-1

Completing the Creation of the Database and Developing SQL Queries

To use the database to trace individual exposure to contextual effects across the two time periods of this study, three separate SQL queries were developed. Collectively, these queries address the central research questions of this study. Queries may also be created as a researcher needs them.

SQL Query 1: Detection of Spatiotemporal Paths with the Least Desirable Simple Scores

First, a query identifies from the data, which individual's spatiotemporal paths have the overall least desirable simple scores. In this data a score greater than 2.5 is considered to be "least desirable". This query identifies the highest average simple score for individuals and their space-time paths by averaging the two simple scores they have in this data for 2001 and 2006. In general, this cumulative score would be calculated as the sum of the path scores divided by the number of scores. This query then returns each component of the individual's space-time path. The query is compatible with additional time periods if that data was made available.

```
SELECT mobjects.mobject_sin, mobject_fname, mobject_lname,
      stcontexts.stcyear, stcontexts.ctuid

FROM mobjects, stcontexts, relationships

WHERE mobjects.mobject_id=relationships.mobject_id
      AND relationships.stcyear=stcontexts.stcyear
      AND relationships.ctuid=stcontexts.ctuid
      AND mobjects.mobject_sin
      IN (

            SELECT mobjects.mobject_sin
```

```
            FROM mobjects, stcontexts, relationships, cccscores

            WHERE mobjects.mobject_id=relationships.mobject_id
                  AND relationships.stcyear=stcontexts.stcyear
                  AND relationships.ctuid=stcontexts.ctuid
                  AND stcontexts.colocation_id=cccscores.colocation_id

            GROUP BY mobjects.mobject_sin
            HAVING avg(cccscore) > 2.5)

ORDER BY mobjects.mobject_sin ASC;
```

SQL Query 2: Identify the Simple Score for all of the Census Tracts in a Census Year

This query may be used to create data sets that could be combined with boundary files to present a risk map where lower-risk areas are presented in blues, neutral areas are presented in grey, and higher-risk areas are presented in reds. This query could also be modified with the addition of a “where” clause to restrict the results to one census tract.

```
SELECT stcontexts.stcyear, stcontexts.ctuid, cccscores.colocation_id,
      cccscore

FROM stcontexts, cccscores

WHERE stcontexts.stcyear=’2006-01-25’
      AND stcontexts.colocation_id=cccscores.colocation_id

ORDER BY stcontexts.ctuid;
```

SQL Query 3: Create a List of Individuals with Their Space-Time Paths Cumulative Score

The third query developed here is used to find the cumulative simple score associated with a person’s space-time path. Results for each person in the database are calculated and are then presented as a list with one row for each person and their score. This query may be used to identify those on poor trajectories and allow for action to be taken to reduce the negative impact of their path.

```
SELECT mobjects.mobject_sin, mobject_fname, mobject_lname,
      avg(cccscore) as avgscore

FROM mobjects, stcontexts, relationships, cccscores

WHERE mobjects.mobject_id=relationships.mobject_id
      AND relationships.stcyear=stcontexts.stcyear
      AND relationships.ctuid=stcontexts.ctuid
      AND stcontexts.colocation_id=cccscores.colocation_id

GROUP BY mobjects.mobject_sin
ORDER BY avgscore DESC;
```

Results: Testing of the Analytical Technique

To assess the technique developed in this study, data from the 2001 and 2006 Census were analyzed and combined with generated individual data. Data describing individuals is not available for this study. To assess the new technique, hypothetical individuals have been created. Popular first and last names have been used to create 1000 non-existent individuals for this work. A 9-digit "social insurance" number" has been created by using the number 20150001 and incrementing this value by one for each "person" in the data. These "individuals" have been randomly assigned to a starting census tract for the year 2001. The spreadsheet randomly determined if these people would relocate for 2006 or not. The probability for relocation is approximately 50% to match the characteristics of the Toronto CMA. Given this fictitious data and the developed mobility patterns, the remaining data needed to complete the database has been imported into the movement objects and relationships tables. Queries were then run to explore the new technique's functionality.

Query 1: Detection of Spatiotemporal Paths with the Least Desirable Simple Scores

The first query provides the functionality to retrieve from the database the individual components of space-time paths for individuals with the worst overall space-time path scores. This allows attention to be focused on those individuals who require it the most. Query results are summarized in Table 7. These results provide each of the components of an individual's space-time path in an unconsolidated format so that the analyst may review trends in the path information.

The query is capable of revealing that some changes in scores experienced by individuals are attributable to changes in context; this is the case for Daniel Lewis at the top of Table 7. Even though he has not relocated, his 2006 simple score is now worse than his 2001 score. Alternatively, Stephen Green who has relocated between the two periods has achieved the same score as Daniel Lewis. Others have maintained a consistently poor score by failing to relocate; this is the case for Sandra Lopez. Each example from the data present individuals who have obtained an average space-time path score of 3.

Table 7: Sample Results for Query 1

mobject_sin	mobject_fname	mobject_lname	stcyear	ctuid	cccscore
201500023	DANIEL	LEWIS	2001	5350401.12	2
201500023	DANIEL	LEWIS	2006	5350401.12	4
201500035	STEVEN	GREEN	2001	5350291.02	2
201500035	STEVEN	GREEN	2006	5350378.11	4
201500032	SANDRA	LOPEZ	2001	5350354.00	3
201500032	SANDRA	LOPEZ	2006	5350354.00	3

Query 2: Identify the Simple Score for all of the Census Tracts in a Census Year

The second query retrieves from the database a list of simple scores for each census tract for the 2006 Census year. This provides a numerical summary of local spatiotemporal contexts. Query results are summarized in Table 8. This query returns 993 rows of results, and in this form, it is not very useful. One of the options available is to use query results to prepare summary results in combination with other data and in different forms. These results have been used with geographic boundary files for the 2006 Census for the Toronto CMA to produce the map presented in Figure 3 in an earlier section. The map for 2001 was produced in a similar fashion using the different year in the "where" clause and the appropriate boundary file.

Query 2 may also be modified to restrict the results to one census tract. This could be done with an additional criteria in the "where" clause. For example, from the results of SQL Query 1, did Stephen Green's original Census tract 5350291.02 have a better performance than his new location? Should he have stayed where he was? This modified query reveals that in 2006 his original census tract maintained a simple score of 2.

Table 8: Sample Results for Query 2

stcyear	ctuid	colocation_id	cccscore
2006	5350001.00	2240	-2
2006	5350002.00	2000	-1
2006	5350004.00	0001	1
2006	5350005.00	0001	1

Query 3: Create a List of Individuals with Their Space-Time Paths Cumulative Score

The third query provides the cumulative results of an individual's space time path. An individual's cumulative is calculated as the sum of the scores they have achieved for each point along their path divided by the number of scores. If a person has 8 scores in the database, then those scores are summed and divided by 8. The output of the third query is presented in Table 9. While this set of results is similar to those presented by Query 1, these results are focused on the overall path performance rather than exploring the components of poorly performing paths.

As an example, the six people presented in Table 9 have the worst scores of the 1000 fictitious people in this study. These results are presented in descending order, and those with the best overall scores are at the bottom of this results list. This query may be used as a starting point to explore potential differences in health outcomes. If it is expected that certain people will have higher cumulative scores (based on these query results), then other queries may be used, or developed, to explore why those scores are so high and to determine if there are any patterns in the results presented by Query 3.

Query 3 could be followed by Query 1 with an additional "where" clause provided to restrict results to an individual of interest. This could provide insight into an individual's movements. Alternatively, the additional "where" clause could specify a census tract rather than an individual and this would allow for the composition of that census tract to be explored. Whether or not a population is similar or dissimilar in their expected health outcomes could be determined in this way.

If health outcomes for individuals are already known, then those outcomes could also be used to explore the relationships between cumulative scores and those outcomes. A new query would be required for this functionality.

Table 9: Sample Results for Query 3

mobject_sin	mobject_fname	mobject_lname	avgscore
201500183	SHAWN	ANDREWS	3.5
201500335	LEO	BARBER	3.5
201500605	FREDRICK	BULLOCK	3.5
201500205	TONY	CARR	3.5
201500221	RODNEY	MORRISON	3.5
201500536	NINA	BRADFORD	3.5

Discussion and Conclusions

This study has three primary research objectives. The first objective is the development of a new analytical technique to gain a better understanding of the spatiotemporal relationships between income-related health inequalities and health outcomes. The operationalization of the new technique is the second objective. Third, discussion of how that operationalized technique could be leveraged to answer research questions.

Development of a New Analytical Technique

This study has synthesized the relevant knowledge from the existing literature to conceptually develop a new analytical technique for the purpose of gaining a better understanding of public health problems involving income-related health inequalities. These problems feature challenges that are just beginning to be dealt with in health geography literature. Locating movement objects in space and time, and accounting for the spatiotemporal context, was a requirement for the technique developed. This work has successfully adapted research by Andrienko et al. (2011) to provide the over-arching framework for this technique. Contributions by other authors including Semple et al. (2013) and Heywood et al. (2006) have been applied to techniques used to complete tasks within that framework.

This technique has been developed around the requirements of answering research questions in the area of income-related health inequalities and the variables of percentage of low education, percentage of immigrants, percentage of visible minority, and percentage of low income have been used to develop the required spatiotemporal context. As this work has successfully developed a technique for these types of spatiotemporal problems, it may be adapted to satisfy the needs of other problems with different movement objects or spatiotemporal contexts.

Operationalization of the Analytical Technique

Not all research is able to be operationalized and this may be due to limitations of available resources including finances, computing power, or even required data. Additional research may also be needed before theories may be put into action. The technique developed in this work has been moved from theory to an operational state.

The existing literature describes the challenges of identifying suitable tools to provide solutions for spatiotemporal problems. Many of the challenges described are related to the visualization of data and/or results. Spatiotemporal problems may be addressed without a visually-focused tool, such as geographic information software (gis), being used in all stages of the analysis. This strategy has been used here. The technique developed in this study has been operationalized with a common spreadsheet application and open source software tools including: the statistical analysis software “R”, GeoDa for completing tests for spatial autocorrelation, and MySQL to provide a database platform for the technique. ArcGIS has been utilized for some data preparation tasks and to prepare presentation quality maps. The operationalization of the technique developed has allowed for the research questions described in this work to be addressed.

Exploiting the New Technique to Answer Research Questions

The analytical technique developed and operationalized in this research has been motivated by the need to answer specific research questions. The spatiotemporal variants of the spatial questions asked in the “Introduction” section included: "Can we verify that those with poor health outcomes have lived in one or more places when negative influences on health were present?", "Do certain space-time paths show greater health risks to individuals?", "Can we identify those on poor trajectories and take action to reduce the negative impact of their path?", "Are there any patterns of 'path bundling' that are of interest?”. The SQL database and query language operationalized in this work has been used to develop queries to answer these research questions.

The new technique, and the way it has been operationalized, is flexible. Additional research questions may be developed and the database queried. As SQL is a well-established database language, significant opportunity exists to leverage the potential information that may be extracted by using the new technique.

Limitations and Opportunities of the Developed Analytical Technique

While the methodology provides a strong foundation for a technique that may be used to better understand the relationship between income-related health inequalities and health outcomes, there are some opportunities where additional research has the potential to improve the approaches presented here.

The ability to achieve the best results from the technique is dependent on the quality of the context estimated for a certain spatiotemporal event that has occurred. A “simple score” has been applied for this research, but access to expert knowledge could be leveraged to better calibrate the context used by the technique and this would provide a more accurate evaluation of each individual spatiotemporal context. While this is true, there are questions of scalability. The current approach of developing a spatiotemporal context yields 5^n potential contexts where “n” is the number of variables in the study. An approach with more variables than the four used here could become unmanageable. Identifying a method to handle a larger number of variables, and therefore more complex problems, would be useful.

Not all variables have a linear relationship with health outcomes. Ng (2011) demonstrated that recent immigrants can be very healthy. An area with a high percentage of recent immigrants could possess a high level of health. This is a contradiction to the assumptions of the developed technique and this could lead to interpreting local cluster classifications of "high – high or "low – low" incorrectly. The simple score produced in this situation would also have a built-in error. A similar issue could be found with areas exhibiting the "ethnic density" effect as presented by Pickett and Wilkinson (2008). Additional research is required to expand on the foundation established here to better handle variables with characteristics of non-linearity.

Kwan (2012a) has previously informed on the "uncertain geographic context problem" and it is important to be aware that those who live in a given census tract do not live identical lives. Two people who live beside each other are likely to have significantly different activity spaces. The existing research uses census data that is published every five years to develop the spatiotemporal context. Is there an opportunity to provide a more frequent measure of this context? Can location-aware devices assist in the process of informing on context? What time intervals are most appropriate to support these kinds of spatiotemporal problems? There is a lot of room for additional research in this area.

The availability of data has the potential to become a challenge to many fields of research. Recent changes to the Canadian Census have raised questions by local governments, businesses, and institutional users as to whether there are quality issues that could make the Census an unreliable source of data. If the Census is not able to provide for data needs, then what suitable alternatives exist? Aside from the Census, individual-level data can also be challenging to obtain. Current attitudes on personal privacy have made it more difficult to obtain personal information. The public needs to recognize the value in supporting and participating in projects that have the goal of improving the population's health. Through the use of location-aware devices and "apps" some businesses are gathering a significant amount of data on individuals and their behaviours. Is there any reason why a business should be considered to have a greater need than those applications that contribute to the public good and the improvement of public health?

A Future Direction

This research has contributed a technique that has been developed with the aid of fictitious data on individuals. This work may now be expanded upon by applying the developed technique to data on real individuals. As discussed, calibration of the spatiotemporal context may be required. If data on individual's health outcomes are also available, then additional new research questions may be asked, and explored, by expanding on the technique developed in this work. New database tables and queries may be created to support this expansion. At present, the new technique estimates a future relative quality of health based on past spatiotemporal context and an individual's historical space-time path. With additional health outcome data for individuals, greater insight into the relationship between determinants of health and health outcomes may be obtained.

This technique offers expandability as a feature in its original design and this provides for the opportunity to resolve current limitations or provide new functionality. For example, if individual-level data becomes available it might be possible to do more than just plot a space-time path from this data. An individual's characteristics such as gender, income, family status, or employment status could be used to develop an "individualized context" to replace what has been inferred by aggregate-level census data. Further, new datasets could provide for context that is truly spatial. Examples could be physical characteristics of space such as the amount of pollution or greenspace/parks. Other characteristics such as a neighbourhood's social capital could also be used to contribute to spatiotemporal context. This approach could incorporate the concepts of compositional and contextual factors of health, as described by Kawachi et al. (2002). Further, if the technique was expanded upon in this way, data might be available to explore cases of specific diseases. If breast cancer was the disease of interest, a query could be written to consider only women who have had breast cancer and then consider their space-time paths.

The technique developed in this research presents an opportunity to gain a better understanding of income-related health inequalities and the relationship to health outcomes. With this greater understanding, the opportunity to take action exists. The product of this research has the potential to inform public health policy decisions and impact the quality of health of individuals. This technique demonstrates that it is now possible to explore the kinds of spatiotemporal questions asked in this research. Availability and access to data will now determine the quality of the answers that the technique is able to offer.

Appendix A: Techniques in Spatial Statistics

Techniques to test for spatial autocorrelation have been identified in the "Literature Review" section of this research. This appendix presents each of the equations for the techniques identified and reviews how to interpret the results provided by these techniques. Spatial weight matrices are also reviewed as these matrices provide an input required by the equations reviewed here.

Techniques for Testing for Spatial Autocorrelation

The Global Moran's I is calculated using Equation 1 in Table 10. Results range from approximately -1 to 1 where positive results will indicate a degree of autocorrelation, negative results will indicate a degree of negative autocorrelation, and values around 0 will indicate an absence of pattern. Geary's C is another global statistic and is calculated using Equation 2 in Table 10. Values for C will range from 0 to 2. Where spatial autocorrelation exists values will trend toward 0, negative spatial autocorrelation result in values trending towards 2, and an absence of spatial autocorrelation will result in values around 1. Chun and Griffith (2013) explain that the Getis Ord statistic, presented as Equation 3 in Table 10, is interpreted in a similar fashion to the Moran's I. Positive results of G indicate spatial clustering of high values while negative results of G indicate clustering of low values for a given variable. Results for all of these methods must also be significant according to a significance test. Rogerson and Yamada (2009) demonstrate that interpreting the local statistic results is similar to the global statistics. Table 10 presents the Local Moran Statistic in Equation 4 along with variations of the Getis G_i in Equations 5 and 6. Chun and Griffith (2013) clarify that the Gi does not include i in the calculation while the G_i* version does.

