

THE URBAN HEALTH INDEX

A HANDBOOK FOR ITS CALCULATION AND USE

**A VERSATILE APPROACH TO MEASURING
AND MAPPING HEALTH INEQUALITIES**

THE URBAN HEALTH INDEX

A HANDBOOK FOR ITS CALCULATION AND USE

WHO Library Cataloguing-in-Publication Data

The Urban Health Index: A Handbook for its Calculation and Use
I. World Health Organization.

ISBN 978 92 4 150780 6

Subject headings are available from WHO institutional repository

© World Health Organization 2014

All rights reserved. Publications of the World Health Organization are available on the WHO web site (www.who.int) or can be purchased from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 3264; fax: +41 22 791 4857; e-mail: bookorders@who.int).

Requests for permission to reproduce or translate WHO publications –whether for sale or for non-commercial distribution– should be addressed to WHO Press through the WHO website (http://www.who.int/about/licensing/copyright_form/en/index.html).

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use.

The named authors alone are responsible for the views expressed in this publication.

Prepared for:
World Health Organization Centre for Health Development
Kobe, Japan

Authors:
Georgia State University, Atlanta, United States of America
Scott R. Weaver, Ph.D.
Research Assistant Professor, Division of Epidemiology and Biostatistics, School of Public Health
Director, Data and Research Services, Center of Excellence on Health Disparities Research, School of Public Health
Dajun Dai, Ph.D.
Assistant Professor, Department of Geosciences
Christine Stauber, Ph.D.
Assistant Professor, Division of Environmental Health, School of Public Health
Ruiyan Luo, Ph.D.
Assistant Professor, Division of Epidemiology and Biostatistics, School of Public Health
Richard Rothenberg, M.D., M.P.H.
Regents' Professor, Division of Epidemiology and Biostatistics, School of Public Health
Director, Research Core, Center of Excellence on Health Disparities Research, School of Public Health

Design by Phoenix Aid Design, Denmark.
Printed in Japan.

Suggested citation:
The Urban Health Index: A handbook for its calculation and use. Kobe, Japan: World Health Organization; 2014.

TABLE OF CONTENTS

TABLE OF FIGURES	VI
ACKNOWLEDGEMENTS	VII
INTRODUCTION TO THE URBAN HEALTH INDEX AND THIS HANDBOOK	2
WHAT IS THE URBAN HEALTH INDEX?	2
WHO SHOULD READ THIS HANDBOOK?	3
HOW IS THIS HANDBOOK ORGANIZED?	3
OVERVIEW OF THE PROCESS FOR CONSTRUCTING THE URBAN HEALTH INDEX (UHI)	4
SELECTING INDICATORS FOR THE URBAN HEALTH INDEX	5
SELECTION OF INDICATORS: A FRAMEWORK	6
RATIONALE FOR INDICATOR SELECTION	6
EXAMPLES OF DATA SOURCES	9
SELECTING A GEOGRAPHIC UNIT OF ANALYSIS	10
AN ILLUSTRATION: SELECTING INDICATORS FOR AN APPLICATION OF THE URBAN HEALTH INDEX FOR SALVADOR, BAHIA, BRAZIL	11
STEPS FOR CONSTRUCTING THE URBAN HEALTH INDEX	13
1. PREPARING THE DATA	13
2. CALCULATING THE URBAN HEALTH INDEX	15
Standardization of the Indicators	15
Amalgamation of the Standardized Indicators	17
3. CALCULATING THE ERROR VARIANCE OF THE URBAN HEALTH INDEX	17
4. CALCULATING SUMMARY STATISTICS FOR THE URBAN HEALTH INDEX	18
5. CREATING AN INDEX PLOT OF THE URBAN HEALTH INDEX	19
6. QUANTIFYING GEOGRAPHIC DISPARITIES USING THE URBAN HEALTH INDEX	20
Assessing Disparities Among Areas in the Middle Range of the UHI: The Disparity Gradient	20
Assessing the Extent of Disparity between the Outer UHI Quantiles: The Disparity Difference and Ratio	21
MAPPING THE URBAN HEALTH INDEX	22
ANALYSIS AND REPORT WRITING	26
THE ONE PAGE SUMMARY	26
THE DETAILED REPORT	27
LIMITATIONS AND CONDITIONS FOR INTERPRETATION OF THE URBAN HEALTH INDEX	30
RELIABILITY AND VALIDITY OF INDICATORS	30
CONSISTENCY OF INDICATORS	32
UHI: RELATIVE VERSUS ABSOLUTE INTERPRETATIONS	33
AVOID EXTRAPOLATING BEYOND THE INDICATORS	33
INTERACTIONS WITH PLANNERS AND DECISION MAKERS—USING THE URBAN HEALTH INDEX TO SET PRIORITIES AND GOALS FOR POLICY CHANGE	34
APPENDIX	36
APPENDIX A: GUIDE TO AN EXCEL-BASED TOOL (V. 1) FOR CALCULATION AND VISUALIZATION OF THE URBAN HEALTH INDEX	36
Tools and Capabilities	36
Instructions for Using the Tool	37
APPENDIX B: GUIDANCE FOR A WEIGHTED URBAN HEALTH INDEX	39
APPENDIX C: DERIVATION OF THE VARIANCE FOR THE URBAN HEALTH INDEX	41
Derivation of the Variance for UHI: Without Assuming Normality	41
Derivation of the Variance for UHI: With Assumed Normality	42
APPENDIX D: GUIDANCE AND ILLUSTRATION OF MAPPING THE URBAN HEALTH INDEX WITH QGIS AND ARCGIS	45
Using ArcGIS for Mapping the UHI	45
Using QGIS for Mapping the Urban Health Index	50
APPENDIX E: A VISUAL SUMMARY FOR COMMUNICATING THE URBAN HEALTH INDEX	60
APPENDIX F: A POWERPOINT TEMPLATE FOR COMMUNICATING THE URBAN HEALTH INDEX RESULTS	62

TABLE OF FIGURES

Figure 1: Flow Diagram for Construction of an Urban Health Index	3
Figure 2: Framework for Classifying Measures of Urban Health.	6
Figure 3: Health Indicator Data for Tokyo Prefecture, Japan (2003-2007) in Spreadsheet Format	12
Figure 4: Illustrating the Calculation of Using the Sample Minimum as the Lower Goalpost.	15
Figure 5: Illustrating the Calculation of Using Zero as the Lower Goalpost.	15
Figure 6: Illustrating the Calculation of the UHI	16
Figure 7: Example UHI Index Plot for Tokyo Prefecture, Japan	18
Figure 8: Flowchart of Steps for Mapping the UHI.	22
Figure 9: An Example: Mapping the UHI in Atlanta, Georgia.. . . .	23
Figure 10: The UHI Representation Using Point (left) and Surface (right) After Smoothing	24
Figure 11: Results of the Weighting Simulations	39
Figure 12: Illustrated Calculation of the Weighted Geometric Mean.	40
Figure 13: The User Interface of ArcMap to add a Shapefile (left) and the Resulting Screen After the Shapefile is Added (right).	44
Figure 14: The User Interface to Open an Attribute Table (left) and the Resulting Attribute Table (right)	45
Figure 15: The UHI Excel file Imported into ArcMap (left) and the ArcMap Table of Contents (right).. . . .	46
Figure 16: Concept Model of the Join Function in ArcMap	46
Figure 17: Process to Select the Join Function (left), and the Join Data window (right)	47
Figure 18: Symbology Setting (left) and the Choropleth Map (right)	47
Figure 19: Flip Option in Layer Properties Window (left) and the Map after Flipping Map Colors (right).	48
Figure 20: Insert Menu to Add Map Elements (left) and the Final Map (right)..	48
Figure 21: QGIS Interface of Adding Data (left) and the Browser Tree (right)	49
Figure 22: Plugins Interface.	50
Figure 23: Interface for the Layers Window.	50
Figure 24: Checking the Data Type for UHI Values When Joining UHI Data to the Shapefile	51
Figure 25: Joining the UHI File in CSV Format to a Boundary Shapefile.	51
Figure 26: Viewing Fields in the Layer Properties Window.	52
Figure 27: Changing Layer Properties.	52
Figure 28: Add Column Window.	53
Figure 29: Layer Properties Window after the New Column is Added	53
Figure 30: Field Calculator Window.	54
Figure 31: Attribute Table Window UHI Data.	55
Figure 32: Creating a Choropleth Map in the Layer Properties Window.	55
Figure 33: UHI Results in a Choropleth Map	56
Figure 34: Paper and Quality Set-Up in Print Composer Window	57
Figure 35: Add, Select, and Move a Map in the Print Composer Window	57
Figure 36: Item Properties in the Print Composer Window	58
Figure 37: Selecting Map Elements to Print in the Composer Window	58

ACKNOWLEDGEMENTS

This handbook was written and developed by Scott R. Weaver, Dajun Dai, Christine Stauber, Ruiyan Luo, and Richard Rothenberg of Georgia State University (GSU) in Atlanta, Georgia, United States of America. It is a product of collaboration since 2010 between the GSU research team based in the School of Public Health, and the World Health Organization Centre for Health Development, or WHO Kobe Centre (WKC), to develop the Urban Health Index. It also reflects inputs from a broad group of experts over the course of this collaboration from both within the two collaborating institutions and beyond.