Table 10: Global and Local Spatial Statistics Equations

Equation 1: Global Moran's I	
$I = \frac{m \sum_i^m \sum_j^m w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{(\sum_i^m \sum_j^m w_{ij}) \sum_i^m (y_i - \bar{y})^2}$	
Equation 2: Geary's C (Global Statistic)	
$C = \frac{(m-1)\sum_{i=1}^m \sum_{j=1}^m w_{ij}(y_i - y_j)^2}{2(\sum_{i=1}^m \sum_{j \neq i} w_{ij}) \sum_{i=1}^m (y_i - \bar{y})^2}$	
Equation 3: Getis and Ord's Global Statistic	
$G(d) = \frac{\sum_{i=1}^m \sum_{j=1}^m w_{ij}(d) y_i y_j}{\sum_{i=1}^m \sum_{j=1}^m y_i y_j}, i \neq j$	
Equation 4: Local Moran Statistic	
$I_i = \frac{m(y_i - \bar{y})}{\sum_j (y_j - \bar{y})^2} \sum_j w_{ij}(y_j - \bar{y})$	
Equation 5: Getis' Gi Statistic (Local Statistic)	
$G_i = \frac{\sum_j w_{ij}(d) x_j - W_i \bar{x}}{s(i) \left\{ [(m-1) S_{1i} - W_i^2] \big/ (m-2) \right\}^{1/2}}, j \neq i$	$where\ W_i = \sum_{j \neq i} w_{ij}(d), and$ $S_{1i} = \sum_{j \neq i} \{w_{ij}(d)\}^2$
Equation 6: Getis Gi* Statistic (Local Statistic)	
$G_i^* = \frac{\sum_j w_{ij}(d) x_j - W_i^* \bar{x}}{s \left\{ [m S_{1i}^* - W_i^{*2}] \big/ (m-1) \right\}^{1/2}}, all\ j$	$where\ W_i^* = \sum_j w_{ij}(d), and$ $S_{1i}^* = \sum_j \{w_{ij}(d)\}^2$

Notes:

1) The term i is the region of interest, and the term j is a region not of interest.
2) The term m is the number of regions in the study area.
3) The term d is the distance between regions i and j.
4) The term y_i is the observed value from the area of interest while $\bar{y}$ is the mean of y_i.
5) The term w_{ij} refers to the value in the spatial weights matrix that defines the spatial relationship between *i* and *j* being considered.
6) The terms $\bar{x}$ and s are the sample mean and standard deviation respectively.

Spatial Weight Matrices

To operationalize the global or local spatial statistic equations and approaches already discussed, some measure of what is and is not "local" needs to be specified. This measure is represented in each of the equations in Table 10 by the w_{ij} element. Unwin and Unwin (1998) provide guidance on this topic. One of the typical approaches to provide for w_{ij} is to create a "contiguity" or "adjacency" matrix. This matrix identifies the relationship between an area i and every other area j. For example, if there are ten areas, then a 10 x 10 adjacency matrix will be created. Areas cannot have a relationship with themselves and these cells in the matrix are usually filled with a "0". It should also be noted that the matrix will be symmetrical across the line of diagonal zeros. It is typical to enter a value of "1" into the matrix where two areas are considered to have a local relationship and a "0" where they do not.

There are a variety of methods in determining whether two areas should be deemed to have a local relationship, Unwin and Unwin continue to explain that if two areas share a border, then they are considered to have a local relationship and a "1" would be entered into the matrix cell for this i and j. Freisthler et al. (2006) expand on the techniques available for determining adjacency. Configurations matching chess pawn movements are referenced as typical choices. A queen configuration would consider all surrounding neighbours, while a rook configuration would ignore the diagonal relationships that have smaller shared borders. The "order" of the relationship is also discussed by Freisthler et al. (2006) and this simply refers to whether neighbours of neighbours are considered. For example, on a chess board, a first order queen configuration would have eight neighbours, while a second order queen configuration would extend out beyond the first eight neighbours to the next eight squares on the same "line of sight" to yield a total of 16 squares that would be considered as neighbours. This would result in sixteen "1" entries into the contiguity matrix.

While adjacency is one method of developing the "weights" matrix, there are other options. Cromley and McLafferty (2012) present that distance measures may be used to develop a weighting scheme. If the distance between the centroids of two areas is within a specified threshold, then they may be considered close enough to each other to be neighbours and have a "1" entered into the matrix; if they are farther apart than the threshold, then a "0" is entered.

Getis and Aldstadt (2004) warn that some care should be given to the development of a weights matrix. If the weights matrix is not reflective of the actual relationships that exist in the study area then a misspecification error may occur. If prior knowledge of the spatial relationships is known, then this should guide the development of the weights matrix.

Bivariate Moran's I

Locating spatial clusters with local spatial statistics is usually done for one variable at a time. There are, however, a few software products that will allow for the inclusion of two variables for a local Moran's I test. The GeoDa application is one of these. According the GeoDa user's guide (Anselin, 2003) a Bivariate LISA test, employing steps similar to those used for the univariate version, may be performed. Dijkstra, Janssen, De Bakker, Bos, Lub, Van Wissen, and Hak (2013) make use of this functionality to assess colocation of a predictor neighbour variable with the local outcome variable.

Semple et al. (2013) also considered the bivariate LISA technique for their research; however, they found that the tool did not perform as they expected. The tool, as provided in GeoDa, offers the ability to test an area of interest on one variable and neighbouring areas on a second variable. To illustrate, this might be the case of an area with a high incidence of crime being surrounded by a high incidence of low income areas. This does not show that the area of interest is also a low income area, and this is the type of result that is usually desired when exploring variables for colocation.

Ideally, GeoDa would provide the functionality to identify clusters where both variables are considered for both the area of interest and the neighbours. As an example, an area of interest with high incidence of poor health might also have high incidence of low income. This area of interest would be surrounded with neighbours with similarly high incidences of poor health and low income. Following this approach, a researcher would be able to identify where clusters of high or low values across multiple variables exist. Alternative techniques are required to identify clusters defined by more than one variable.

Appendix B: Global Spatial Autocorrelation Test Results (Moran's I Scatterplots)

The scatterplots that appear in this section are the result of conducting tests for spatial autocorrelation using the Moran's I equation. Each of the five variables discussed in the "Methodology" section have been tested for each of the two years covered in this research. The 2001 results appear followed by those for 2006. Note that the Moran's I coefficient appears at the top of the vertical axis.

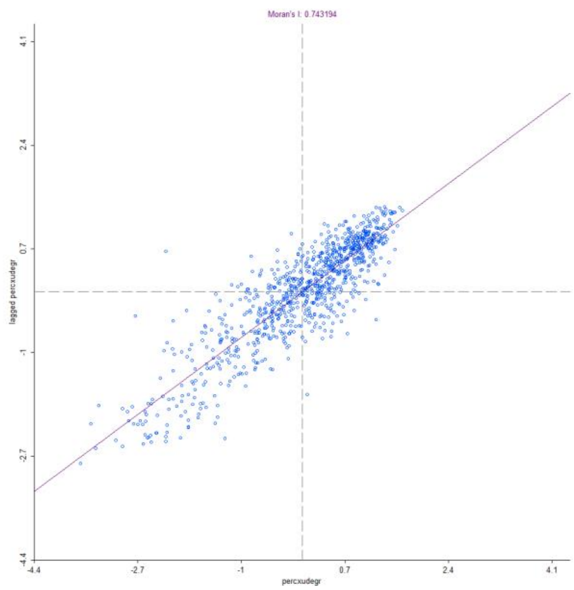

Figure 5: Scatterplot for Global Moran's I Test on 2001 Percent of Education less than University Degree

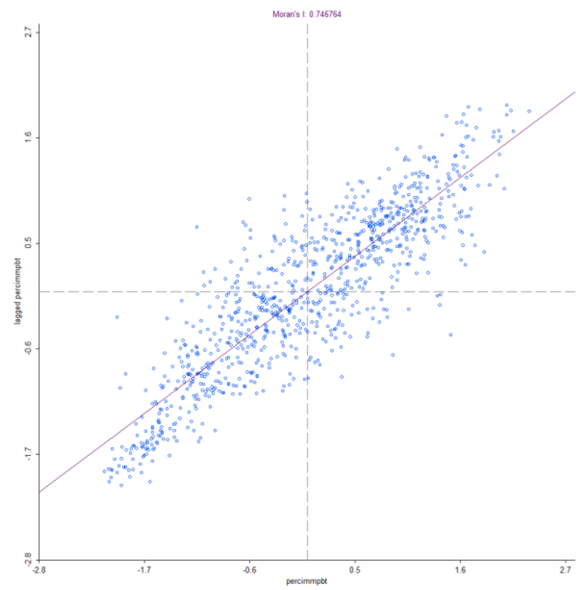

Figure 6: Scatterplot for Global Moran's I Test on 2001 Percent of Immigrants

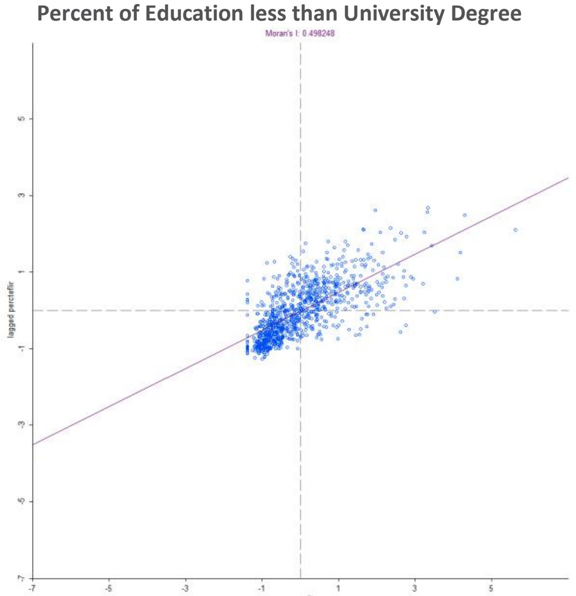

Figure 7: Scatterplot for Global Moran's I Test on 2001 Percent of Economic Families with Low Income Status

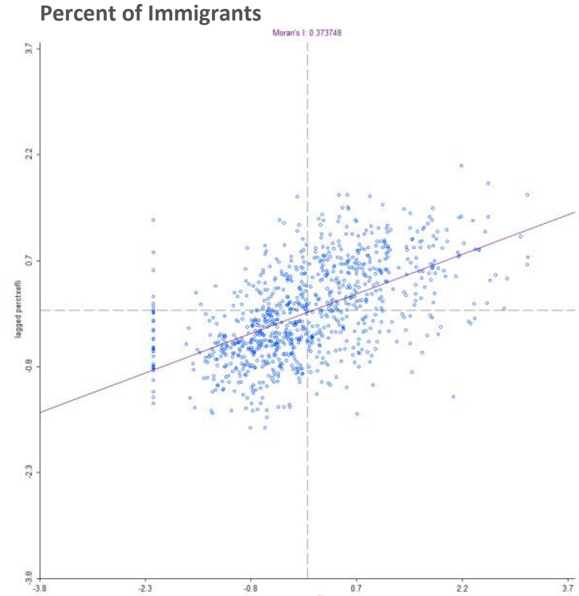

Figure 8: Scatterplot for Global Moran's I Test on 2001 Percent of Individuals not in Economic Families with Low Income Status

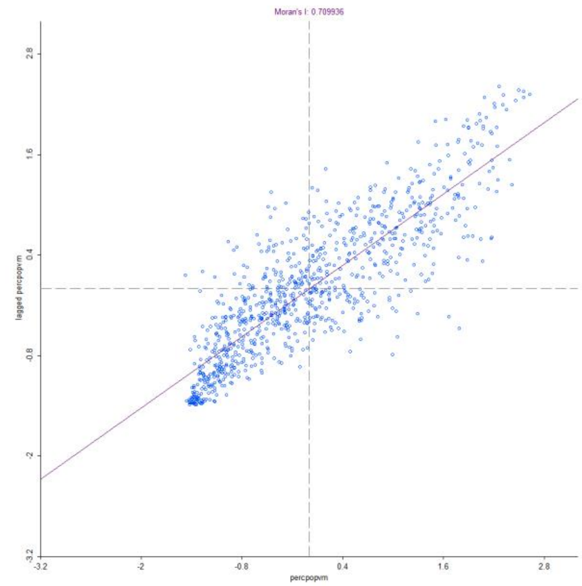

Figure 9: Scatterplot for Global Moran's I Test on 2001 Percent of Visible Minorities

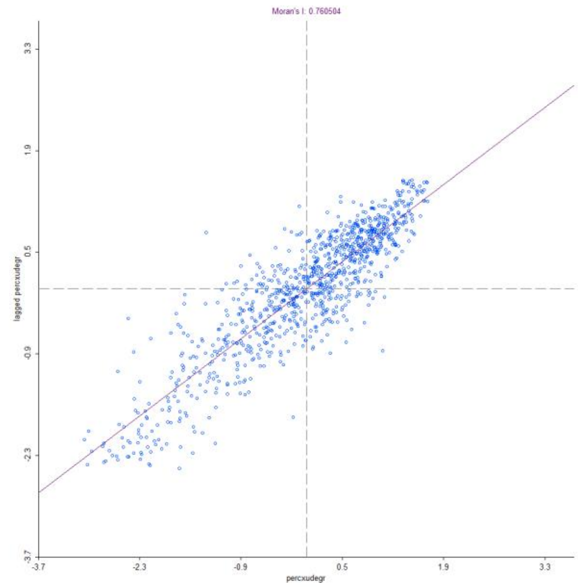

Figure 10: Scatterplot for Global Moran's I Test on 2006 Percent of Education less than University Degree

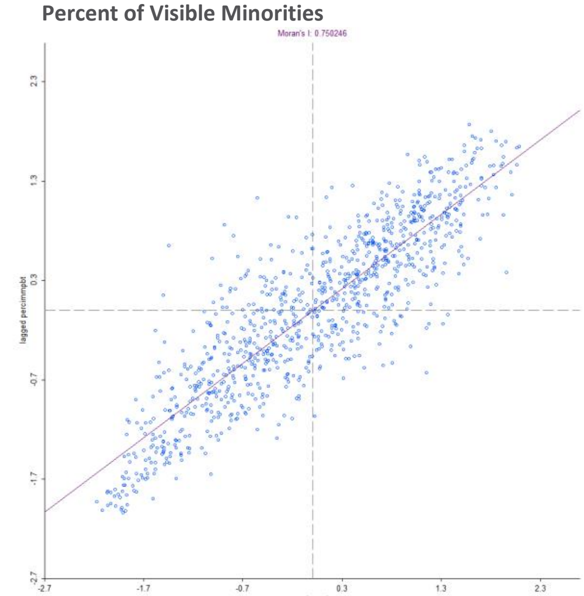

Figure 11: Scatterplot for Global Moran's I Test on 2006 Percent of Immigrants

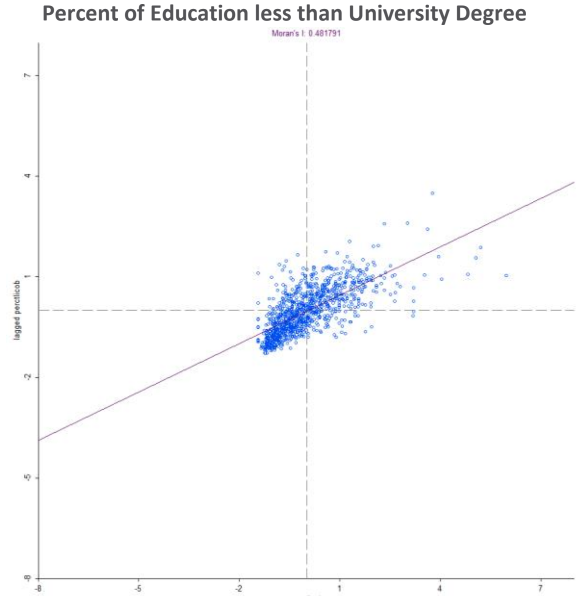

Figure 12: Scatterplot for Global Moran's I Test on 2006 Percent of Economic Families with Low Income Status

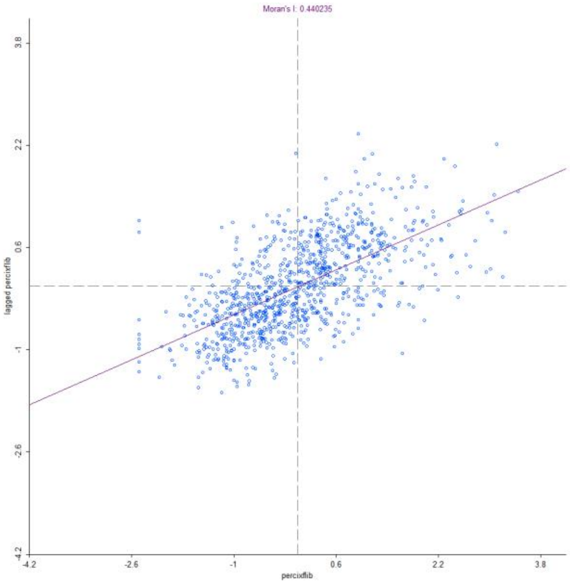

Figure 13: Scatterplot for Global Moran's I Test on 2006 Percent of Individuals not in Economic Families with Low Income Status

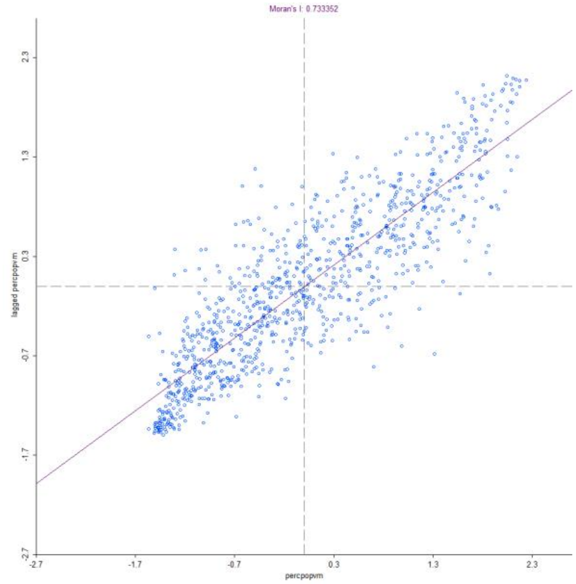

Figure 14: Scatterplot for Global Moran's I Test on 2006 Percent of Visible Minorities

Appendix C: Local Moran's I Results for Development of Spatial Context

The Local Moran's I test assigns a result value to each of the regions in the study area. The possible results are summarized in Table 11. Local Moran's I were calculated for each of the five variables developed for the 2001 and 2006 Census years.