Funding for the development of this handbook was provided by the WKC. Megumi Kano (WKC) commissioned the work and provided overall technical guidance. Amit Prasad (WKC) provided technical inputs and contributed to earlier drafts of this handbook. Hui Eva Liang (GSU) provided technical assistance. Martin Bortz (University of Heidelberg) applied the Urban Health Index in Rio de Janeiro, Brazil, and produced the results and figures presented in Appendix E. Members of the GHIFT collaborative contributed technical inputs. Michael Eriksen, Founding Dean of the Georgia State University School of Public Health, provided technical and moral support. Kathleen Sartori (GSU) performed copy editing. Matthew Simmonds (GSU) developed the graphic design. Sophia Banton, Byungwoo Cho and Kumiko Miyake (GSU) assisted in the translation, data management, and analysis.

INTRODUCTION TO THE URBAN HEALTH INDEX AND THIS HANDBOOK

The Urban Health Index (UHI) is a single metric that can be used to measure and map the disparities in health determinants and outcomes in urban areas. This Handbook is primarily intended for those who want to calculate the UHI for a particular geographic area of interest.

What is the Urban Health Index?

The UHI provides a flexible approach to selection, amalgamation, and presentation of health data. Its purpose is to furnish visual, graphical, and statistical insight into various health indicators and health determinants within particular geographic boundaries and health disparities with a focus on capturing intra-urban health disparities. The UHI may be used by public health workers, evaluators, statisticians, program managers, academic researchers, and decision makers to examine the current status of urban areas, to assess change and the effect of program interventions, and to plan for urban improvements.

The UHI was developed through a set of papers and consultations commissioned by the World Health Organization Centre for Health Development (WHO Kobe Centre). The original motivating question was whether or not a single urban metric was feasible and could capture critical information about an urban area. The decision from these discussions was that a single metric—by default a composite statistic—would not serve the many purposes demanded of it, and a multipurpose tool was likely to be more advantageous. The WHO Kobe Centre and its consultants envision a method that could be applied to health indicators and health determinants, would have built-in measures for disparities, and would lend itself to geographic visualization.

The tool that emerged was not predicated on new methods, but rather built on a methodology that has been under development for many years. A review of extant indicator databases and current approaches to the formation of indices revealed diversity in terminology but considerable concordance in the types of indicators used.¹ Numerous indices have been proposed, most predicated on the inclusion of specific indicators, and many employing arbitrary weighting schemes. The UHI described here builds on the considerable correlation among indicators of the same type (for example, total mortality and its subsets), and eschews weighting in favor of different indicator combinations.² The method for the UHI construction has drawn on the approach used by the Human Development Index (HDI)³ that standardizes indicators by converting them to a proportion of their range, and combines them using the geometric mean. The UHIs for contiguous areas are rank ordered. A *disparity ratio* is calculated from the extremes of the distribution, and a *disparity slope* is calculated by the angle of increase. UHIs for contiguous areas are mapped to provide an immediate visual grasp of the extent and distribution of disparities.

-
- 1 Rothenberg R, Stauber C, Weaver SR, Dai D, Prasad A, Kano M. Review and commentary: urban health indicators and indices – current status. Unpublished work.
 - 2 Rothenberg R, Weaver SR, Dai D, Stauber C, Prasad A, Kano M. A flexible Urban Health Index of small area disparities. J Urban Health. 2014. doi:10.1007/s11524-014-9867-6.
 - 3 United Nations Development Programme (UNDP). Human Development Report 2011. Sustainability and equity: A better future for all. New York: Palgrave Macmillan; 2011 (http://hdr.undp.org/sites/default/files/reports/271/hdr_2011_en_complete.pdf, accessed 11 September 2014).

This approach permits freedom to choose the scale (from small area estimates to national comparisons), the indicators (largely dependent on data availability), and the mode of presentation. The following chapters outline the method in detail, provide examples of prior and potential use, and furnish resources for further development and applications. Building on existing work provides connections to the considerable ongoing efforts in this field, and offers pathways to evaluation and innovation in a specific geographic locale.

Who should read this handbook?

This handbook is written for a diverse audience. Its primary target audiences are the public health workers and policy analysts who require a method for quantifying urban health disparities and then presenting such information to policy planners and decision makers. However, research and academic communities may also have an interest. In its Appendices, the handbook offers a rigorous approach to the statistical aspects and visualization of the UHI. Finally, the UHI may prove to be a cogent tool for the media to reach a lay audience concerned with health inequalities.

How is this handbook organized?

The sections of this handbook are organized to guide the reader through the step-by-step calculation of the UHI, from choosing health indicators and determinants to communicating with decision makers and others about the index. Throughout the handbook, actual applications of the UHI in Atlanta (USA), Tokyo (Japan), Shanghai (China), and Rio de Janeiro (Brazil) are provided as illustrative examples.

We begin with a brief Overview of the process for constructing a UHI. Initial considerations involved in this construction are described in the next section, Selecting Indicators for the Urban Health Index. This section also provides guidance for selecting and gathering data for valid and reliable measures of selected indicators and determinants at the desired level of measurement or unit of analysis. Next, the primary steps of calculating the UHI values for an urban area of interest are described and illustrated. This section includes: (1) the preparation and examination of the selected indicator and determinant variables; (2) the use of statistical or spreadsheet software programs to standardize, then aggregate the indicators into the UHI; and (3) the plotting and statistical summation of UHI values for capturing the extent of health disparities within an urban area or among urban areas.

Next, we describe the geographic information systems (GIS) tools for mapping and analyzing the UHI. These GIS tools complement and greatly expand upon the more simplified tools described in the previous section. The UHI and the graphs and statistics computed from it can be used by policy makers and others to gauge the level of health disparities of an urban area and how to track these elements over time.

With the goal of using the UHI as a tool to affect public policy, the following two sections provide guidance on incorporating the UHI findings into a report for dissemination and information on the factors that may limit or qualify interpretation of the UHI. The final section provides general guidance and suggestions for consideration when preparing to communicate the UHI results with decision makers.