Table 11: Possible Results and Interpretation for a GeoDa LISA Test

Result Value	Classification	Abbreviation	Interpretation
0	"Not Significant"	NS	Results of the test are not significant for this region.
1	"High – High"	HH	High values exist in the tested region and high values exist in neighbouring regions.
2	"Low – Low"	LL	Low values exist in the test region and low values exist in neighbouring regions.
3	"Low – High"	LH	Low values exist in the test region and high values exist in neighbouring regions.
4	"High – Low"	HL	High values exist in the test region and low values exist in neighbouring regions.

Ten Local Indicators of Spatial Autocorrelation (LISA) maps appear below and each of these present the pattern of local clustering for the variable being tested. For this research, "high – high" clusters are presented in dark red and indicate clusters of characteristics that are believed to have a greater negative impact on health outcomes. Alternatively, "low – low" clusters are indicated with a dark blue colour and are thought to have a lessor negative impact on those outcomes. Some key patterns are evident in the maps that have been generated, and these are discussed below. These characteristics will impact the final spatiotemporal context developed for the new analytical technique developed.

Percentage of Low Educational Attainment

For 2001, the pattern of educational attainment in the Toronto CMA is presented in Figure 15 and this pattern reflects that those who live in the downtown core have the most favourable status; this is indicated by the "low – low" clusters in blue. This cluster extends north along Yonge Street into the census subdivisions north of Toronto. The pattern also shows some pockets on both the east and west sides of the primary cluster located close to the lake. In 2001, those with the lowest educational attainment were located in the north and north-east parts of the Toronto CMA. Other clusters of similarly low educational attainment extend from the city of Toronto's north-west to the Toronto CMA north-west. A few clusters appear in Toronto east and in the Durham region. The pattern appearing in

2006 is presented in Figure 16 and this is fairly consistent with the 2001 period. The one noticeable evolution between the two years is that the two low educational attainment clusters in the north of the CMA have joined into one large cluster.

Percentage of Immigrants

Immigration clusters for both of the 2001 and 2006 Census years are presented in Figures 17 and 18 respectively. For both periods, at the core of the city of Toronto there exists a “low – low” cluster indicating fewer immigrants. Immigrants are observed to be clustered (“high” – “high”) along the north and north-west borders of the city with some pockets located just within neighbouring census subdivisions. The outer regions of the CMA all have lower representation by immigrants.

Percentage of Low Income

Low income characteristics are presented in Figures 19 to 22 for 2001 and 2006 for each of the two percentage of low income variables. Clusters of higher percentages of low income in the Toronto CMA are primarily located within the city of Toronto for both 2001 and 2006. As indicated by the lower value of the global Moran’s I statistic (see the “Methodology” section or Appendix B) for these variables, the clustering pattern is not as cohesive as with the other three variables. The general trend, in both years, is that clustering of higher percentages of low income (“high” – “high” clusters) occurs from about the downtown core of Toronto to each of the city’s north-east and north-west corners. This presents the familiar “U” pattern that is often observed in maps of Toronto.

Percentage of Visible Minorities

Lastly, Figures 23 and 24 present the results the 2001 and 2006 Census years. Visible minorities are indicated to be present in higher percentages in the north-east and north-west of the city of Toronto. In some places this clustering extends into nearby Census subdivisions such as Mississauga, Brampton, and Markham. The outer regions of the CMA have lower percentages of immigrants.

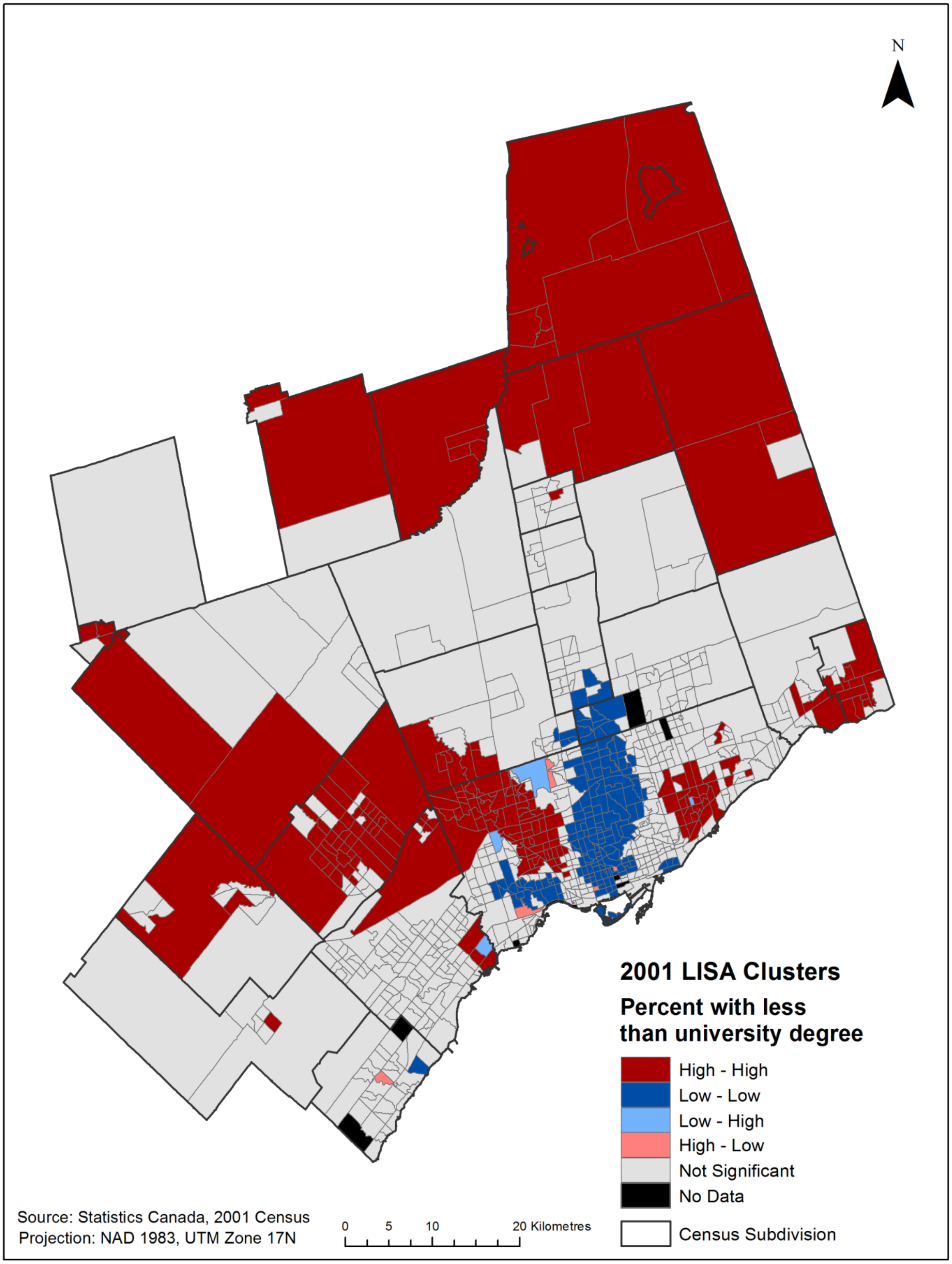

Figure 15: 2001 Toronto CMA LISA Clusters for Percentage with less than University Degree

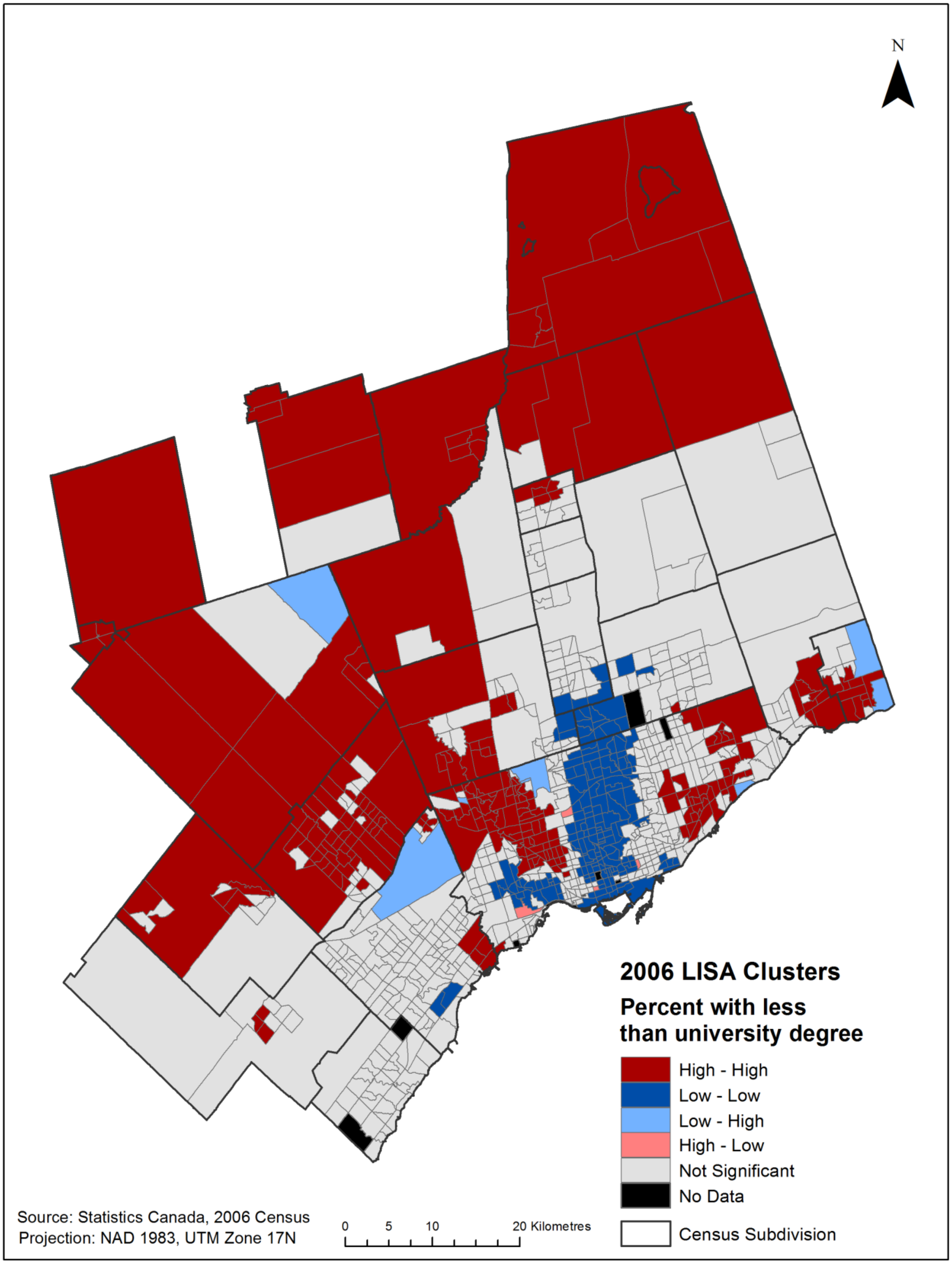

Figure 16: 2006 Toronto CMA LISA Clusters for Percentage with less than University Degree