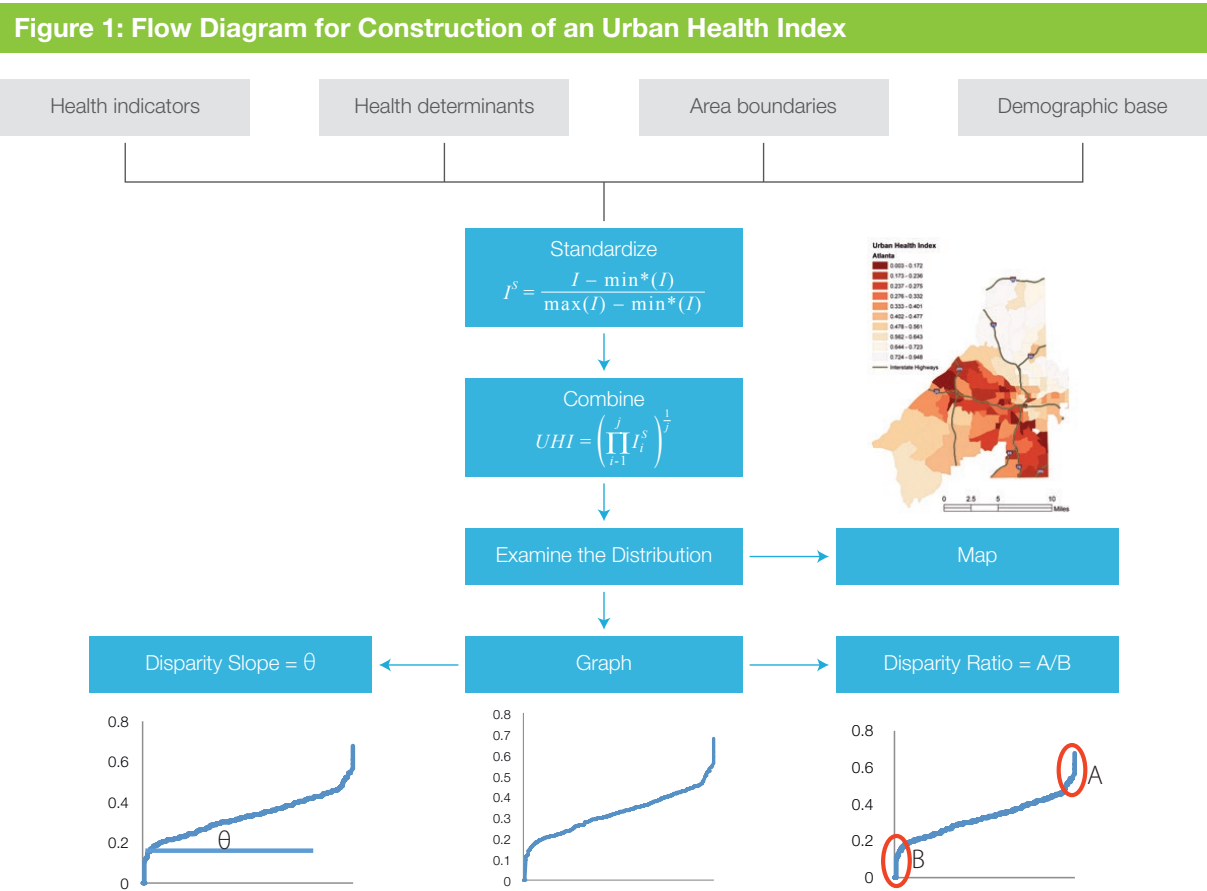
Finally, the technical appendices describe an accompanying Microsoft Excel-based tool⁴ for calculating the UHI and delve into more technical topics, including the calculation of a weighted UHI, the calculation of the variance for the UHI, and more powerful GIS tools for mapping the UHI.

4 Both this handbook and the Excel-based calculation tool are available for download from the WHO Centre for Health Development's website at http://www.who.int/kobe_centre/publications/urban_health_index_toolkit/en/ (accessed 16 October 2014).

OVERVIEW OF THE PROCESS FOR CONSTRUCTING THE URBAN HEALTH INDEX (UHI)

The construction of the UHI involves standardizing and combining indicators of health determinants or outcomes by geographic area using simple formulae, and then characterizing their statistical and geographic distribution.

Figure 1 provides an overview of the required elements and their interrelationship for the construction of the UHI. Each step in the flow diagram will be discussed in detail in later sections. As shown in the figure, an unspecified set of health indicators or determinants can be standardized and combined. This aggregate is the UHI value for a given geographic area. When ranked and graphed, this set of UHI values often appear as a rising line with distorted tails (the low end drops off precipitously; the high end rises steeply). The slope of the line and the ratio of the tails constitute measures of health disparity or inequality. If the areas are contiguous, geovisualization is accomplished by dividing the distribution into deciles and coloring them by graded hues, with darker hues denoting increasing disadvantages. The map provides a direct visual comprehension of the distribution of disparities.



SELECTING INDICATORS FOR THE URBAN HEALTH INDEX

The UHI is a single metric that can be used to measure and map the disparities in health determinants and outcomes in urban areas. This Handbook is primarily intended for those who want to calculate the Urban Health Index for a particular geographic area of interest based on locally-selected indicators.

In its simplest form, an index is constructed from a set of variables for indicators that have been transformed (standardized, normalized, and scaled) so that they are directly comparable, and then combined together. The UHI is more aptly defined as a framework and methodology for an index of the health in urban areas. A particular index developed using the UHI framework will derive its meaning and usefulness from the pre-selected indicators that comprise it. As such, the most critical task in constructing the UHI is selecting the best indicators for the intended purpose. This task will necessarily be a local decision that will vary over time, across different circumstances, and among stakeholders. In fact, the choices for the indicators used to develop the UHI will often be constrained by the availability or, rather, lack of suitable data on the desired indicators for the geopolitical level of interest. However, caution should be taken when relying on this available data-driven approach, as the data available should still be selected through some type of conceptual framework. For this reason, the application of the UHI is not prescriptive in terms of the specific indicators that must or should be selected. Rather, the application is flexible and open to modification.

In fact, there is nothing intrinsic in the UHI linked to “urban”. The intention behind the development of the UHI and its application was focused on urban areas, but the approach and methodology of the index could be applied to rural areas, too. However, while the application of the UHI is not limited to urban areas, its focus on health and health disparities at small geographic area level makes it more applicable to population dense areas of urban places than the sparsely populated rural areas. Application to rural areas often requires a larger geographic unit of analysis.

The guidance provided in this section will highlight the need to consider a framework for indicator selection that is focused on the intended purpose and use of the index. In providing a framework and suggestions for the process of selecting indicators, it will reference the works of others the reader may want to consult. This section provides a practical guide for locating and using data given the chosen or desired indicators. The Handbook discusses alternate reference sources, such as those on environmental measures, population distribution, indicators of health disparities, and available services. Links will be made to existing, relevant frameworks, such as the WHO's *Urban HEART*.⁵ This section will also offer some guidance on selecting the geopolitical or geo-statistical units, with consideration of the availability and reliability of data, number of units, homogeneity of populations within units, meaningfulness of units for enacting policy, and temporal stability of boundaries, and number of units.

5 Urban HEART: Urban Health Equity Assessment and Response Tool. Kobe: WHO; 2010 (http://www.who.int/kobe_centre/publications/urban_heart.pdf?ua=1, accessed 11 September 2014).

Selection of Indicators: A Framework

To determine which indicators would be best for the application of an UHI, it is important to discuss the various types of indicators available. A review of urban health indicators and indices found great similarity in the areas from which these major indices were developed.⁶ To address these similarities, we propose the following framework for classifying the available measures. This framework contains three levels of measurement: rubrics, domains, and indicators (see Figure 2). *Rubrics* represent societal factors that affect health, either directly or as determinants. *Domains* are specific factors within a rubric for which measurements are available (for example: Sociodemographics [Rubric] → Infant mortality [Domain]). Indicators, and the disparities that are derived from them, are the specific measures actually used (for example: Environment [Rubric] → Air quality [Domain] → Proportion of households living within 300 m of major industrial stationary sources of air pollution [Indicator]). The framework provided in Figure 2 highlights the ability to select indicators that exist within various domains or larger groupings without being restricted to specific measures that may not be available for different applications of the UHI. The Handbook will provide examples of some different approaches for selecting indicators for the UHI.

Many of the currently developed indices are drawn from similar sets of indicators. Nine domains (see Figure 2 for examples) appear in many of the aggregations. Three of them (health care, infant mortality, and education) appear in more than two-thirds of indices reviewed. Thus, despite disagreement about the particulars, there is some evidence of agreement about basic content. This observation suggests that even local-level data can be available for creating the UHI.

Rationale for Indicator Selection

The choice of indicators used to develop the UHI is intended to be flexible and will vary with each application and purpose. When choosing a list of indicators, it is important to consider a framework and set of criteria for selection. Flowers and colleagues⁷ provide a checklist of 20 aspects to consider, some of which are title, origin, rationale, routine or special collection, and frequency. They proposed that other characteristics of indicators should also be considered, including the strengths, weaknesses, and ability to influence practice or behavior. A simple, and perhaps more forceful summary of ideal characteristics of indicators is provided by Etches et al.,⁸ whose keywords bear repeating: consensual, conceptual, valid, sensitive, specific, feasible, reliable, sustainable, understandable, timely, comparable, and flexible.

When considering indicators, it is important to determine the purpose of the index. In the application of the UHI to determinants of health, the purpose may be to examine a broad range of indicators from a wide range of social, geographical, economic and environmental determinants of health. This selection of indicators will provide a composite index that is not heavily influenced by one specific domain. The WHO's *Urban HEART*, for example, presents an indicator framework of domains including health outcomes, physical environment and infrastructure, social and human development, governance and economy. One challenge with this approach is that when the selection of indicators are drawn from many different fields, when combined, these indicators may obscure differences as a result of the interaction of the various indicators.

6 Rothenberg R, Stauber C, Weaver SR, Dai D, Prasad A, Kano M. Review and commentary: urban health indicators and indices – current status. Unpublished work.

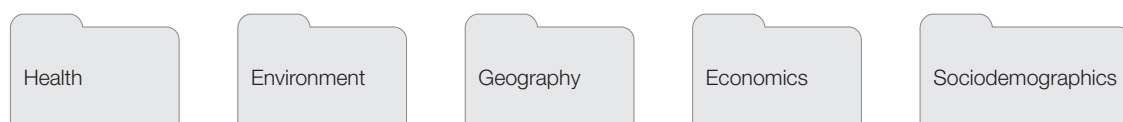
7 Flowers J, Hall P, Pencheon D. Public health indicators. *Public Health*. 2005;119:239–245.

8 Etches V, Frank J, DiRuggiero E, Manuel D. Measuring population health: a review of indicators. *Annu Rev Public Health*. 2006;27:29–55.

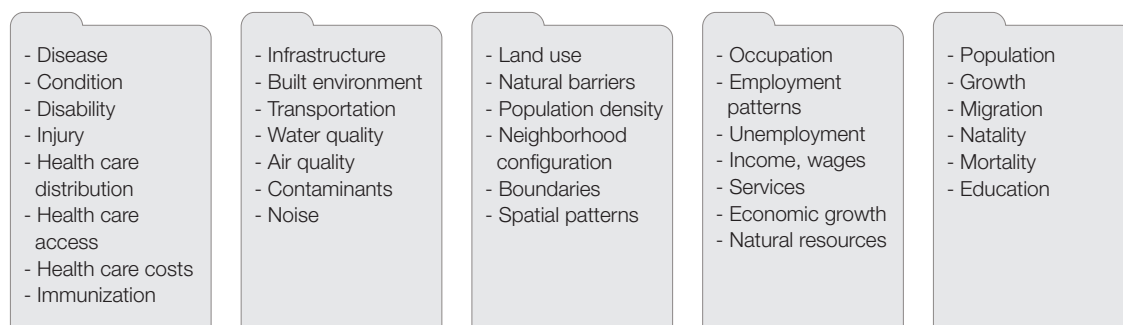
In application, many indices choose to draw from a wide variety of domains. For example, Stephens et al.,⁹ created an Index of Deprivation comparing Accra, Ghana with Sao Paulo, Brazil. Interestingly, groups working on the two areas devolved on the same five domains: income, education of head of household, number of persons per room, sanitation, and safe water access, but had to use different indicators within those domains. The indicators selected were socio-demographic, economic and environmental conditions for each location. The collected data produced an overall picture that concealed substantial differences between the two areas. Those differences were revealed, however, by a simple choropleth map comparing the two cities by using four levels of socio-environmental conditions. The index generated was a composite of multiple determinants that can impact health. This may be a more balanced approach to examining the various urban health facets without over-emphasizing one component, such as education or economic conditions, over another, such as environmental conditions.

Figure 2: Framework for Classifying Measures of Urban Health

Rubrics



Domains



Indicators



The selection of exact indicators forming the UHI will likely vary with each location (and availability) and purpose. As another example, in an examination of sociodemographic inequalities and neonatal conditions in the city of Salvador, Brazil, Gonçalves and colleagues¹⁰ selected five variables in five different domains from the Brazilian Census data to construct a Living Conditions Index. The variables they selected were from the following domains: income, education, access to water, living in slum conditions

9 Stephens C, Akerman M, Avle S, Maia PB, Companario P, Doe B, Tetteh D. Urban equity and urban health: using existing data to understand inequalities in health and environment in Accra, Ghana and Sao Paulo, Brazil. *Environ and Urban*. 1997;9:181–202.

10 Gonçalves AC, Costa MCN, Paim JS, Silva LMV, Braga JU, Barreto M. Social inequalities in neonatal mortality and living condition. *Revista Brasileira de Epidemiologia*. 2013;16(3):682–691.

and crowding. For the purposes of their application, they examined how lower and higher living conditions associated with neonatal mortality. The benefit of selecting indicators from various domains enabled the authors to construct a composite view of living conditions in the city of Salvador.

In construction of the UHI for Atlanta, the indicators were selected from three domains: demographic, economic and educational.¹¹ An environmental indicator, such as those mentioned in the composite indices above, would not be a valid indicator for the city of Atlanta, as access to safe drinking water does not vary within the city. Indicators that vary across the geographic unit of interest may be more useful in demonstrating disparities within a city. Therefore, when selecting the most appropriate indicators, the variability of a given indicator within the area of interest should also be considered. As shown in Table 1, there are a number of criteria to consider when determining which indicators to select for application in the UHI. These criteria serve to prompt thoughtful selection of indicators and consideration of characteristics for the data ultimately selected for UHI construction. Each criterion is discussed and an example is provided for illustration.

Table 1: Criteria and Examples of Indicator Selection

Criteria to Consider for Selection of Indicator(s)¹²	Discussion	Example of Indicators Demonstrating these Properties (or lack thereof)
Built on consensus	There is general agreement (at the local level or in the relevant literature) that the use of the indicator is justified.	Literacy, as a measure of access to education, has been used as an indicator in the Human Development Index. ¹³ It has also been adopted in a number of other indices such as the Living Conditions Index. ¹⁴
Relevant	There exists a clear rationale (such as a link to policy) for selecting the indicator and its relevance to the goals and objectives of calculating and using the UHI.	The measure of access to improved water is directly linked to the Millennium Development Goal of reducing the proportion of the population without access by half (by 2015), which may be more relevant in some contexts than in others. ¹⁵
Valid and Reliable	The indicator should measure what it purports to measure. The way in which the data are collected should be considered trustworthy and unflinching.	Special attention should be paid to the way the data are collected. If biases are present, they should be acknowledged, and their data limitations should be addressed. Is there an incentive to underreport or underestimate the indicator, such as cure rate for tuberculosis?
Sensitive to differences	The measure should vary across the units of interest to enable detection of differences.	According to administrative data, access to improved water would be 100% across the census tracts for the city of Atlanta and would not help to discern differences (i.e., low sensitivity to variance).
Clear specification	Clear details should be provided about the numerator and denominator used to construct the indicator.	If multiple variables are available that provide information about income, the choice and selection of the most appropriate measure (and its transformation) should be described in detail. If median and mean income variables are available and only one was selected, a rationale should be provided to explain why.
Repeatable	To enable comparison over different time periods, measures should be selected that can be tracked over time.	Consideration should be given if the method of measurement changes over time or place. This is especially important if comparing monetary values over time. For example, the Human Development Index uses purchasing power parity to enable comparisons of monetary values over time.

11 Rothenberg R, Weaver SR, Dai D, Stauber C, Prasad A, Kano M. A flexible Urban Health Index of small area disparities. *J Urban Health*. 2014. doi:10.1007/s11524-014-9867-6.

12 Flowers J, Hall P, Pencheon D. Public health indicators. *Public Health*. 2005;119:239–245.

13 United Nations Development Programme (UNDP). Human Development Report 2013. The rise of the South: human progress in a diverse world. Technical notes. New York: UNDP;2013 (http://hdr.undp.org/sites/default/files/hdr_2013_en_technotes.pdf, accessed 11 September 2014).

14 Goncalves AC, Costa MCN, Paim JS, Silva LMV, Braga JU, Barreto M. Social inequalities in neonatal mortality and living condition. *Revista Brasileira de Epidemiologia*. 2013;16(3):682–691.

15 United Nations. Millennium Development Goals [website]. (<http://www.un.org/millenniumgoals/>, accessed 11 September 2014).