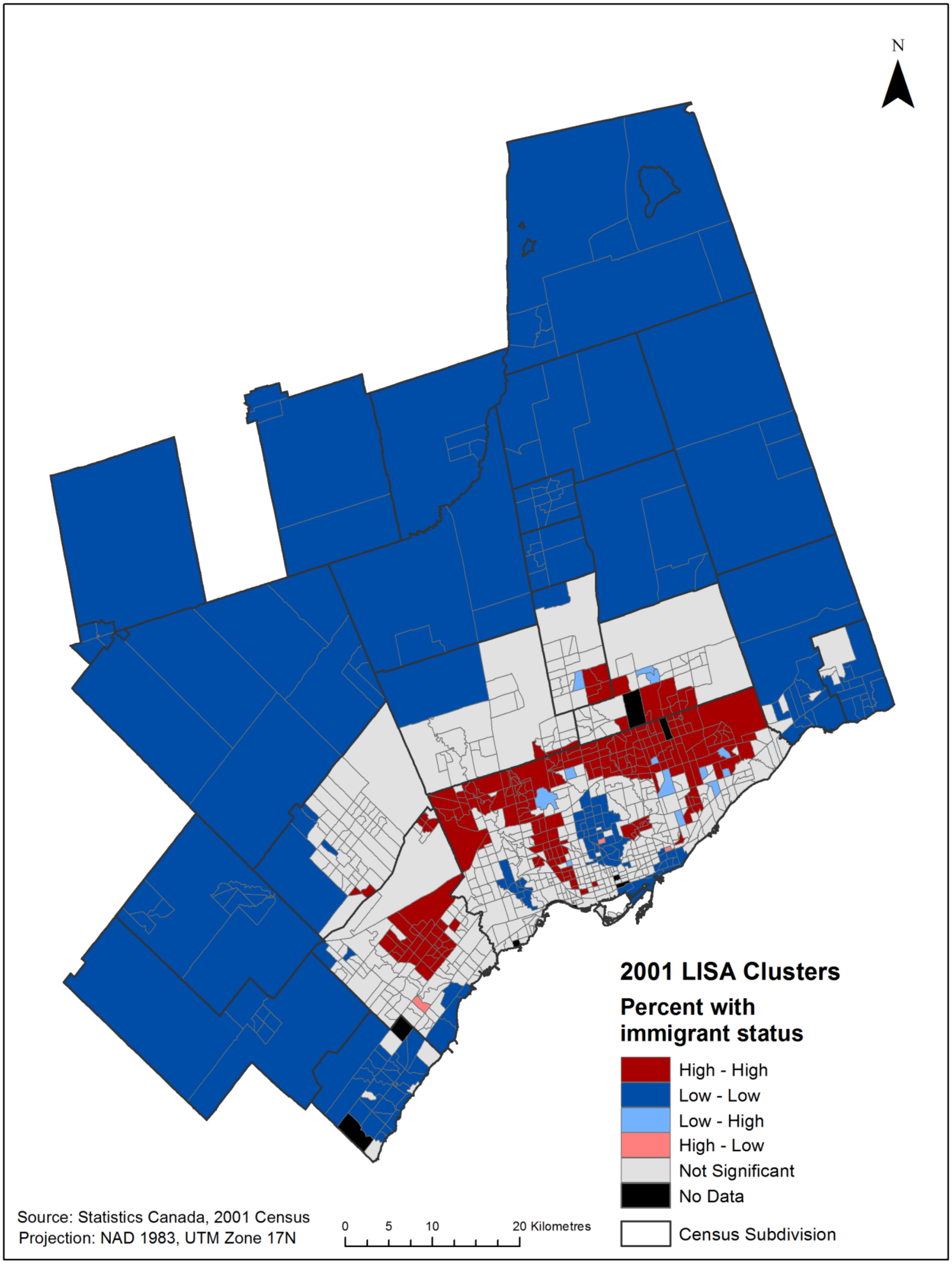

Figure 17: 2001 Toronto CMA LISA Clusters for Percentage with Immigrant Status

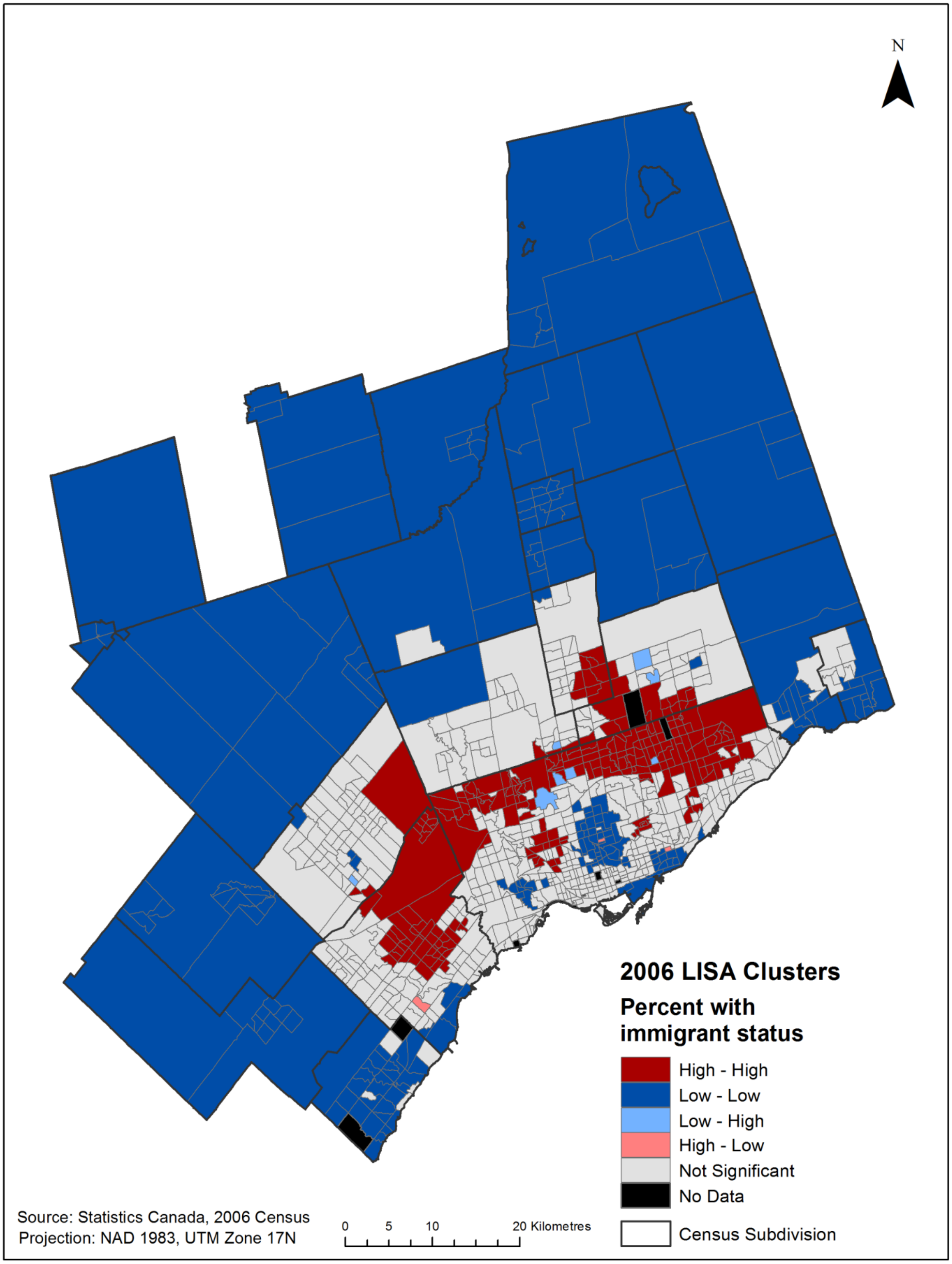

Figure 18: 2006 Toronto CMA LISA Clusters for Percentage with Immigrant Status

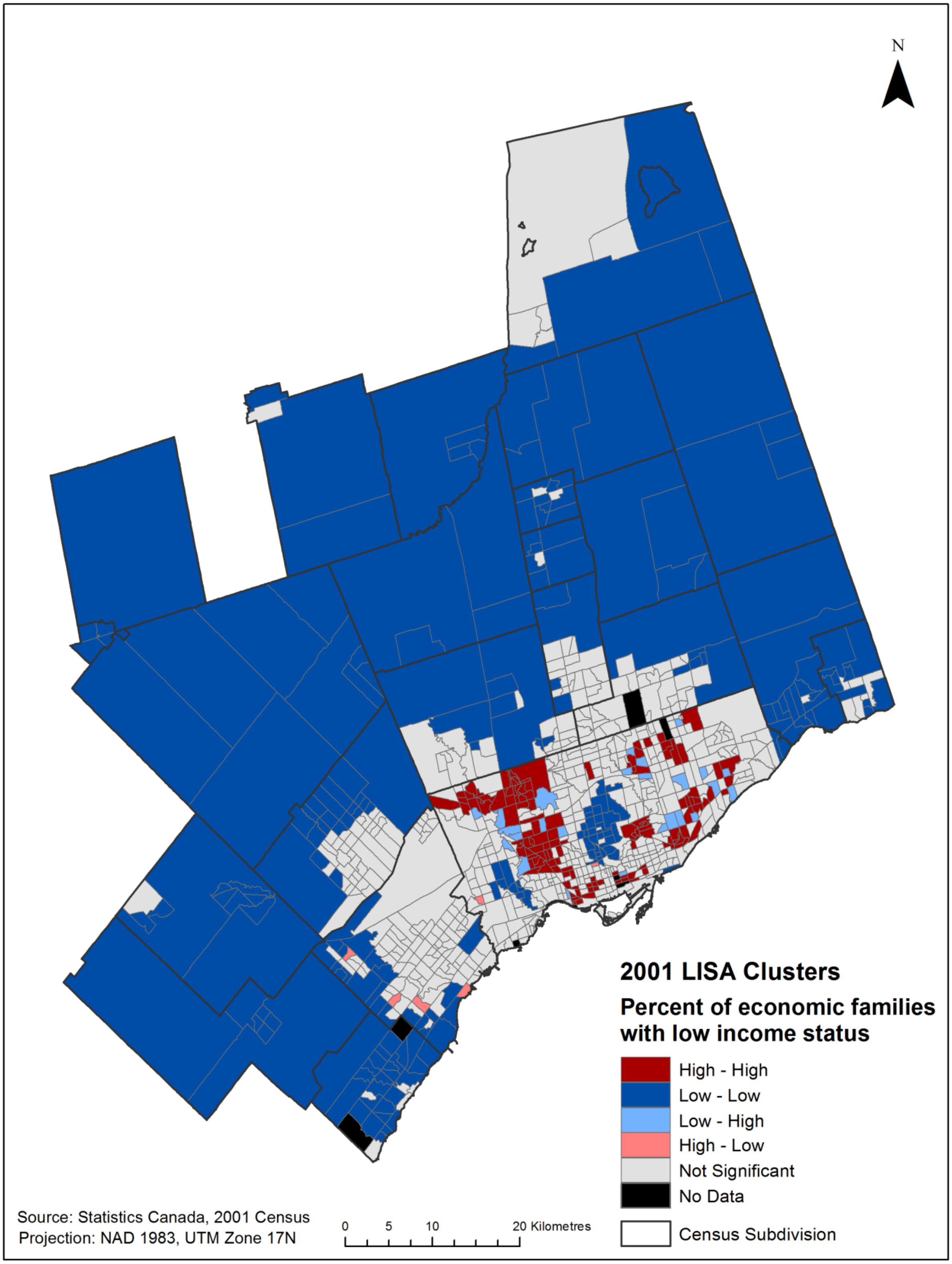

Figure 19: 2001 Toronto CMA LISA Clusters for Percentage of Economic Families with Low Income Status

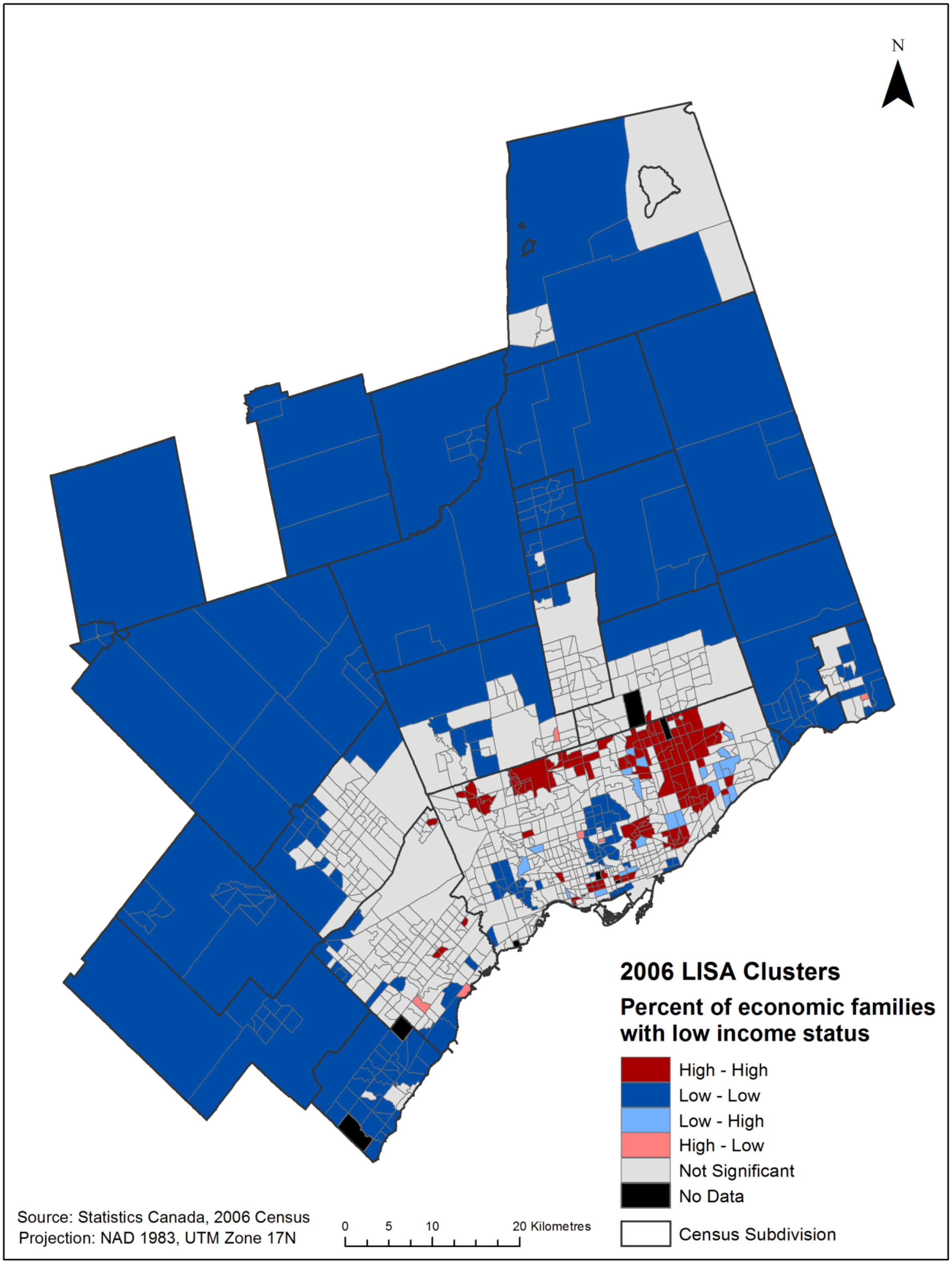

Figure 20: 2006 Toronto CMA LISA Clusters for Percentage of Economic Families with Low Income Status

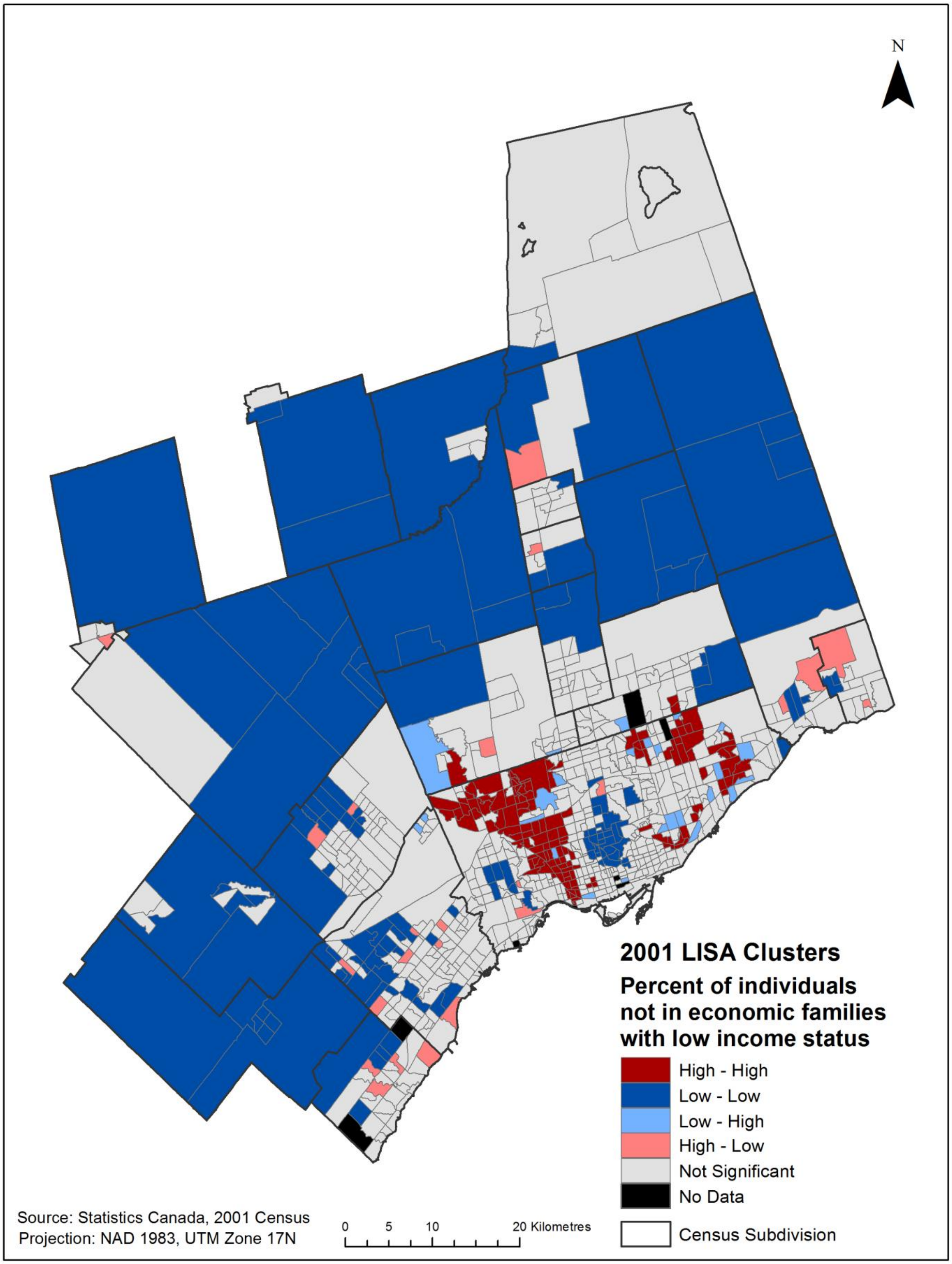

Figure 21: 2001 Toronto CMA LISA Clusters for Percentage of Individuals not in Economic Families with Low Income Status

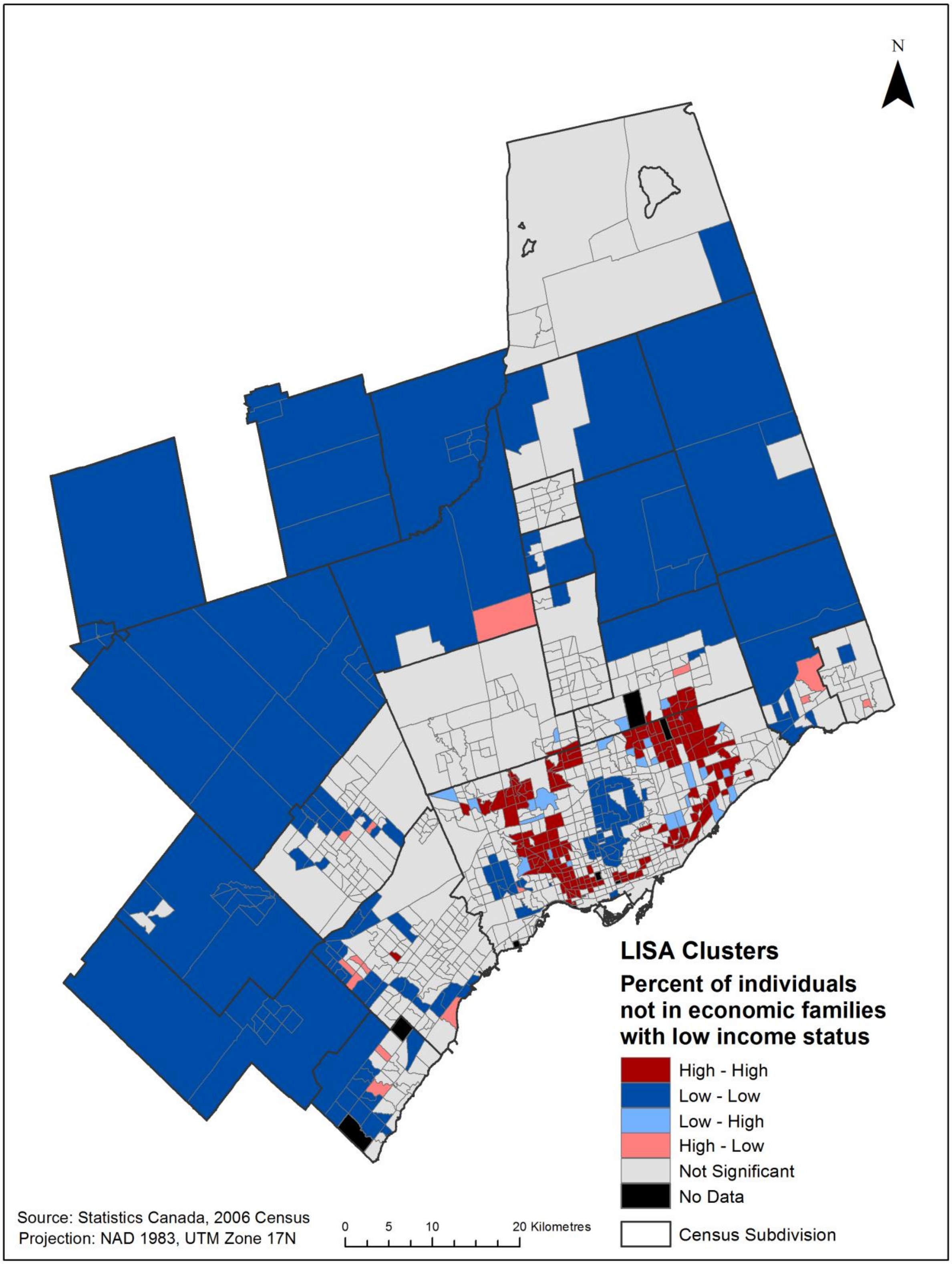

Figure 22: 2006 Toronto CMA LISA Clusters for Percentage of Individuals not in Economic Families with Low Income Status

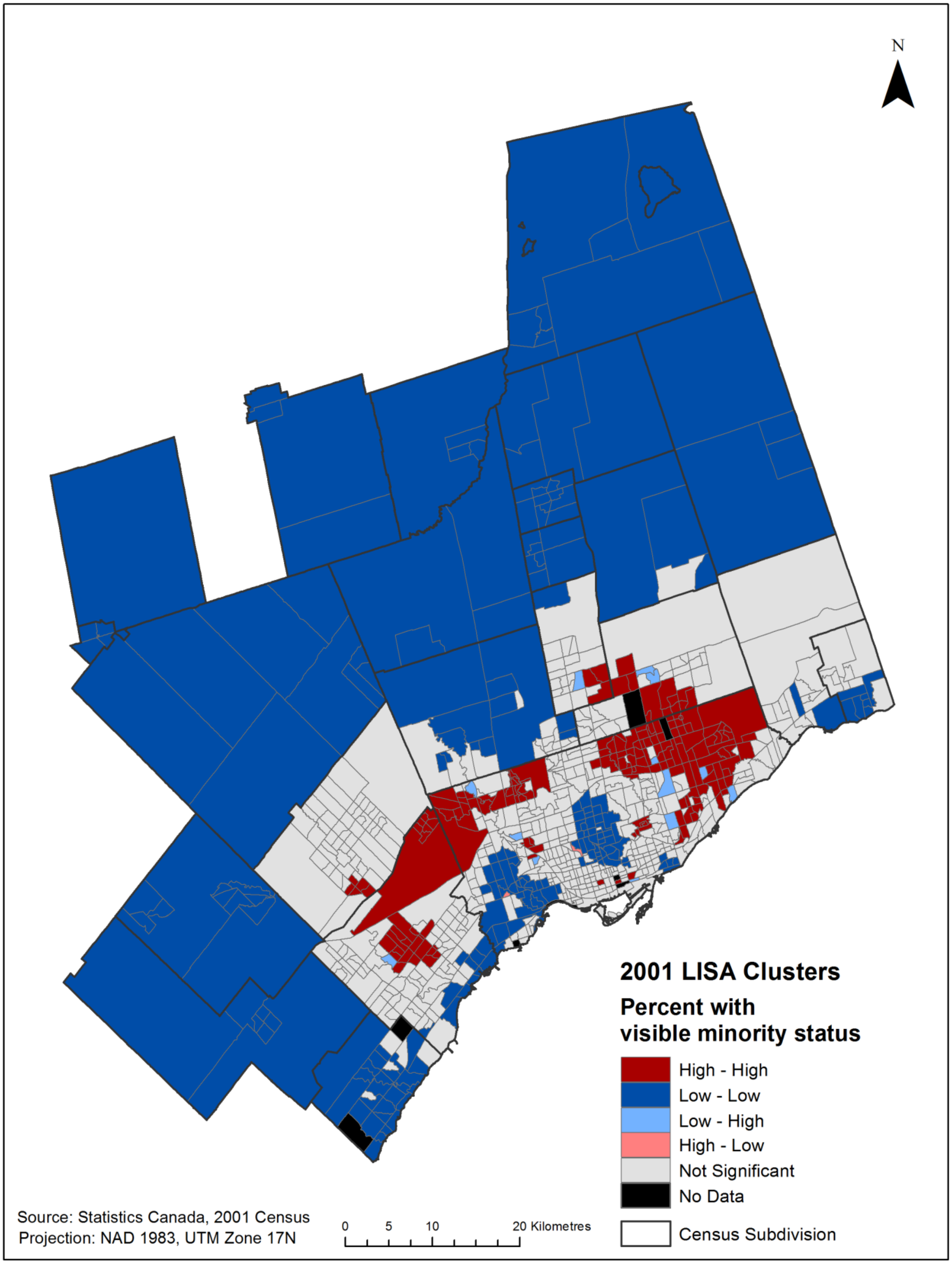

Figure 23: 2001 Toronto CMA LISA Clusters for Percentage with Visible Minority Status

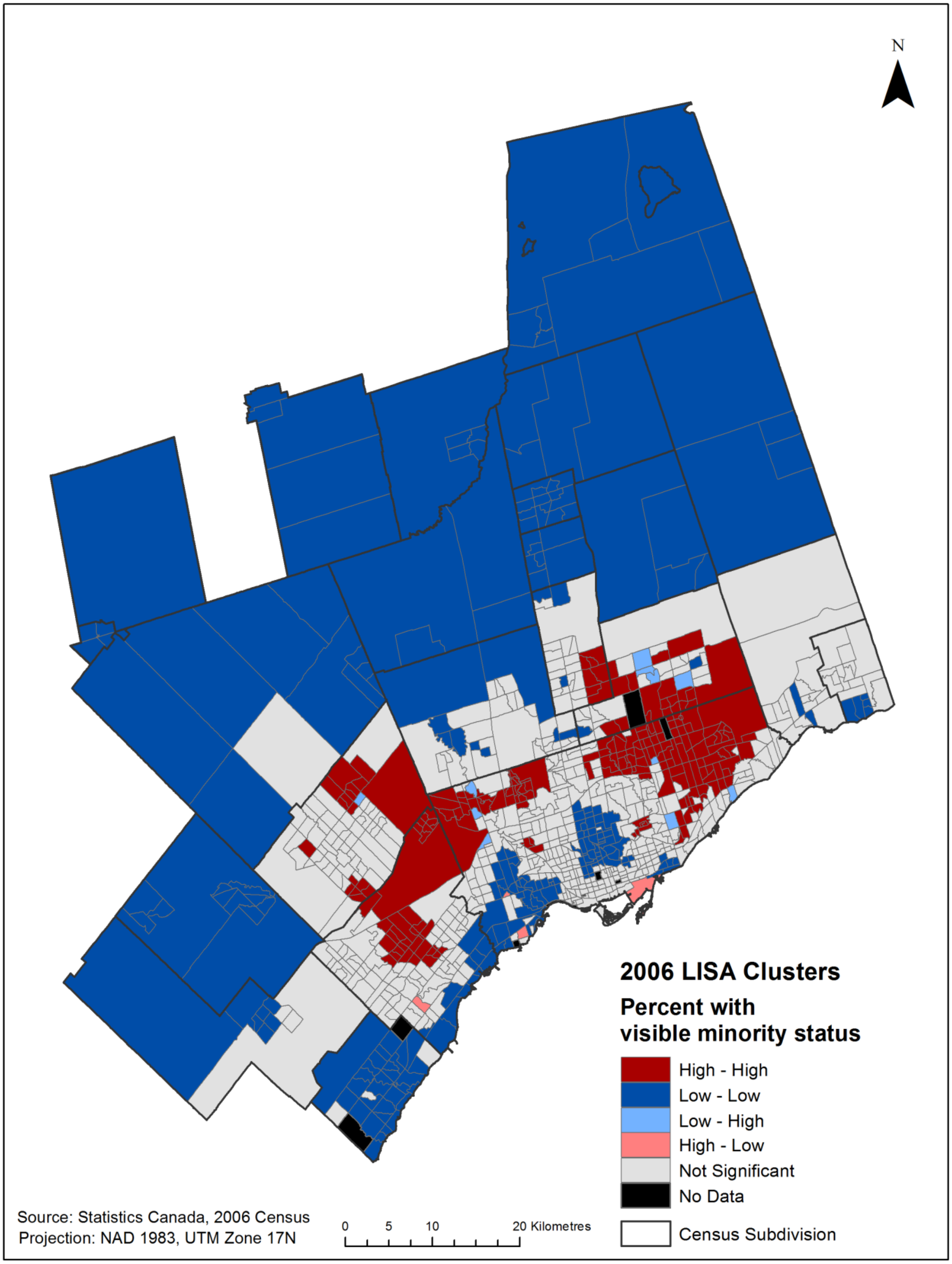

Figure 24: 2006 Toronto CMA LISA Clusters for Percentage with Visible Minority Status

Appendix D: Development of Spatiotemporal Context Scores for Colocation Combinations

The methodology presented in this research depends on the ability to reference scores that have previously been determined for each of the possible spatiotemporal contexts that exist within the scope of this study. For this work, there are four variables that contribute to the spatiotemporal context, these include: percentage of low education, percentage of immigrants, percentage of visible minority, and percentage of low income. Each variable may contribute a state to the spatiotemporal context and this state is determined by the results of tests for spatial autocorrelation. Possible states include: "0" for "Not Significant" (NS), "1" for "High – High" (HH), "2" for "Low – Low" (LL), "3" for "Low – High" (LH), and "4" for "High – Low" (HL). Given that there are four variables with five possible states, and only one state is possible at a particular time, the number of unique spatiotemporal contexts may be calculated as:

$$\binom{5}{1}\binom{5}{1}\binom{5}{1}\binom{5}{1} = 5 \times 5 \times 5 \times 5 = 625$$

Table 12 below uses each variable's state to identify each of the 625 unique combinations of collocated clusters. The state for each variable is recorded in columns 3 through 6 and each combination of the four values is unique. The states for each variable have also been concatenated to create the "colocation identifier" presented in column 2. This identifier is used in the database tables to create a link between spatiotemporal context data and the meaning of that context.

To develop a score for each potential spatiotemporal context that could occur with this set of variables, counts of the number of occurrences of each state have been calculated. For example, if a clustering combination consisted of a "High – High" cluster for each of the low education and low income variables, the tally would be "2". Columns 7 through 11 present the tallies for each of the five states. Scores are then calculated based on these tallies.

"High – High" classifications are indicative of the clustering of greater amounts of more negative phenomena. For example, a higher percentage of low education people in a Census tract with neighbouring Census tracts that also have a higher percentage of low education people would yield a "High – High" cluster. Column 12 uses the number of "High – High" clusters in each possible spatiotemporal context to calculate a score for that context; this is simply the value in column 7 multiplied by 1. These values have been multiplied by a positive number as “High – High” clustering is associated with a higher number of undesirable health outcomes.

“Low – Low” classifications represent the clustering of less negative phenomena. From the perspective of relative numbers, these classifications represent the more favorable places to live (based only on the variable that has been considered). The “Low – Low” clusters work in the same way as the “High – High” clusters except they are based on lower percentages of the phenomena. Column 13 multiplies the number of “Low –Low” clusters identified in column 8 by -1. This value has been made negative as “Low – Low” clustering is associated with a lower number of undesirable health outcomes.

The “simple score” presented in column 14 is the sum of columns 12 and 13. This approach assumes that the impact of each variable is equally weighted and that the less negative impact of one variable can cancel the more negative impact of another variable. The resulting value has a range of -4 to 4 and may be any integer value within this range. This column has been colour-coded; as the score gets higher (implying more negative consequences) the red colour becomes more saturated, and as the score moves lower (implying less negative consequences) the blue becomes more saturated. Contexts with a “simple score” of 0 are shaded in grey and these are reflective of a neutral spatiotemporal context. It should also be noted that scores based on states of “Low – High” or “High – Low” clusters have not been calculated. The implication of these classifications is less clear, and for this research, they are considered to neutral and not included in the calculation of the simple score.

The simple score calculated here has been completed to demonstrate an approach to manage spatiotemporal problems. It is expected that an expert in a health-related field would provide more accurate scores for the 625 possible spatiotemporal contexts based on their expert knowledge if this research was to be used to support a non-academic and professional application.

Table 12: Unique Colocation Cluster Combinations and Simple Scores

(1) C	(2) Colocation Identifier	(3) Low Educ.	(4) Immigrant	(5) Visible Minority	(6) Low Income	(7) Count of HH	(8) Count of LL	(9) Count of LH	(10) Count of HL	(11) Count of NS	(12) HH Score	(13) LL Score	(14) Simple Score
1	1111	1	1	1	1	4	0	0	0	0	4	0	4
2	0111	0	1	1	1	3	0	0	0	1	3	0	3
3	1011	1	0	1	1	3	0	0	0	1	3	0	3
4	1101	1	1	0	1	3	0	0	0	1	3	0	3
5	1110	1	1	1	0	3	0	0	0	1	3	0	3
6	1113	1	1	1	3	3	0	1	0	0	3	0	3
7	1114	1	1	1	4	3	0	0	1	0	3	0	3
8	1131	1	1	3	1	3	0	1	0	0	3	0	3
9	1141	1	1	4	1	3	0	0	1	0	3	0	3
10	1311	1	3	1	1	3	0	1	0	0	3	0	3
11	1411	1	4	1	1	3	0	0	1	0	3	0	3
12	3111	3	1	1	1	3	0	1	0	0	3	0	3
13	4111	4	1	1	1	3	0	0	1	0	3	0	3
14	0011	0	0	1	1	2	0	0	0	2	2	0	2
15	0101	0	1	0	1	2	0	0	0	2	2	0	2
16	0110	0	1	1	0	2	0	0	0	2	2	0	2
17	0113	0	1	1	3	2	0	1	0	1	2	0	2
18	0114	0	1	1	4	2	0	0	1	1	2	0	2
19	0131	0	1	3	1	2	0	1	0	1	2	0	2
20	0141	0	1	4	1	2	0	0	1	1	2	0	2
21	0311	0	3	1	1	2	0	1	0	1	2	0	2
22	0411	0	4	1	1	2	0	0	1	1	2	0	2
23	1001	1	0	0	1	2	0	0	0	2	2	0	2
24	1010	1	0	1	0	2	0	0	0	2	2	0	2
25	1013	1	0	1	3	2	0	1	0	1	2	0	2
26	1014	1	0	1	4	2	0	0	1	1	2	0	2
27	1031	1	0	3	1	2	0	1	0	1	2	0	2
28	1041	1	0	4	1	2	0	0	1	1	2	0	2
29	1100	1	1	0	0	2	0	0	0	2	2	0	2
30	1103	1	1	0	3	2	0	1	0	1	2	0	2
31	1104	1	1	0	4	2	0	0	1	1	2	0	2
32	1112	1	1	1	2	3	1	0	0	0	3	-1	2
33	1121	1	1	2	1	3	1	0	0	0	3	-1	2
34	1130	1	1	3	0	2	0	1	0	1	2	0	2
35	1133	1	1	3	3	2	0	2	0	0	2	0	2
36	1134	1	1	3	4	2	0	1	1	0	2	0	2
37	1140	1	1	4	0	2	0	0	1	1	2	0	2
38	1143	1	1	4	3	2	0	1	1	0	2	0	2
39	1144	1	1	4	4	2	0	0	2	0	2	0	2
40	1211	1	2	1	1	3	1	0	0	0	3	-1	2
41	1301	1	3	0	1	2	0	1	0	1	2	0	2