Examples of Data Sources

A national or regional census is an important source to consider in application of the UHI, particularly because many nations already conduct one. For many countries, a national census will often provide quality data on a variety of measures that can be used in constructing an UHI of health determinants, such as income, education, housing, race or ethnicity, and more. A census aims to count everyone in the population, and thus provides some of the most accurate data available for the small geographic areas.

Governments will usually place access restrictions on individual-level census data to protect the privacy of individuals, but if access can be obtained it could provide the greatest flexibility to aggregate data to the geographic level and boundaries of choice. Even with restricted access, small-area aggregate data are often available for the development of the UHI.

Sometimes, in tandem with a national census, a long-form statistical survey will be administered to a sampled subpopulation, as was done in the American Community Survey.¹⁶ The indicators covered in long-form statistical surveys will be broader, and the analyst will need to contend with the impact of sampling error and less data availability in a smaller geographical area. For example, 1-year estimates from the American Community Survey are only available for areas with populations greater than 65,000. Three-year estimates are available for areas with population greater than 20,000, and 5-year estimates are available for census block groups.

Important considerations when selecting national census data include: the population enumeration methodology, response rates, coverage of population and possible biases, such as undercoverage of underrepresented populations and institutionalized populations, and how recent data were collected. Knowledge of the methodology can assist in selection of the most appropriate geographic level of analysis and temporal period. Key considerations for national census and statistical survey data include:

- Is data collected from the entire population or a sample?
- If a sample is used, what geographic unit is the data intended to represent?
- What time period is it intended to represent?
- What are the limitations with comparing across time?

Generally, census data are intended to be compared and therefore are collected in a format that lends itself to comparison over time; however, comparisons may not be advised between overlapping periods or periods of different lengths (i.e., 1-year vs. 3-years).

National vital statistics systems are another source of data that should be considered for developing the UHI for health outcomes. In the United States, the National Vital Statistics System is one of the oldest surveillance systems in the country.¹⁷ Most nations have a vital statistics system that tracks births and mortality on a continuous basis. Similar among many vital statistics systems is the use of the International Classification of Diseases (ICD) coding system. This classification system facilitates the international comparability of mortality statistics.¹⁸ In the initial development of the UHI, data from the National Japanese Health Statistics Database were used to construct the Index using mortality

16 US Census Bureau. American Community Survey [website]. (http://www.census.gov/acs/www/about_the_survey/american_community_survey/, accessed 11 September 2014).

17 Centers for Disease Control and Prevention. National Vital Statistics System [website]. (http://www.cdc.gov/nchs/nvss/about_nvss.htm, accessed 11 September 2014).

18 Centers for Disease Control and Prevention. Classification of Diseases, Functioning, and Disability [website]. (<http://www.cdc.gov/nchs/icd/icd10.htm>, accessed 11 September 2014).

data for the entire country.¹⁹ If using vital statistics, consider the level at which data can be accessed and if it is sufficiently reliable. For example, in the U.S., micro-level data are publicly available only at the county level. Sub-county level data can sometimes be obtained if a contact can be made within a local- or state-level agency. Because mortality *counts* vary by population size, they are not well-suited for the UHI; rather, mortality *rates* or standardized mortality *ratios* will be more appropriate. Consider whether the rates should be standardized for age or other demographic variables and the pros and cons of different approaches for standardizing mortality rates.

The Demographic and Health Surveys (DHS) are another possible data source to consider for the UHI.²⁰ The DHS are a nationally representative set of surveys conducted in 85 countries since 1984.²¹ They are publicly available for download and have a number of key indicators, as well as some (mostly self-reported) health outcomes. Important features of the DHS are the comparability of surveys over time and in different locations. Over the years, they have increasingly incorporated biomarkers and other measures of health, and in many instances, have georeferenced data available for analysis and that could be used for generating city- and subcity- level measures for large urban areas. As with other data sources, key considerations include the sampling frame and survey methodology. DHS provides thorough details on the data in manuals that should be consulted when considering application for the UHI.

Other sources of data include, but are not limited to, surveys from national agencies such as Ministries of Education, Health, Labor, Statistics, and municipal surveillance data collected by local governments. For example, the city of Atlanta in the U.S. provides links to crime data with street address and latitude and longitude coordinates.²² *Urban HEART*, a tool developed by the WHO for local municipalities to identify and reduce health inequities, highlights partnerships that can support access to a variety of data sets that could be employed in the calculation of the UHI.²³

Selecting a Geographic Unit of Analysis

In addition to the data sources and indicators selected, another important consideration is the selected geographic unit of analysis for UHI construction. Given the UHI's primary focus on describing intra-urban disparities, it is recommended to identify the smallest and most relevant geographic unit of analysis for the purpose of the particular application of the UHI. As an example, whereas data in the United States' American Community Survey may be available at the census block group (5-year estimates), it might make more sense to identify a geographic unit that is relevant to local policy, such as a neighborhood. Neighborhoods, while not always governmentally administrative units within municipal areas, may be relevant because they have a social and historical context that is meaningful to the populations within the larger area under study. They may also be more homogenous given forces that operate on how people select into neighborhoods, which may accentuate differences or inequalities between neighborhoods within an urban area.

Another consideration is whether the data sources for the selected indicators are available for the same geographic level. The limited availability of data on many indicators at the smallest area level will usually be a limiting factor and often forces the analyst to choose a geographic level of analysis larger

19 Rothenberg R, Weaver SR, Dai D, Stauber C, Prasad A, Kano M. A flexible Urban Health Index of small area disparities. *J Urban Health*. 2014. doi:10.1007/s11524-014-9867-6.

20 ICF International. The Demographic and Health Surveys (DHS) Program [website]. (<http://www.dhsprogram.com/>, accessed 11 September 2014).

21 Corsi DJ, Neuman M, Finlay JE, Subramanian SV. Demographic and health surveys: A profile. *Int J Epidemiol*. 2012;41:1602–1613.

22 City of Atlanta/Atlanta Police Department. Find my zone [website]. (<http://www.atlantapd.org/findmyzone.aspx>, accessed 11 September 2014).

23 Urban HEART: Urban Health Equity Assessment and Response Tool. Kobe: WHO; 2010 (http://www.who.int/kobe_centre/publications/urban_heart.pdf?ua=1, accessed 11 September 2014).

than would otherwise be desired. It may be necessary to settle for a larger aggregate or to employ GIS techniques to render area data concordant among different geographies. This is particularly true when geographic boundaries change over time, such as when municipalities are merged or divided or population shifts lead to changes in statistical area boundaries. Combining data sets from different sources may present challenges when selecting a more focused but more relevant unit of analysis. For example, Vital Statistics data sets, such as those available from the National Vital Statistics System,²⁴ are often only available at aggregated units that do not allow for small-area evaluation. As will become clearer in later sections of this Handbook, certain quantities derived from the UHI, such as the health disparities ratio, will tend to be less stable when fewer areas are evaluated. Although it is difficult to define a minimum, it is often preferable to have more than 50 areas. Although still informative when the number of areas is less than this 50, some visualizations and derived quantities (discussed later) may be less interpretable. Finally, one should consider that the patterns emerging from the application of the UHI to one geographic level or set of geographic boundaries could differ considerably from another geographic level or set of geographic boundaries – for example, a zip code or set of zip codes versus a census tract. When feasible, calculate the UHI for multiple geographic levels of an urban area of interest and interpret similarities and differences in the patterns, taking into consideration the processes that form the boundaries and the residency selection forces.²⁵

An Illustration: Selecting Indicators for an Application of the Urban Health Index for Salvador, Bahia, Brazil

Indicator selection for the UHI for Salvador was performed with an emphasis on criteria for indicator selections found in Table 1. Initial indicator selection was informed by examining literature that had applied similar approaches to census data from the city of Salvador. Based on this review, four indicators (of the five used in previously published research) were selected: percent receiving > 2X the minimum monthly wage, percent of literate 10 to 14 year-old children, percent of population with access to piped water, and the number of people per household as a proxy measure for population density. Three additional indicators were also selected: percent of literate adults aged 15 years and older, monthly income per capita, and percent of population with access to piped sewerage. These last three were added based on the following criteria: consensus, validity and reliability, repeatability, and sensitivity to differences. See Table 2 for a discussion of the selection criteria for these indicators.