42	1310	1	3	1	0	2	0	1	0	1	2	0	2
43	1313	1	3	1	3	2	0	2	0	0	2	0	2
44	1314	1	3	1	4	2	0	1	1	0	2	0	2
45	1331	1	3	3	1	2	0	2	0	0	2	0	2
46	1341	1	3	4	1	2	0	1	1	0	2	0	2
47	1401	1	4	0	1	2	0	0	1	1	2	0	2
48	1410	1	4	1	0	2	0	0	1	1	2	0	2
49	1413	1	4	1	3	2	0	1	1	0	2	0	2
50	1414	1	4	1	4	2	0	0	2	0	2	0	2
51	1431	1	4	3	1	2	0	1	1	0	2	0	2
52	1441	1	4	4	1	2	0	0	2	0	2	0	2
53	2111	2	1	1	1	3	1	0	0	0	3	-1	2
54	3011	3	0	1	1	2	0	1	0	1	2	0	2
55	3101	3	1	0	1	2	0	1	0	1	2	0	2
56	3110	3	1	1	0	2	0	1	0	1	2	0	2
57	3113	3	1	1	3	2	0	2	0	0	2	0	2
58	3114	3	1	1	4	2	0	1	1	0	2	0	2
59	3131	3	1	3	1	2	0	2	0	0	2	0	2
60	3141	3	1	4	1	2	0	1	1	0	2	0	2
61	3311	3	3	1	1	2	0	2	0	0	2	0	2
62	3411	3	4	1	1	2	0	1	1	0	2	0	2
63	4011	4	0	1	1	2	0	0	1	1	2	0	2
64	4101	4	1	0	1	2	0	0	1	1	2	0	2
65	4110	4	1	1	0	2	0	0	1	1	2	0	2
66	4113	4	1	1	3	2	0	1	1	0	2	0	2
67	4114	4	1	1	4	2	0	0	2	0	2	0	2
68	4131	4	1	3	1	2	0	1	1	0	2	0	2
69	4141	4	1	4	1	2	0	0	2	0	2	0	2
70	4311	4	3	1	1	2	0	1	1	0	2	0	2
71	4411	4	4	1	1	2	0	0	2	0	2	0	2
72	0001	0	0	0	1	1	0	0	0	3	1	0	1
73	0010	0	0	1	0	1	0	0	0	3	1	0	1
74	0013	0	0	1	3	1	0	1	0	2	1	0	1
75	0014	0	0	1	4	1	0	0	1	2	1	0	1
76	0031	0	0	3	1	1	0	1	0	2	1	0	1
77	0041	0	0	4	1	1	0	0	1	2	1	0	1
78	0100	0	1	0	0	1	0	0	0	3	1	0	1
79	0103	0	1	0	3	1	0	1	0	2	1	0	1
80	0104	0	1	0	4	1	0	0	1	2	1	0	1
81	0112	0	1	1	2	2	1	0	0	1	2	-1	1
82	0121	0	1	2	1	2	1	0	0	1	2	-1	1
83	0130	0	1	3	0	1	0	1	0	2	1	0	1
84	0133	0	1	3	3	1	0	2	0	1	1	0	1
85	0134	0	1	3	4	1	0	1	1	1	1	0	1

86	0140	0	1	4	0	1	0	0	1	2	1	0	1
87	0143	0	1	4	3	1	0	1	1	1	1	0	1
88	0144	0	1	4	4	1	0	0	2	1	1	0	1
89	0211	0	2	1	1	2	1	0	0	1	2	-1	1
90	0301	0	3	0	1	1	0	1	0	2	1	0	1
91	0310	0	3	1	0	1	0	1	0	2	1	0	1
92	0313	0	3	1	3	1	0	2	0	1	1	0	1
93	0314	0	3	1	4	1	0	1	1	1	1	0	1
94	0331	0	3	3	1	1	0	2	0	1	1	0	1
95	0341	0	3	4	1	1	0	1	1	1	1	0	1
96	0401	0	4	0	1	1	0	0	1	2	1	0	1
97	0410	0	4	1	0	1	0	0	1	2	1	0	1
98	0413	0	4	1	3	1	0	1	1	1	1	0	1
99	0414	0	4	1	4	1	0	0	2	1	1	0	1
100	0431	0	4	3	1	1	0	1	1	1	1	0	1
101	0441	0	4	4	1	1	0	0	2	1	1	0	1
102	1000	1	0	0	0	1	0	0	0	3	1	0	1
103	1003	1	0	0	3	1	0	1	0	2	1	0	1
104	1004	1	0	0	4	1	0	0	1	2	1	0	1
105	1012	1	0	1	2	2	1	0	0	1	2	-1	1
106	1021	1	0	2	1	2	1	0	0	1	2	-1	1
107	1030	1	0	3	0	1	0	1	0	2	1	0	1
108	1033	1	0	3	3	1	0	2	0	1	1	0	1
109	1034	1	0	3	4	1	0	1	1	1	1	0	1
110	1040	1	0	4	0	1	0	0	1	2	1	0	1
111	1043	1	0	4	3	1	0	1	1	1	1	0	1
112	1044	1	0	4	4	1	0	0	2	1	1	0	1
113	1102	1	1	0	2	2	1	0	0	1	2	-1	1
114	1120	1	1	2	0	2	1	0	0	1	2	-1	1
115	1123	1	1	2	3	2	1	1	0	0	2	-1	1
116	1124	1	1	2	4	2	1	0	1	0	2	-1	1
117	1132	1	1	3	2	2	1	1	0	0	2	-1	1
118	1142	1	1	4	2	2	1	0	1	0	2	-1	1
119	1201	1	2	0	1	2	1	0	0	1	2	-1	1
120	1210	1	2	1	0	2	1	0	0	1	2	-1	1
121	1213	1	2	1	3	2	1	1	0	0	2	-1	1
122	1214	1	2	1	4	2	1	0	1	0	2	-1	1
123	1231	1	2	3	1	2	1	1	0	0	2	-1	1
124	1241	1	2	4	1	2	1	0	1	0	2	-1	1
125	1300	1	3	0	0	1	0	1	0	2	1	0	1
126	1303	1	3	0	3	1	0	2	0	1	1	0	1
127	1304	1	3	0	4	1	0	1	1	1	1	0	1
128	1312	1	3	1	2	2	1	1	0	0	2	-1	1
129	1321	1	3	2	1	2	1	1	0	0	2	-1	1

130	1330	1	3	3	0	1	0	2	0	1	1	0	1
131	1333	1	3	3	3	1	0	3	0	0	1	0	1
132	1334	1	3	3	4	1	0	2	1	0	1	0	1
133	1340	1	3	4	0	1	0	1	1	1	1	0	1
134	1343	1	3	4	3	1	0	2	1	0	1	0	1
135	1344	1	3	4	4	1	0	1	2	0	1	0	1
136	1400	1	4	0	0	1	0	0	1	2	1	0	1
137	1403	1	4	0	3	1	0	1	1	1	1	0	1
138	1404	1	4	0	4	1	0	0	2	1	1	0	1
139	1412	1	4	1	2	2	1	0	1	0	2	-1	1
140	1421	1	4	2	1	2	1	0	1	0	2	-1	1
141	1430	1	4	3	0	1	0	1	1	1	1	0	1
142	1433	1	4	3	3	1	0	2	1	0	1	0	1
143	1434	1	4	3	4	1	0	1	2	0	1	0	1
144	1440	1	4	4	0	1	0	0	2	1	1	0	1
145	1443	1	4	4	3	1	0	1	2	0	1	0	1
146	1444	1	4	4	4	1	0	0	3	0	1	0	1
147	2011	2	0	1	1	2	1	0	0	1	2	-1	1
148	2101	2	1	0	1	2	1	0	0	1	2	-1	1
149	2110	2	1	1	0	2	1	0	0	1	2	-1	1
150	2113	2	1	1	3	2	1	1	0	0	2	-1	1
151	2114	2	1	1	4	2	1	0	1	0	2	-1	1
152	2131	2	1	3	1	2	1	1	0	0	2	-1	1
153	2141	2	1	4	1	2	1	0	1	0	2	-1	1
154	2311	2	3	1	1	2	1	1	0	0	2	-1	1
155	2411	2	4	1	1	2	1	0	1	0	2	-1	1
156	3001	3	0	0	1	1	0	1	0	2	1	0	1
157	3010	3	0	1	0	1	0	1	0	2	1	0	1
158	3013	3	0	1	3	1	0	2	0	1	1	0	1
159	3014	3	0	1	4	1	0	1	1	1	1	0	1
160	3031	3	0	3	1	1	0	2	0	1	1	0	1
161	3041	3	0	4	1	1	0	1	1	1	1	0	1
162	3100	3	1	0	0	1	0	1	0	2	1	0	1
163	3103	3	1	0	3	1	0	2	0	1	1	0	1
164	3104	3	1	0	4	1	0	1	1	1	1	0	1
165	3112	3	1	1	2	2	1	1	0	0	2	-1	1
166	3121	3	1	2	1	2	1	1	0	0	2	-1	1
167	3130	3	1	3	0	1	0	2	0	1	1	0	1
168	3133	3	1	3	3	1	0	3	0	0	1	0	1
169	3134	3	1	3	4	1	0	2	1	0	1	0	1
170	3140	3	1	4	0	1	0	1	1	1	1	0	1
171	3143	3	1	4	3	1	0	2	1	0	1	0	1
172	3144	3	1	4	4	1	0	1	2	0	1	0	1
173	3211	3	2	1	1	2	1	1	0	0	2	-1	1

174	3301	3	3	0	1	1	0	2	0	1	1	0	1
175	3310	3	3	1	0	1	0	2	0	1	1	0	1
176	3313	3	3	1	3	1	0	3	0	0	1	0	1
177	3314	3	3	1	4	1	0	2	1	0	1	0	1
178	3331	3	3	3	1	1	0	3	0	0	1	0	1
179	3341	3	3	4	1	1	0	2	1	0	1	0	1
180	3401	3	4	0	1	1	0	1	1	1	1	0	1
181	3410	3	4	1	0	1	0	1	1	1	1	0	1
182	3413	3	4	1	3	1	0	2	1	0	1	0	1
183	3414	3	4	1	4	1	0	1	2	0	1	0	1
184	3431	3	4	3	1	1	0	2	1	0	1	0	1
185	3441	3	4	4	1	1	0	1	2	0	1	0	1
186	4001	4	0	0	1	1	0	0	1	2	1	0	1
187	4010	4	0	1	0	1	0	0	1	2	1	0	1
188	4013	4	0	1	3	1	0	1	1	1	1	0	1
189	4014	4	0	1	4	1	0	0	2	1	1	0	1
190	4031	4	0	3	1	1	0	1	1	1	1	0	1
191	4041	4	0	4	1	1	0	0	2	1	1	0	1
192	4100	4	1	0	0	1	0	0	1	2	1	0	1
193	4103	4	1	0	3	1	0	1	1	1	1	0	1
194	4104	4	1	0	4	1	0	0	2	1	1	0	1
195	4112	4	1	1	2	2	1	0	1	0	2	-1	1
196	4121	4	1	2	1	2	1	0	1	0	2	-1	1
197	4130	4	1	3	0	1	0	1	1	1	1	0	1
198	4133	4	1	3	3	1	0	2	1	0	1	0	1
199	4134	4	1	3	4	1	0	1	2	0	1	0	1
200	4140	4	1	4	0	1	0	0	2	1	1	0	1
201	4143	4	1	4	3	1	0	1	2	0	1	0	1
202	4144	4	1	4	4	1	0	0	3	0	1	0	1
203	4211	4	2	1	1	2	1	0	1	0	2	-1	1
204	4301	4	3	0	1	1	0	1	1	1	1	0	1
205	4310	4	3	1	0	1	0	1	1	1	1	0	1
206	4313	4	3	1	3	1	0	2	1	0	1	0	1
207	4314	4	3	1	4	1	0	1	2	0	1	0	1
208	4331	4	3	3	1	1	0	2	1	0	1	0	1
209	4341	4	3	4	1	1	0	1	2	0	1	0	1
210	4401	4	4	0	1	1	0	0	2	1	1	0	1
211	4410	4	4	1	0	1	0	0	2	1	1	0	1
212	4413	4	4	1	3	1	0	1	2	0	1	0	1
213	4414	4	4	1	4	1	0	0	3	0	1	0	1
214	4431	4	4	3	1	1	0	1	2	0	1	0	1
215	4441	4	4	4	1	1	0	0	3	0	1	0	1
216	0000	0	0	0	0	0	0	0	0	4	0	0	0
217	0003	0	0	0	3	0	0	1	0	3	0	0	0

218	0004	0	0	0	4	0	0	0	1	3	0	0	0
219	0012	0	0	1	2	1	1	0	0	2	1	-1	0
220	0021	0	0	2	1	1	1	0	0	2	1	-1	0
221	0030	0	0	3	0	0	0	1	0	3	0	0	0
222	0033	0	0	3	3	0	0	2	0	2	0	0	0
223	0034	0	0	3	4	0	0	1	1	2	0	0	0
224	0040	0	0	4	0	0	0	0	1	3	0	0	0
225	0043	0	0	4	3	0	0	1	1	2	0	0	0
226	0044	0	0	4	4	0	0	0	2	2	0	0	0
227	0102	0	1	0	2	1	1	0	0	2	1	-1	0
228	0120	0	1	2	0	1	1	0	0	2	1	-1	0
229	0123	0	1	2	3	1	1	1	0	1	1	-1	0
230	0124	0	1	2	4	1	1	0	1	1	1	-1	0
231	0132	0	1	3	2	1	1	1	0	1	1	-1	0
232	0142	0	1	4	2	1	1	0	1	1	1	-1	0
233	0201	0	2	0	1	1	1	0	0	2	1	-1	0
234	0210	0	2	1	0	1	1	0	0	2	1	-1	0
235	0213	0	2	1	3	1	1	1	0	1	1	-1	0
236	0214	0	2	1	4	1	1	0	1	1	1	-1	0
237	0231	0	2	3	1	1	1	1	0	1	1	-1	0
238	0241	0	2	4	1	1	1	0	1	1	1	-1	0
239	0300	0	3	0	0	0	0	1	0	3	0	0	0
240	0303	0	3	0	3	0	0	2	0	2	0	0	0
241	0304	0	3	0	4	0	0	1	1	2	0	0	0
242	0312	0	3	1	2	1	1	1	0	1	1	-1	0
243	0321	0	3	2	1	1	1	1	0	1	1	-1	0
244	0330	0	3	3	0	0	0	2	0	2	0	0	0
245	0333	0	3	3	3	0	0	3	0	1	0	0	0
246	0334	0	3	3	4	0	0	2	1	1	0	0	0
247	0340	0	3	4	0	0	0	1	1	2	0	0	0
248	0343	0	3	4	3	0	0	2	1	1	0	0	0
249	0344	0	3	4	4	0	0	1	2	1	0	0	0
250	0400	0	4	0	0	0	0	0	1	3	0	0	0
251	0403	0	4	0	3	0	0	1	1	2	0	0	0
252	0404	0	4	0	4	0	0	0	2	2	0	0	0
253	0412	0	4	1	2	1	1	0	1	1	1	-1	0
254	0421	0	4	2	1	1	1	0	1	1	1	-1	0
255	0430	0	4	3	0	0	0	1	1	2	0	0	0
256	0433	0	4	3	3	0	0	2	1	1	0	0	0
257	0434	0	4	3	4	0	0	1	2	1	0	0	0
258	0440	0	4	4	0	0	0	0	2	2	0	0	0
259	0443	0	4	4	3	0	0	1	2	1	0	0	0
260	0444	0	4	4	4	0	0	0	3	1	0	0	0
261	1002	1	0	0	2	1	1	0	0	2	1	-1	0

262	1020	1	0	2	0	1	1	0	0	2	1	-1	0
263	1023	1	0	2	3	1	1	1	0	1	1	-1	0
264	1024	1	0	2	4	1	1	0	1	1	1	-1	0
265	1032	1	0	3	2	1	1	1	0	1	1	-1	0
266	1042	1	0	4	2	1	1	0	1	1	1	-1	0
267	1122	1	1	2	2	2	2	0	0	0	2	-2	0
268	1200	1	2	0	0	1	1	0	0	2	1	-1	0
269	1203	1	2	0	3	1	1	1	0	1	1	-1	0
270	1204	1	2	0	4	1	1	0	1	1	1	-1	0
271	1212	1	2	1	2	2	2	0	0	0	2	-2	0
272	1221	1	2	2	1	2	2	0	0	0	2	-2	0
273	1230	1	2	3	0	1	1	1	0	1	1	-1	0
274	1233	1	2	3	3	1	1	2	0	0	1	-1	0
275	1234	1	2	3	4	1	1	1	1	0	1	-1	0
276	1240	1	2	4	0	1	1	0	1	1	1	-1	0
277	1243	1	2	4	3	1	1	1	1	0	1	-1	0
278	1244	1	2	4	4	1	1	0	2	0	1	-1	0
279	1302	1	3	0	2	1	1	1	0	1	1	-1	0
280	1320	1	3	2	0	1	1	1	0	1	1	-1	0
281	1323	1	3	2	3	1	1	2	0	0	1	-1	0
282	1324	1	3	2	4	1	1	1	1	0	1	-1	0
283	1332	1	3	3	2	1	1	2	0	0	1	-1	0
284	1342	1	3	4	2	1	1	1	1	0	1	-1	0
285	1402	1	4	0	2	1	1	0	1	1	1	-1	0
286	1420	1	4	2	0	1	1	0	1	1	1	-1	0
287	1423	1	4	2	3	1	1	1	1	0	1	-1	0
288	1424	1	4	2	4	1	1	0	2	0	1	-1	0
289	1432	1	4	3	2	1	1	1	1	0	1	-1	0
290	1442	1	4	4	2	1	1	0	2	0	1	-1	0
291	2001	2	0	0	1	1	1	0	0	2	1	-1	0
292	2010	2	0	1	0	1	1	0	0	2	1	-1	0
293	2013	2	0	1	3	1	1	1	0	1	1	-1	0
294	2014	2	0	1	4	1	1	0	1	1	1	-1	0
295	2031	2	0	3	1	1	1	1	0	1	1	-1	0
296	2041	2	0	4	1	1	1	0	1	1	1	-1	0
297	2100	2	1	0	0	1	1	0	0	2	1	-1	0
298	2103	2	1	0	3	1	1	1	0	1	1	-1	0
299	2104	2	1	0	4	1	1	0	1	1	1	-1	0
300	2112	2	1	1	2	2	2	0	0	0	2	-2	0
301	2121	2	1	2	1	2	2	0	0	0	2	-2	0
302	2130	2	1	3	0	1	1	1	0	1	1	-1	0
303	2133	2	1	3	3	1	1	2	0	0	1	-1	0
304	2134	2	1	3	4	1	1	1	1	0	1	-1	0
305	2140	2	1	4	0	1	1	0	1	1	1	-1	0

306	2143	2	1	4	3	1	1	1	1	0	1	-1	0
307	2144	2	1	4	4	1	1	0	2	0	1	-1	0
308	2211	2	2	1	1	2	2	0	0	0	2	-2	0
309	2301	2	3	0	1	1	1	1	0	1	1	-1	0
310	2310	2	3	1	0	1	1	1	0	1	1	-1	0
311	2313	2	3	1	3	1	1	2	0	0	1	-1	0
312	2314	2	3	1	4	1	1	1	1	0	1	-1	0
313	2331	2	3	3	1	1	1	2	0	0	1	-1	0
314	2341	2	3	4	1	1	1	1	1	0	1	-1	0
315	2401	2	4	0	1	1	1	0	1	1	1	-1	0
316	2410	2	4	1	0	1	1	0	1	1	1	-1	0
317	2413	2	4	1	3	1	1	1	1	0	1	-1	0
318	2414	2	4	1	4	1	1	0	2	0	1	-1	0
319	2431	2	4	3	1	1	1	1	1	0	1	-1	0
320	2441	2	4	4	1	1	1	0	2	0	1	-1	0
321	3000	3	0	0	0	0	0	1	0	3	0	0	0
322	3003	3	0	0	3	0	0	2	0	2	0	0	0
323	3004	3	0	0	4	0	0	1	1	2	0	0	0
324	3012	3	0	1	2	1	1	1	0	1	1	-1	0
325	3021	3	0	2	1	1	1	1	0	1	1	-1	0
326	3030	3	0	3	0	0	0	2	0	2	0	0	0
327	3033	3	0	3	3	0	0	3	0	1	0	0	0
328	3034	3	0	3	4	0	0	2	1	1	0	0	0
329	3040	3	0	4	0	0	0	1	1	2	0	0	0
330	3043	3	0	4	3	0	0	2	1	1	0	0	0
331	3044	3	0	4	4	0	0	1	2	1	0	0	0
332	3102	3	1	0	2	1	1	1	0	1	1	-1	0
333	3120	3	1	2	0	1	1	1	0	1	1	-1	0
334	3123	3	1	2	3	1	1	2	0	0	1	-1	0
335	3124	3	1	2	4	1	1	1	1	0	1	-1	0
336	3132	3	1	3	2	1	1	2	0	0	1	-1	0
337	3142	3	1	4	2	1	1	1	1	0	1	-1	0
338	3201	3	2	0	1	1	1	1	0	1	1	-1	0
339	3210	3	2	1	0	1	1	1	0	1	1	-1	0
340	3213	3	2	1	3	1	1	2	0	0	1	-1	0
341	3214	3	2	1	4	1	1	1	1	0	1	-1	0
342	3231	3	2	3	1	1	1	2	0	0	1	-1	0
343	3241	3	2	4	1	1	1	1	1	0	1	-1	0
344	3300	3	3	0	0	0	0	2	0	2	0	0	0
345	3303	3	3	0	3	0	0	3	0	1	0	0	0
346	3304	3	3	0	4	0	0	2	1	1	0	0	0
347	3312	3	3	1	2	1	1	2	0	0	1	-1	0
348	3321	3	3	2	1	1	1	2	0	0	1	-1	0
349	3330	3	3	3	0	0	0	3	0	1	0	0	0

350	3333	3	3	3	3	0	0	4	0	0	0	0	0
351	3334	3	3	3	4	0	0	3	1	0	0	0	0
352	3340	3	3	4	0	0	0	2	1	1	0	0	0
353	3343	3	3	4	3	0	0	3	1	0	0	0	0
354	3344	3	3	4	4	0	0	2	2	0	0	0	0
355	3400	3	4	0	0	0	0	1	1	2	0	0	0
356	3403	3	4	0	3	0	0	2	1	1	0	0	0
357	3404	3	4	0	4	0	0	1	2	1	0	0	0
358	3412	3	4	1	2	1	1	1	1	0	1	-1	0
359	3421	3	4	2	1	1	1	1	1	0	1	-1	0
360	3430	3	4	3	0	0	0	2	1	1	0	0	0
361	3433	3	4	3	3	0	0	3	1	0	0	0	0
362	3434	3	4	3	4	0	0	2	2	0	0	0	0
363	3440	3	4	4	0	0	0	1	2	1	0	0	0
364	3443	3	4	4	3	0	0	2	2	0	0	0	0
365	3444	3	4	4	4	0	0	1	3	0	0	0	0
366	4000	4	0	0	0	0	0	0	1	3	0	0	0
367	4003	4	0	0	3	0	0	1	1	2	0	0	0
368	4004	4	0	0	4	0	0	0	2	2	0	0	0
369	4012	4	0	1	2	1	1	0	1	1	1	-1	0
370	4021	4	0	2	1	1	1	0	1	1	1	-1	0
371	4030	4	0	3	0	0	0	1	1	2	0	0	0
372	4033	4	0	3	3	0	0	2	1	1	0	0	0
373	4034	4	0	3	4	0	0	1	2	1	0	0	0
374	4040	4	0	4	0	0	0	0	2	2	0	0	0
375	4043	4	0	4	3	0	0	1	2	1	0	0	0
376	4044	4	0	4	4	0	0	0	3	1	0	0	0
377	4102	4	1	0	2	1	1	0	1	1	1	-1	0
378	4120	4	1	2	0	1	1	0	1	1	1	-1	0
379	4123	4	1	2	3	1	1	1	1	0	1	-1	0
380	4124	4	1	2	4	1	1	0	2	0	1	-1	0
381	4132	4	1	3	2	1	1	1	1	0	1	-1	0
382	4142	4	1	4	2	1	1	0	2	0	1	-1	0
383	4201	4	2	0	1	1	1	0	1	1	1	-1	0
384	4210	4	2	1	0	1	1	0	1	1	1	-1	0
385	4213	4	2	1	3	1	1	1	1	0	1	-1	0
386	4214	4	2	1	4	1	1	0	2	0	1	-1	0
387	4231	4	2	3	1	1	1	1	1	0	1	-1	0
388	4241	4	2	4	1	1	1	0	2	0	1	-1	0
389	4300	4	3	0	0	0	0	1	1	2	0	0	0
390	4303	4	3	0	3	0	0	2	1	1	0	0	0
391	4304	4	3	0	4	0	0	1	2	1	0	0	0
392	4312	4	3	1	2	1	1	1	1	0	1	-1	0
393	4321	4	3	2	1	1	1	1	1	0	1	-1	0

394	4330	4	3	3	0	0	0	2	1	1	0	0	0
395	4333	4	3	3	3	0	0	3	1	0	0	0	0
396	4334	4	3	3	4	0	0	2	2	0	0	0	0
397	4340	4	3	4	0	0	0	1	2	1	0	0	0
398	4343	4	3	4	3	0	0	2	2	0	0	0	0
399	4344	4	3	4	4	0	0	1	3	0	0	0	0
400	4400	4	4	0	0	0	0	0	2	2	0	0	0
401	4403	4	4	0	3	0	0	1	2	1	0	0	0
402	4404	4	4	0	4	0	0	0	3	1	0	0	0
403	4412	4	4	1	2	1	1	0	2	0	1	-1	0
404	4421	4	4	2	1	1	1	0	2	0	1	-1	0
405	4430	4	4	3	0	0	0	1	2	1	0	0	0
406	4433	4	4	3	3	0	0	2	2	0	0	0	0
407	4434	4	4	3	4	0	0	1	3	0	0	0	0
408	4440	4	4	4	0	0	0	0	3	1	0	0	0
409	4443	4	4	4	3	0	0	1	3	0	0	0	0
410	4444	4	4	4	4	0	0	0	4	0	0	0	0
411	0002	0	0	0	2	0	1	0	0	3	0	-1	-1
412	0020	0	0	2	0	0	1	0	0	3	0	-1	-1
413	0023	0	0	2	3	0	1	1	0	2	0	-1	-1
414	0024	0	0	2	4	0	1	0	1	2	0	-1	-1
415	0032	0	0	3	2	0	1	1	0	2	0	-1	-1
416	0042	0	0	4	2	0	1	0	1	2	0	-1	-1
417	0122	0	1	2	2	1	2	0	0	1	1	-2	-1
418	0200	0	2	0	0	0	1	0	0	3	0	-1	-1
419	0203	0	2	0	3	0	1	1	0	2	0	-1	-1
420	0204	0	2	0	4	0	1	0	1	2	0	-1	-1
421	0212	0	2	1	2	1	2	0	0	1	1	-2	-1
422	0221	0	2	2	1	1	2	0	0	1	1	-2	-1
423	0230	0	2	3	0	0	1	1	0	2	0	-1	-1
424	0233	0	2	3	3	0	1	2	0	1	0	-1	-1
425	0234	0	2	3	4	0	1	1	1	1	0	-1	-1
426	0240	0	2	4	0	0	1	0	1	2	0	-1	-1
427	0243	0	2	4	3	0	1	1	1	1	0	-1	-1
428	0244	0	2	4	4	0	1	0	2	1	0	-1	-1
429	0302	0	3	0	2	0	1	1	0	2	0	-1	-1
430	0320	0	3	2	0	0	1	1	0	2	0	-1	-1
431	0323	0	3	2	3	0	1	2	0	1	0	-1	-1
432	0324	0	3	2	4	0	1	1	1	1	0	-1	-1
433	0332	0	3	3	2	0	1	2	0	1	0	-1	-1
434	0342	0	3	4	2	0	1	1	1	1	0	-1	-1
435	0402	0	4	0	2	0	1	0	1	2	0	-1	-1
436	0420	0	4	2	0	0	1	0	1	2	0	-1	-1
437	0423	0	4	2	3	0	1	1	1	1	0	-1	-1

438	0424	0	4	2	4	0	1	0	2	1	0	-1	-1
439	0432	0	4	3	2	0	1	1	1	1	0	-1	-1
440	0442	0	4	4	2	0	1	0	2	1	0	-1	-1
441	1022	1	0	2	2	1	2	0	0	1	1	-2	-1
442	1202	1	2	0	2	1	2	0	0	1	1	-2	-1
443	1220	1	2	2	0	1	2	0	0	1	1	-2	-1
444	1223	1	2	2	3	1	2	1	0	0	1	-2	-1
445	1224	1	2	2	4	1	2	0	1	0	1	-2	-1
446	1232	1	2	3	2	1	2	1	0	0	1	-2	-1
447	1242	1	2	4	2	1	2	0	1	0	1	-2	-1
448	1322	1	3	2	2	1	2	1	0	0	1	-2	-1
449	1422	1	4	2	2	1	2	0	1	0	1	-2	-1
450	2000	2	0	0	0	0	1	0	0	3	0	-1	-1
451	2003	2	0	0	3	0	1	1	0	2	0	-1	-1
452	2004	2	0	0	4	0	1	0	1	2	0	-1	-1
453	2012	2	0	1	2	1	2	0	0	1	1	-2	-1
454	2021	2	0	2	1	1	2	0	0	1	1	-2	-1
455	2030	2	0	3	0	0	1	1	0	2	0	-1	-1
456	2033	2	0	3	3	0	1	2	0	1	0	-1	-1
457	2034	2	0	3	4	0	1	1	1	1	0	-1	-1
458	2040	2	0	4	0	0	1	0	1	2	0	-1	-1
459	2043	2	0	4	3	0	1	1	1	1	0	-1	-1
460	2044	2	0	4	4	0	1	0	2	1	0	-1	-1
461	2102	2	1	0	2	1	2	0	0	1	1	-2	-1
462	2120	2	1	2	0	1	2	0	0	1	1	-2	-1
463	2123	2	1	2	3	1	2	1	0	0	1	-2	-1
464	2124	2	1	2	4	1	2	0	1	0	1	-2	-1
465	2132	2	1	3	2	1	2	1	0	0	1	-2	-1
466	2142	2	1	4	2	1	2	0	1	0	1	-2	-1
467	2201	2	2	0	1	1	2	0	0	1	1	-2	-1
468	2210	2	2	1	0	1	2	0	0	1	1	-2	-1
469	2213	2	2	1	3	1	2	1	0	0	1	-2	-1
470	2214	2	2	1	4	1	2	0	1	0	1	-2	-1
471	2231	2	2	3	1	1	2	1	0	0	1	-2	-1
472	2241	2	2	4	1	1	2	0	1	0	1	-2	-1
473	2300	2	3	0	0	0	1	1	0	2	0	-1	-1
474	2303	2	3	0	3	0	1	2	0	1	0	-1	-1
475	2304	2	3	0	4	0	1	1	1	1	0	-1	-1
476	2312	2	3	1	2	1	2	1	0	0	1	-2	-1
477	2321	2	3	2	1	1	2	1	0	0	1	-2	-1
478	2330	2	3	3	0	0	1	2	0	1	0	-1	-1
479	2333	2	3	3	3	0	1	3	0	0	0	-1	-1
480	2334	2	3	3	4	0	1	2	1	0	0	-1	-1
481	2340	2	3	4	0	0	1	1	1	1	0	-1	-1

482	2343	2	3	4	3	0	1	2	1	0	0	-1	-1
483	2344	2	3	4	4	0	1	1	2	0	0	-1	-1
484	2400	2	4	0	0	0	1	0	1	2	0	-1	-1
485	2403	2	4	0	3	0	1	1	1	1	0	-1	-1
486	2404	2	4	0	4	0	1	0	2	1	0	-1	-1
487	2412	2	4	1	2	1	2	0	1	0	1	-2	-1
488	2421	2	4	2	1	1	2	0	1	0	1	-2	-1
489	2430	2	4	3	0	0	1	1	1	1	0	-1	-1
490	2433	2	4	3	3	0	1	2	1	0	0	-1	-1
491	2434	2	4	3	4	0	1	1	2	0	0	-1	-1
492	2440	2	4	4	0	0	1	0	2	1	0	-1	-1
493	2443	2	4	4	3	0	1	1	2	0	0	-1	-1
494	2444	2	4	4	4	0	1	0	3	0	0	-1	-1
495	3002	3	0	0	2	0	1	1	0	2	0	-1	-1
496	3020	3	0	2	0	0	1	1	0	2	0	-1	-1
497	3023	3	0	2	3	0	1	2	0	1	0	-1	-1
498	3024	3	0	2	4	0	1	1	1	1	0	-1	-1
499	3032	3	0	3	2	0	1	2	0	1	0	-1	-1
500	3042	3	0	4	2	0	1	1	1	1	0	-1	-1
501	3122	3	1	2	2	1	2	1	0	0	1	-2	-1
502	3200	3	2	0	0	0	1	1	0	2	0	-1	-1
503	3203	3	2	0	3	0	1	2	0	1	0	-1	-1
504	3204	3	2	0	4	0	1	1	1	1	0	-1	-1
505	3212	3	2	1	2	1	2	1	0	0	1	-2	-1
506	3221	3	2	2	1	1	2	1	0	0	1	-2	-1
507	3230	3	2	3	0	0	1	2	0	1	0	-1	-1
508	3233	3	2	3	3	0	1	3	0	0	0	-1	-1
509	3234	3	2	3	4	0	1	2	1	0	0	-1	-1
510	3240	3	2	4	0	0	1	1	1	1	0	-1	-1
511	3243	3	2	4	3	0	1	2	1	0	0	-1	-1
512	3244	3	2	4	4	0	1	1	2	0	0	-1	-1
513	3302	3	3	0	2	0	1	2	0	1	0	-1	-1
514	3320	3	3	2	0	0	1	2	0	1	0	-1	-1
515	3323	3	3	2	3	0	1	3	0	0	0	-1	-1
516	3324	3	3	2	4	0	1	2	1	0	0	-1	-1
517	3332	3	3	3	2	0	1	3	0	0	0	-1	-1
518	3342	3	3	4	2	0	1	2	1	0	0	-1	-1
519	3402	3	4	0	2	0	1	1	1	1	0	-1	-1
520	3420	3	4	2	0	0	1	1	1	1	0	-1	-1
521	3423	3	4	2	3	0	1	2	1	0	0	-1	-1
522	3424	3	4	2	4	0	1	1	2	0	0	-1	-1
523	3432	3	4	3	2	0	1	2	1	0	0	-1	-1
524	3442	3	4	4	2	0	1	1	2	0	0	-1	-1
525	4002	4	0	0	2	0	1	0	1	2	0	-1	-1