Table 2: Rationale for Indicator Selection for Salvador, Bahia, Brazil

Indicator	Criterion	Discussion about selection
% of those 15 years and older and literate	Consensus, repeatable, valid and reliable	Unlike the indicator applied by others in the city of Salvador (i.e., % of 10 to 14 year- old and literate), this measure is used in the HDI. There is a general consensus that this is an important indicator of the development of a nation. Furthermore, it might be available for comparison in other data sets internationally, and the data have been collected and compared over time for Brazil.
Income per capita	Valid and reliable	Income per capita can be modified to enable comparisons unlike other measures, such as minimum wage, as minimum wage varies per country and over time. This measure can also allow calculation of purchasing power parity and enable comparison over time.
% with access to piped sewerage	Consensus, sensitive to differences	Globally, there is much evidence of the tremendous health gains associated with excreta removal via piped sewerage. In addition, a much smaller proportion of the world's population has achieved this metric, and the data are likely to demonstrate more differences within urbanized areas (as compared to other indicators, such as access to piped water).

24 Centers for Disease Control and Prevention. National Vital Statistics System [website]. (http://www.cdc.gov/nchs/nvss/about_nvss.htm, accessed 11 September 2014).

25 Here, selection forces are those factors that influence where people choose or are constrained to live. For example, a high quality school will often lead to people, especially families with school-aged children, to live within its catchment if they can afford to do so; whereas higher rental or home ownership costs may restrict others from living within that area.

At the time of writing this Handbook, there have been several applications of the UHI to various cities spanning the globe. These applications are highlighted in Table 3 and visualized in illustrations throughout this handbook.

Table 3: Recent Applications of the Urban Health Index

Location	Data source and Year(s)	Geographic unit of analysis	Domains Included	Indicators Included
Atlanta, Georgia, U.S.	American Community Survey (2005-2009)	Census tract	Employment, Income, Education, Demographic	% employment, % above poverty level, % with high school degree or higher, % with bachelor's degree or higher, household mean income, household median income, % of households not headed by a single female with children under 18 years old.
Tokyo, Japan	Ministry of Health, Labour and Welfare of Japan, National Statistics Center. 2003-2007 Vital statistics by local public health center and municipality, 2009 update.	Smallest level municipality (non-designated cities, villages, towns, city wards of designated cities, and special wards of Tokyo)	Health outcomes (Mortality/Diseases)	Sex-specific standardized mortality ratios (SMRs) for all-cause mortality and cause-specific mortality for cerebral vascular disease, heart disease, malignant neoplasm, and pneumonia.
Rio de Janeiro, Brazil	National Health Information System (2000-2012)	Ward, Administrative District	Health outcomes (Mortality/Diseases)	Mortality Rates for diabetes, ischemic heart disease, breast and cervical cancer, HIV, tuberculosis, infant mortality, traffic accidents and homicides.
Salvador, Brazil	2010 Brazil Census	Census tract	Water Quality, Sanitation, Education, Income, Demographic	% of household with access to a piped water supply, % of households with access to sewerage, % of 10-14 literate, % of those >14 literate, income per capita, % of population earning 2X or more of monthly minimum wage, people per household.
Shanghai, China	China Data Center, Shanghai 2000 Population Census Data Assembly	Town	Education, Employment patterns, Occupation, Sanitation, Infrastructure, Water Quality	% non-agricultural population, % employed, % professional and managerial jobs, % with middle-school degree or higher, % with high school degree or higher, % with junior-college degree or higher, % of household cooking using gas and electricity, % households with tap water, % households with own bath heating facilities, and % households with own lavatory.

STEPS FOR CONSTRUCTING THE URBAN HEALTH INDEX

Constructing the UHI involves collecting and preparing the data; standardizing indicators and calculating the index values and summary statistics; plotting the index for visual assessment of its distribution; and finally, quantifying the disparities in terms of the gap between extremes and the gradient of disparity.

Once indicators have been chosen and data gathered, the steps for calculating the UHI are not complex. However, at each point, there are options or choices to consider. What follows is a step-by-step calculation of the UHI and the array of options available for tailoring the construction of the UHI to the data and needs of the analyst.

1. Preparing the Data

This step entails preparing the data for the steps that follow. It is assumed that the data are arranged in a spreadsheet (such as, Microsoft Excel or Google Docs) or statistical software (such as, SAS, Stata, SPSS, R) with the non-overlapping geographic units displayed in each row (one row per geographic unit, with the top row potentially containing the column labels) along with the indicators and geographic data (e.g., place name, geographic ID code, GIS coordinates) along the columns.²⁶ Figure 3 depicts a screen shot of such a layout. There are many choices for software programs that can be used for the purposes of calculating the UHI, a choice which we leave to the user. Developed in tandem with this Handbook is a tool (see Appendix A) in Microsoft Excel for calculating the UHI. The user can manually enter or copy and paste the data into this tool and select among the available options for automatic calculation of the UHI.

Figure 3: Health Indicator Data for Tokyo Prefecture, Japan (2003-2007) in Spreadsheet Format²⁷

	A	B	C	D	E	F	G	H	I	J
1	Area Code	Area Name	Indicator 1	Indicator 2	Indicator 3	Indicator 4	Indicator 5	Indicator 6	Indicator 7	Indicator 8
2	13101	CHIYODA-KU	95.1	99.1	106.8	117.4	95.1	89.2	81.8	98
3	13102	CHUO-KU	96.2	111.7	106.2	123.3	96.1	107	83.6	111.4
4	13103	MINATO-KU	87.7	100.4	97.1	116.3	82.5	97.5	84.5	86.8
5	13104	SHINJYUKU-KU	95.6	100.2	95.3	107.5	103.1	105.3	93.4	92.1
6	13105	BUNKYO-KU	94.7	99.4	101.4	103.8	92.2	99.7	89.8	97.4
7	13106	TAITO-KU	122.9	114.6	119.2	128.6	121.9	110.7	121.2	109
8	13107	SUMIDA-KU	114.2	113.1	119.1	118.3	116.2	107.1	117.2	122.5
9	13108	KOTO-KU	113	112.3	115.9	116.7	106	103.8	123	114.9
10	13109	SHINAGAWA-KU	101.6	102.8	106.9	108.2	108.9	102.6	92.2	99.4
11	13110	MEGURO-KU	82.8	87	88.7	97.1	92.2	85.9	66.8	77
12	13111	OTA-KU	100.3	104.6	106.4	109.7	101	105.5	98.1	100.3
13	13112	SETAGAYA-KU	85	92.1	92.5	102.3	85.2	90.9	75.2	86.2
14	13113	SHIBUYA-KU	85.7	93.2	90.8	107	87.2	91.7	75.6	81.1

In addition to the indicators, the user will need to identify the geographic unit. In this illustration, there are two columns for the municipal identifier: one is alphanumeric (Column A), which, in this case, is

²⁶ If the analyst wishes to quantify the uncertainty in the UHI, then it would also be necessary to provide the variances or standard errors for all indicators that will compose the UHI.

²⁷ Ministry of Health, Labour and Welfare of Japan. 2003-2007 Vital statistics by local public health center and municipality, 2009 update [Japanese]. Incorporated Administrative Agency National Statistics Center (<http://www.e-stat.go.jp/>, accessed 11 September 2014).

the area code the Japanese government has assigned to its municipalities, and the other is text (Column B) (for example, “13101” in Column A is the unique area code for “Chiyoda-ku” ward, shown in Column B). The alphanumeric data will be particularly useful if later merging the UHI with other data, such as census or GIS data. An initial step is to be sure the data are ready to use, which means all values are within the correct range and there are no typographical errors, or in other words, “clean”. A key issue for data readiness is the possibility of missing data. It is important the user considers the amount of missing data on indicators and how the missing data are coded, as well as how the software chosen handles this missing data. A simple approach to handling missing data is to drop units or areas for which data are missing for any of the indicators. This will result in missing UHI values for the missing unit. This will often be tolerable when the number of missing units is smaller in relation to the total number of units. When missing data are more extensive and spread across units, it will result in an unacceptable loss of units from the UHI computation and analysis. In these cases, the solutions for handling the missing data can be more complex and include a variety of methods for imputation (assigning values when missing) of data. For this reason, those with limited statistical resources and expertise at their disposal will do best to avoid indicators for which many data points are missing or seek assistance from a statistical expert familiar with such methods.