526	4020	4	0	2	0	0	1	0	1	2	0	-1	-1
527	4023	4	0	2	3	0	1	1	1	1	0	-1	-1
528	4024	4	0	2	4	0	1	0	2	1	0	-1	-1
529	4032	4	0	3	2	0	1	1	1	1	0	-1	-1
530	4042	4	0	4	2	0	1	0	2	1	0	-1	-1
531	4122	4	1	2	2	1	2	0	1	0	1	-2	-1
532	4200	4	2	0	0	0	1	0	1	2	0	-1	-1
533	4203	4	2	0	3	0	1	1	1	1	0	-1	-1
534	4204	4	2	0	4	0	1	0	2	1	0	-1	-1
535	4212	4	2	1	2	1	2	0	1	0	1	-2	-1
536	4221	4	2	2	1	1	2	0	1	0	1	-2	-1
537	4230	4	2	3	0	0	1	1	1	1	0	-1	-1
538	4233	4	2	3	3	0	1	2	1	0	0	-1	-1
539	4234	4	2	3	4	0	1	1	2	0	0	-1	-1
540	4240	4	2	4	0	0	1	0	2	1	0	-1	-1
541	4243	4	2	4	3	0	1	1	2	0	0	-1	-1
542	4244	4	2	4	4	0	1	0	3	0	0	-1	-1
543	4302	4	3	0	2	0	1	1	1	1	0	-1	-1
544	4320	4	3	2	0	0	1	1	1	1	0	-1	-1
545	4323	4	3	2	3	0	1	2	1	0	0	-1	-1
546	4324	4	3	2	4	0	1	1	2	0	0	-1	-1
547	4332	4	3	3	2	0	1	2	1	0	0	-1	-1
548	4342	4	3	4	2	0	1	1	2	0	0	-1	-1
549	4402	4	4	0	2	0	1	0	2	1	0	-1	-1
550	4420	4	4	2	0	0	1	0	2	1	0	-1	-1
551	4423	4	4	2	3	0	1	1	2	0	0	-1	-1
552	4424	4	4	2	4	0	1	0	3	0	0	-1	-1
553	4432	4	4	3	2	0	1	1	2	0	0	-1	-1
554	4442	4	4	4	2	0	1	0	3	0	0	-1	-1
555	0022	0	0	2	2	0	2	0	0	2	0	-2	-2
556	0202	0	2	0	2	0	2	0	0	2	0	-2	-2
557	0220	0	2	2	0	0	2	0	0	2	0	-2	-2
558	0223	0	2	2	3	0	2	1	0	1	0	-2	-2
559	0224	0	2	2	4	0	2	0	1	1	0	-2	-2
560	0232	0	2	3	2	0	2	1	0	1	0	-2	-2
561	0242	0	2	4	2	0	2	0	1	1	0	-2	-2
562	0322	0	3	2	2	0	2	1	0	1	0	-2	-2
563	0422	0	4	2	2	0	2	0	1	1	0	-2	-2
564	1222	1	2	2	2	1	3	0	0	0	1	-3	-2
565	2002	2	0	0	2	0	2	0	0	2	0	-2	-2
566	2020	2	0	2	0	0	2	0	0	2	0	-2	-2
567	2023	2	0	2	3	0	2	1	0	1	0	-2	-2
568	2024	2	0	2	4	0	2	0	1	1	0	-2	-2
569	2032	2	0	3	2	0	2	1	0	1	0	-2	-2

570	2042	2	0	4	2	0	2	0	1	1	0	-2	-2
571	2122	2	1	2	2	1	3	0	0	0	1	-3	-2
572	2200	2	2	0	0	0	2	0	0	2	0	-2	-2
573	2203	2	2	0	3	0	2	1	0	1	0	-2	-2
574	2204	2	2	0	4	0	2	0	1	1	0	-2	-2
575	2212	2	2	1	2	1	3	0	0	0	1	-3	-2
576	2221	2	2	2	1	1	3	0	0	0	1	-3	-2
577	2230	2	2	3	0	0	2	1	0	1	0	-2	-2
578	2233	2	2	3	3	0	2	2	0	0	0	-2	-2
579	2234	2	2	3	4	0	2	1	1	0	0	-2	-2
580	2240	2	2	4	0	0	2	0	1	1	0	-2	-2
581	2243	2	2	4	3	0	2	1	1	0	0	-2	-2
582	2244	2	2	4	4	0	2	0	2	0	0	-2	-2
583	2302	2	3	0	2	0	2	1	0	1	0	-2	-2
584	2320	2	3	2	0	0	2	1	0	1	0	-2	-2
585	2323	2	3	2	3	0	2	2	0	0	0	-2	-2
586	2324	2	3	2	4	0	2	1	1	0	0	-2	-2
587	2332	2	3	3	2	0	2	2	0	0	0	-2	-2
588	2342	2	3	4	2	0	2	1	1	0	0	-2	-2
589	2402	2	4	0	2	0	2	0	1	1	0	-2	-2
590	2420	2	4	2	0	0	2	0	1	1	0	-2	-2
591	2423	2	4	2	3	0	2	1	1	0	0	-2	-2
592	2424	2	4	2	4	0	2	0	2	0	0	-2	-2
593	2432	2	4	3	2	0	2	1	1	0	0	-2	-2
594	2442	2	4	4	2	0	2	0	2	0	0	-2	-2
595	3022	3	0	2	2	0	2	1	0	1	0	-2	-2
596	3202	3	2	0	2	0	2	1	0	1	0	-2	-2
597	3220	3	2	2	0	0	2	1	0	1	0	-2	-2
598	3223	3	2	2	3	0	2	2	0	0	0	-2	-2
599	3224	3	2	2	4	0	2	1	1	0	0	-2	-2
600	3232	3	2	3	2	0	2	2	0	0	0	-2	-2
601	3242	3	2	4	2	0	2	1	1	0	0	-2	-2
602	3322	3	3	2	2	0	2	2	0	0	0	-2	-2
603	3422	3	4	2	2	0	2	1	1	0	0	-2	-2
604	4022	4	0	2	2	0	2	0	1	1	0	-2	-2
605	4202	4	2	0	2	0	2	0	1	1	0	-2	-2
606	4220	4	2	2	0	0	2	0	1	1	0	-2	-2
607	4223	4	2	2	3	0	2	1	1	0	0	-2	-2
608	4224	4	2	2	4	0	2	0	2	0	0	-2	-2
609	4232	4	2	3	2	0	2	1	1	0	0	-2	-2
610	4242	4	2	4	2	0	2	0	2	0	0	-2	-2
611	4322	4	3	2	2	0	2	1	1	0	0	-2	-2
612	4422	4	4	2	2	0	2	0	2	0	0	-2	-2
613	0222	0	2	2	2	0	3	0	0	1	0	-3	-3

614	2022	2	0	2	2	0	3	0	0	1	0	-3	-3
615	2202	2	2	0	2	0	3	0	0	1	0	-3	-3
616	2220	2	2	2	0	0	3	0	0	1	0	-3	-3
617	2223	2	2	2	3	0	3	1	0	0	0	-3	-3
618	2224	2	2	2	4	0	3	0	1	0	0	-3	-3
619	2232	2	2	3	2	0	3	1	0	0	0	-3	-3
620	2242	2	2	4	2	0	3	0	1	0	0	-3	-3
621	2322	2	3	2	2	0	3	1	0	0	0	-3	-3
622	2422	2	4	2	2	0	3	0	1	0	0	-3	-3
623	3222	3	2	2	2	0	3	1	0	0	0	-3	-3
624	4222	4	2	2	2	0	3	0	1	0	0	-3	-3
625	2222	2	2	2	2	0	4	0	0	0	0	-4	-4

Appendix E: SQL Code for Database Creation

SQL code appearing in this appendix may be used to create the MySQL database and tables described in the Methodology sections of this research. Note that dependencies exist and this will require that the tables and table data are created in an appropriate order. Tables may be populated with data at the time the tables are created if they are created in the order presented below.

Database Table: CCCScores

```
CREATE TABLE cccscores (
colocation_id SMALLINT UNSIGNED NOT NULL,
cccscore TINYINT NOT NULL,
PRIMARY KEY (colocation_id)
);
```

Database Table: STContexts

```
CREATE TABLE stcontexts (
stcyear DATE NOT NULL,
ctuid CHAR(10) NOT NULL,
colocation_id SMALLINT UNSIGNED NOT NULL,
PRIMARY KEY (stcyear, ctuid),
FOREIGN KEY(colocation_id) REFERENCES cccscores(colocation_id)
);
```

Database Table: MObjects

```
CREATE TABLE mobjects (
mobject_id MEDIUMINT UNSIGNED NOT NULL AUTO_INCREMENT,
mobject_sin INT UNSIGNED NOT NULL,
mobject_fname VARCHAR(30) NOT NULL,
mobject_lname VARCHAR(30) NOT NULL,
PRIMARY KEY (mobject_id)
);

ALTER TABLE mobjects ADD INDEX mobjectsin (mobject_sin),
ADD INDEX mobjectlname (mobject_lname);
```

Database Table: RELATIONSHIPS

```
CREATE TABLE relationships (
mobject_id MEDIUMINT UNSIGNED NOT NULL,
stcyear DATE NOT NULL,
ctuid VARCHAR(10) NOT NULL,
PRIMARY KEY (mobject_id, stcyear, ctuid),
FOREIGN KEY(mobject_id) REFERENCES mobjects(mobject_id),
FOREIGN KEY(stcyear, ctuid) REFERENCES stcontexts(stcyear, ctuid)
);
```

References

Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S. I., Jern, M., Kraak, M-J., Shumann, H., Tominski, C. (2010). Space, time and visual analytics. International Journal of Geographical Information Science, 24 (10), 1577-1600.

Andrienko, G., Andrienko, N., Heurich, M. (2011). An event-based conceptual model for context-aware movement analysis. International Journal of Geographical Information Science, 25 (9), 1347-1370.

Anselin, L. (2003). GeoDa 0.9 user's guide. Retrieved from https://geodacenter.asu.edu/software/documentation on December 6, 2014.

Beale, L., Abellan, J. J., Hodgson, S., Jarup, L. (2008). Methodological issues and approaches to spatial epidemiology. Environmental Health Perspectives, 116 (8), 1105-1110.

Chun, Y., Griffith, D. A. (2013). Spatial statistics & geostatistics. Los Angeles: SAGE Publications Inc.

City of Toronto. Toronto Public Health. (2008). The unequal city: income and health inequalities in Toronto, 2008. Toronto: Toronto Public Health.

City of Toronto. (2014). Release of the 2006 Census on language, immigration, citizenship, mobility/migration. Retrieved from http://www1.toronto.ca/city_of_toronto/social_development_finance__administration/files/pdf/2006_lang_imm_citizenship_mobility_backgrounder.pdf on November 29, 2014.

Cromley, E. K., McLafferty, S. L. (2012). GIS and public health. (2nd Edition). New York: The Guilford Press.

Dijkstra, A., Janssen, F., De Bakker, M., Bos, J., Lub, R., Van Wissen, L. J. G., Hak, E. (2013). Using spatial analysis to predict health care at the local level: a case study of type 2 diabetes medication use and its association with demographic change and socioeconomic status. PLoS ONE, 8 (8): e72730.

Entwisle, B. (2007). Putting people into place. Demography, 44 (4), 687-703.

Freisthler, B., Lery, B., Gruenewald, P. J., Chow, J. (2006). Methods and challenges of analyzing spatial data for social work… Social Work Research, 30 (4), 198-210.

Gatrell, A., Löytönen, M. (Eds.). (1998). GIS and health. Philadelphia: Taylor & Francis Inc.

Geist, C., McManus, P. (2008). Geographical mobility over the life course: motivations and implications. Population, Space and Place, 14, 283-303.

Getis, A., Aldstadt, J. (2004). Constructing the spatial weights matrix using a local statistic. Geographical Analysis, 36 (2), 90-104.

Hägerstrand, T. (1970). What about people in regional science? Papers of the Regional Science Association, 24, 7-21.

Heywood, I., Cornelius, S., Carver, S. (2006). An introduction to geographic al information systems. (3rd Edition). Edinburgh Gate: Pearson Education Limited.

Hornsby, K., Egenhofer, M. J. (2002). Modeling moving objects over multiple granularities. Annals of Mathematics and Artificial Intelligence, 36, 177-194.

Jacquez, G. M., Greiling, D. A., Kaufmann, A. M. (2005). Design and implementation of a space-time intelligence system for disease surveillance. Journal of Geographical Systems, 7, 7-23.

Jerrett, M., Gale, S., Kontgis, C. (2010). Spatial modeling in environmental and public health research. International Journal of Environmental Research and Public Health, 7, 1302-1329.

Joyce, K., Bambra, C. (2010). Health inequalities in developed nations. Social Alternatives, 29 (2), 21-27.

Kawachi, I., Subramanian, S. V., Almeida-Filho, N. (2002). A glossary for health inequalities. Journal of Epidemiology and Community Health, 56 (9), 647-652.

Kraak, M-J., Ormeling, F. (2003). Cartography: visualization of geospatial data. (2nd Edition). Edinburgh Gate: Pearson Education Limited.

Kwan, M-P. (2012a). The uncertain geographic context problem. Annals of the Association of American Geographers, 102 (5), 958-968.

Kwan, M-P. (2012b). How GIS can help address the uncertain geographic context problem in social science research. Annals of GIS, 18 (4), 245-255.

Lofters, A., Slater, M., Kirst, M., Shankardass, K., Quinonez, C. (2014). How do people attribute income-related inequalities in health? a cross-sectional study in Ontario, Canada. PLoS ONE, 9 (1): e85286.

Longley, P. A., Goodchild, M. F., Magcuire, D. J., Rhind, D. W. (2011). Geographic information systems & science. (3rd Edition). New Jersey: John Wiley & Sons, Inc.

Lundblad, P., Eurenius, O., Heldring, T. (2009). Interactive visualization of weather and ship data. 2009 International Conference Information Visualisation, 379-386.

Marmot, M. G. (2003). Understanding social inequalities in health. Perspectives in Biology and Medicine, 46 (3), S9-S23.

Masseria, C., Hernández-Quevedo, C., Allin, S. (2010). Health inequality: what does it mean and how can we measure it? Expert Reviews, 10 (2), 177-186.

Ng, Edward. (2011). The healthy immigrant effect and mortality rates. Statistics Canada Health Reports, 22 (4), 1-5.

Pickett, K., Wilkinson, R.G. (2008). People like us: ethnic group density effects on health. Ethnicity & Health, 13 (4), 321-334.

Public Health Agency of Canada. (2013). What makes Canadians healthy or unhealthy? Retrieved from http://www.phac-aspc.gc.ca/ph-sp/determinants/determinants-eng.php on December 1, 2014.

Rezaeian, M., Dunn, G., St. Leger, S., Appleby, L. (2004). The production and interpretation of disease maps: a methodological case-study. Soc Psychiatry Psychiatr Epidemiol, 39, 947-954.

Rogerson, P., Yamada, I. (2009). Statistical detection and surveillance of geographic clusters. Boca Raton: Taylor & Francis Group, LLC.

Samarasundera, E., Walsh, T., Cheng, T., Koenig, A., Jattansingh, K., Dawe, A., Soljak, M. (2012). Methods and tools for geographical mapping and analysis in primary health care. Primary Health Care Research and Development, 13, 10-21.

Semple, H. M., Cudnik, M. T., Sayre, M., Keseg, D., Warden, C. R., Sasson, C. (2013). Identification of high-risk communities for unattended out-of-hospital cardiac arrests using GIS. J Community Health, 38, 277-284.

Statistics Canada. (2015a). 2001 Census, profiles of Census tracts. Retrieved from http://dc1.chass.utoronto.ca.ezproxy.lib.ryerson.ca/census/2001/index.html on March 21, 2015.

Statistics Canada. (2015b). 2006 Census, profiles of Census tracts. Retrieved from http://dc1.chass.utoronto.ca.ezproxy.lib.ryerson.ca/census/2006/index.html on March 21, 2015.

Statistics Canada. (2015c). 2001 Census geography boundary files. Retrieved from http://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2001-eng.cfm on March 22, 2015.

Statistics Canada. (2015d). 2006 Census geography boundary files. Retrieved from http://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2006-eng.cfm on March 22, 2015.

Statistics Canada. (2015e). 2006 community profiles. Retrieved from http://www12.statcan.gc.ca/census-recensement/2006/dp-pd/prof/92-591/details/page.cfm?Lang=E&Geo1=CMA&Code1=535&Geo2=PR&Code2=35&Data=Count&SearchText=toronto&SearchType=Begins&SearchPR=01&B1=All&Custom= on March 23, 2015.

Tong, S. (2000). Migration bias in ecological studies. European Journal of Epidemiology, 16, 365-369.

Unwin, A., Unwin, D. (1998). Exploratory spatial data analysis with local statistics. Journal of the Royal Statistical Society. Series D (The Statistician), 47 (3), 415-421.

Wilkinson, R. G. (2006). The impact of inequality. Social Research, 73 (2), 711-733.

Yen, I. H., Kaplan, G. A. (1998). Poverty area residence and changes in physical activity level: Evidence from the Alameda County Study. American Journal of Public Health, 88 (11), 1709-1712.