Indicators must be arranged in a consistent direction. In general, for the UHI, higher values for indicators should be indicative of better health. For example, consider two indicators of mortality, such as age-adjusted mortality rates and life expectancy. Higher values of mortality rates are obviously indicative of a population’s poorer health; higher life expectancy is indicative of better population health. To rectify such situations, reverse the direction of one or more indicators. For example, order mortality rates in descending order and life expectancy in ascending order so that the early values for each indicate poorer health. The analyst can certainly choose the reverse, but consistent ordering is essential.

In general, it could be informative to check the correlation matrix for potential indicators the analyst believes are arranged in a consistent direction. Positive correlations should be expected, though zero or slightly negative correlations are usually acceptable. However, moderate or large negative correlations are a sign that the indicators are not arranged in a consistent direction and/or that an indicator has a complicated relationship with other indicators or health outcomes. Indicators where correlations with other indicators are unexpectedly negative may complicate interpretation of the UHI; therefore, their exclusion from the index is usually advisable.

The direction of health determinants may sometimes be less clear or differ from expected. Consider the 2000 Census data for Shanghai, China. An indicator for the proportion of households that has its own kitchen, initially thought to correlate with better health outcomes, was found to be lower among the higher-educated and more professionally-employed townships and higher among the less-educated townships on the outskirts of Shanghai. Subsequently, this variable was dropped as an indicator for the UHI.

Some types of indicators will remain readily interpretable if reversed. For example, consider poverty, a well-established determinant of population health, and its commonly used indicator of measuring it—the percent below poverty level. To reverse this or any other variable that is on a percent metric, one needs to subtract the value from 1 for all areas (the complement of the proportion). The reversed indicator or variable reflects the percentage *above* poverty. This indicator could then be included with other indicators scored in the same direction for a UHI where higher values reflect areas with better health or determinants of health.²⁸

28 Even if an indicator, such as mortality rates or household income, is not intuitive if reversed, this does not affect the calculation or interpretation of the UHI. For such cases, we recommend using a simple mathematical operation where the indicator to be reversed is multiplied by -1. This transformation will maintain the scale of the original variable. That is, the difference between indicator values for two geographic units will be the same post-transformation as pre-transformation. Another transformation that will reverse the scale is the reciprocal or $1/I$. However, this transformation is non-linear and does not maintain the original scale of the variable. For this reason and because it is undefined when an indicator value is equal to zero, it is not recommended.

2. Calculating the Urban Health Index

Once the data for selected indicators are properly entered into a statistical program or spreadsheet and are checked for accuracy, missing data, and consistency of direction, they are ready for calculating the UHI. There are two primary steps in calculating the UHI: (1) Standardization of indicators, and (2) Amalgamation of the standardized indicators. Each of these steps is mathematically straightforward, but there are a few things to consider at each stage.

Standardization of the Indicators

The purpose of this step is to adjust for the varying metrics and scales of indicators. The indicator score for each unit is standardized as its difference from the unit with the lowest value divided by the difference between the highest value (maximum) and the lowest value (minimum), or range, as follows:

$$I^S = \frac{I - \min^*(I)}{\max(I) - \min^*(I)}$$

where I is the value of an indicator for a unit, $\max(I)$ is maximum value of I the indicator over all units, $\min^*(I)$ is the minimum value of I over all units minus a small value or alternatively chosen value as discussed below, and I^S is the computed standardized value of I and will satisfy $0 < I^S \leq 1$.

The $\min^*(I)$ and $\max(I)$ have been referred to as “goalposts” by the creators of the HDI.²⁹ In some situations, it might make more sense to choose a theoretical or intuitively justifiable value for the lower goalpost, $\min^*(I)$, rather than the observed minimum. For example, the HDI uses subsistence values as the lower goalpost. The chosen $\min^*(I)$ for the life expectancy indicator, e_0 , is 20 years. The inferred rationale is that selecting subsistence values as the lower goalposts will result in a more interpretable dimension index for the HDI than using the minimum from the observed distribution. Thus, in some instances, alternative goalpost values of $\min^*(I)$ for selected indicators in the UHI are preferable. For example, for some indicators, zero may be a meaningful value for $\min^*(I)$ (i.e., income, proportion unemployed, prevalence rates, mortality rates).

There are just two primary considerations for choosing the $\min^*(I)$. First, the choice of $\min^*(I)$ must be less than minimum of the observed distribution of the index. This is necessary so that I^S will be strictly positive (that is, > 0), a necessary condition for the next step. If one opts to use the minimum of the observed distribution of I , it will be necessary to subtract a small constant in order to ensure the I^S is strictly positive. We would suggest using a constant that is equal to or less than one-hundredth of the minimum value for I . The second consideration is that the chosen $\min^*(I)$ should be defensible for decision making. The I^S is unaffected by the choice of $\min^*(I)$, but the choice of $\min^*(I)$ will often influence the ranking for some units. Though such changes in rank will usually be small for most units, some units do experience more substantial movement. For this reason, a defensible choice of $\min^*(I)$ is important to better withstand potential challenges to the index.

On occasion, it may also be desirable to choose a value for $\max(I)$ other than the maximum of the observed distribution of I . For instance, if the UHI will be computed to examine different temporal periods and trends, then it is important that the goalposts for each temporal period be identical. The Excel-based UHI tool described in Appendix A will default to using the observed maximum and observed minimum (minus a small constant to prevent zeroes in I^S). However, the analyst has the option to enter an alternative maximum or minimum.

29 United Nations Development Programme (UNDP). Human Development Report 2013. The rise of the South: human progress in a diverse world. Technical notes. New York: UNDP;2013 (http://hdr.undp.org/sites/default/files/hdr_2013_en_technotes.pdf, accessed 11 September 2014).

The process for calculating the I^S from values of standardized mortality ratios (SMR) for all causes of death (2003-2007) for selected municipalities in Tokyo Prefecture, Japan is shown in Figure 4 and Figure 5. Municipalities are arranged in ascending order by SMR values to show the minimum and maximum values of this indicator, which are needed for calculation of I^S for this health indicator. The calculation of I^S is shown using the sample minimum as the lower goalpost and using zero as the lower goalpost. In the former scenario, as noted, a small constant (0.01) is subtracted from the sample minimum for purposes of computing the lower goalpost to prevent zeroes.

Figure 4: Illustrating the Calculation of I^S Using the Sample Minimum as the Lower Goalpost

	A	B	C	D	E
1	Area Code	Area Name	Indicator 1	$\frac{I - \min^*(I)}{\max(I) - \min^*(I)}$	$= I^S$
2	13214	KOKUBUNJI-SHI	74.6	$\frac{74.6 - (74.6 - 0.01)}{161.9 - (74.6 - 0.01)} =$	0.0001145
58	⋮	⋮	⋮	⋮	⋮
59	13364	OSHIMA-SHICHO KOZUSHIMA-MURA	124.9	$\frac{124.9 - (74.6 - 0.01)}{161.9 - (74.6 - 0.01)} =$	0.5762227
62	13308	NISHITAMA-GUN OKUTAMA-MACHI	161.9	$\frac{161.9 - (74.6 - 0.01)}{161.9 - (74.6 - 0.01)} =$	1.00

Figure 5: Illustrating the Calculation of I^S Using Zero as the Lower Goalpost

	A	B	C	D	E
1	Area Code	Area Name	Indicator 1	$\frac{I - \min^*(I)}{\max(I) - \min^*(I)}$	$= I^S$
2	13214	KOKUBUNJI-SHI	74.6	$\frac{74.6 - 0}{161.9 - 0} =$	0.4607783
58	⋮	⋮	⋮	⋮	⋮
59	13364	OSHIMA-SHICHO KOZUSHIMA-MURA	124.9	$\frac{124.9 - 0}{161.9 - 0} =$	0.7714639
62	13308	NISHITAMA-GUN OKUTAMA-MACHI	161.9	$\frac{161.9 - 0}{161.9 - 0} =$	1.00

Amalgamation of the Standardized Indicators

Once the values of I^S are obtained for all indicators and units, the next step is to amalgamate the I^S into a single, composite index, or the UHI. The UHI is calculated for each unit by taking the geometric mean of the values of I^S for each unit. In its simplest form, to get the geometric mean of two numbers, you multiply them together and take the square root. When there are j numbers, the formula is:

$$UHI = \left(\prod_{i=1}^j I_i^S \right)^{\frac{1}{j}}$$

where j is the number of standardized indicators, and the UHI is computed by multiplying the values of I^S together for each unit and raising the product to the j^{th} root. A useful property of the geometric mean for the purposes of the UHI is that a given percentage change in any indicator has the same effect on the geometric mean, regardless of the indicator's range. (Since the geometric mean is only defined for positive numbers, the small adjustment to the standardization equation described in the prior step is required.) The UHI value for a unit thus reflects that unit's (geometric) average proportional distance from the lower goalpost, $\min^*(I)$, based on all the indicators included in the construction of the UHI. The calculation is shown in Figure 6 using a subset of data on 10 indicators from Tokyo prefecture, Japan. This UHI used 10 indicators.

Figure 6: Illustrating the Calculation of the UHI

I_{01}^S	I_{02}^S	I_{03}^S	I_{04}^S	I_{05}^S	I_{06}^S	I_{07}^S	I_{08}^S	I_{09}^S	I_{10}^S	$\left(\prod_{i=1}^j I_i^S \right)^{\frac{1}{j}} =$	UHI
0.00011	0.02900	0.09327	0.34770	0.04617	0.17971	0.04326	0.00005	0.05378	0.30025	$(.00011 * .029 * \dots * .300)^{\frac{1}{10}} =$	0.02247
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
0.55332	0.37312	0.93941	1.00000	0.42527	0.28721	0.21761	0.22559	0.32885	0.32894	$(.553 * .373 * \dots * .329)^{\frac{1}{10}} =$	0.40735
1.00000	1.00000	0.19282	0.00019	1.00000	1.00000	1.00000	1.00000	0.60844	0.54147	$(1.00 * 1.00 * \dots * .541)^{\frac{1}{10}} =$	0.32293

Another consideration is whether or not to differentially weight the indicators. The UHI formula provided in Figure 6 assigns equal weight to each indicator so that each indicator contributes equally to the UHI. However, in some contexts, it might be desirable to give some indicators greater weight or other indicators lesser weight when computing the UHI. Guidance on constructing a weighted UHI is provided in Appendix B.

3. Calculating the Error Variance of the Urban Health Index

It would be unusual for indicators to be measured without error. One potential source of error for many indicators is sampling error that occurs when measurements of an indicator are taken on a finite sample from the population. This may be especially true for small geographic area data formed by aggregating individual units. As discussed in the prior section, consider the reliability and validity of indicator data during the indicator selection process. Other sources of error, random and systematic, may also be present and nontrivial, including rank assignment of areas.³⁰ Once indicators are chosen

30 Wolff H, Chong H, Auffhammer M. Classification, detection and consequences of data error: evidence from the Human Development Index. The Economic Journal. 2011;121:843–870.

for constructing the UHI, it is possible to estimate the variance and add confidence bars to plots of the UHI if error variance or standard deviations for the indicators are available. When employing equally-weighted indicators, there are two derivations of the variance of the UHI, one in which normality is assumed and one in which normality is not assumed (see Appendix C).

4. Calculating Summary Statistics for the Urban Health Index

Once the UHI has been calculated for all areas, it will be useful to calculate descriptive statistics for the purposes of interpreting the UHI and state of health for the urban area under consideration. The most useful descriptive statistics are those reflecting central tendency (such as, arithmetic mean or median), dispersion or variation (such as, standard deviation or interquartile range), and quantiles. The central tendency measures inform the analyst as to the “average” proportional distance from the lower goalposts of the average unit. The mean of the UHI is calculated as

$$M_{UHI} = \frac{\sum_{k=1}^n UHI_k}{n}$$

The median, or the 50th percentile, is calculated by ranking the n UHI values and selecting the middle value when n is odd or the average of the two middle values when n is even. Both quantities are calculated by statistical software or spreadsheet programs. The median might be preferred when the UHI distribution is marked by outliers or is skewed (i.e., not symmetrical) since the mean is sensitive to these circumstances.

Dispersion statistics provide an indication of the extent to which the UHI data vary across areas or units. They can mark the extent to which geographic inequities in health determinants or disparities in health outcomes exist within an urban area. Two commonly used dispersion statistics are the standard deviation and the interquartile range. The standard deviation of the UHI is calculated as

$$SD_{UHI} = \sqrt{\frac{1}{n} \sum_k^n (UHI_k - M_{UHI})^2}$$

which is the square root of the average squared distance of the UHI values from the average UHI. Greater values of SD_{UHI} indicate more variability in health determinants or health outcomes, which is often indicative of greater health inequities and inequalities. The interquartile range is calculated as

$$IQR_{UHI} = {}_{UHI}Q_3 - {}_{UHI}Q_1$$

where Q_3 and Q_1 are the third and first quartiles, respectively. To calculate Q_3 and Q_1 , arrange the UHI values in ascending order and calculate the median as described above. Then, split the UHI values into two halves by dividing at the median, but excluding the median. Q_1 , also the 25th percentile, is equal to the median of the lower half or the values below which 25 percent of the UHI fall. Q_3 , also the 75th percentile, is equal to the median of the upper half or the values below which 75 percent of the UHI fall. In practice, interquartile range and standard deviation are readily calculated with statistical software or spreadsheet program.

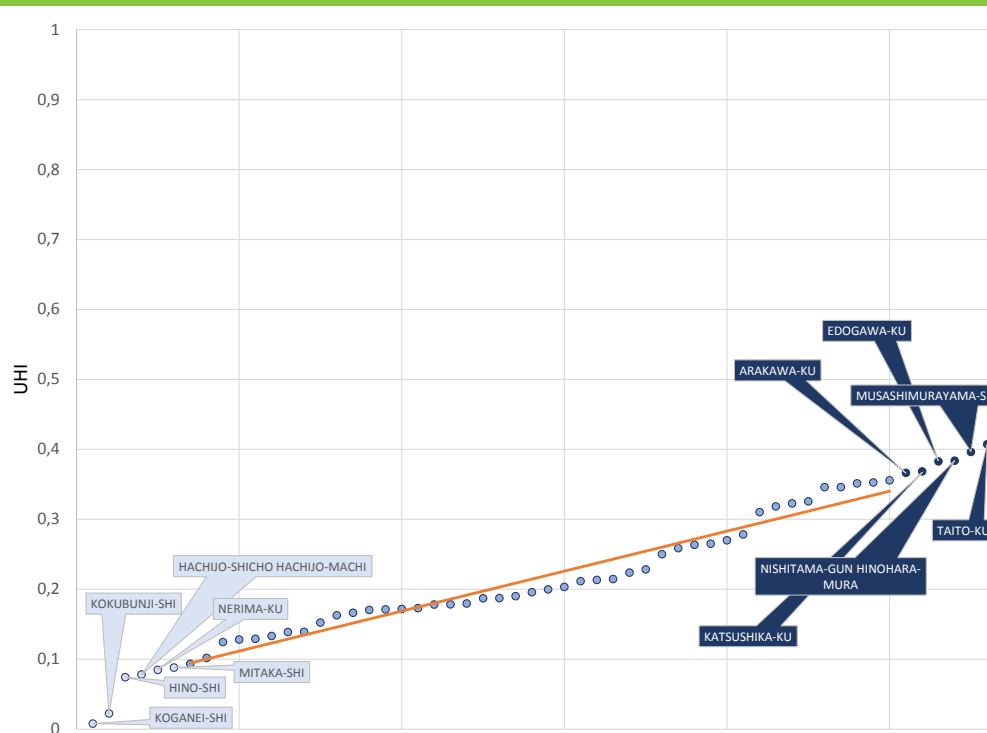
Quantiles are points, plotted at intervals, that divide an ordered variable into equal or nearly equal subsets. Quantiles calculated on the UHI can be useful for grouping areas by their UHI value and demarcating areas that are in especially poor or good health. The analyst may also wish to separate areas into tertiles, which would divide the areas into three equal groups according to their UHI values.

The first tertile would comprise those areas where UHI values are in the bottom 33%; the third tertile would comprise those areas where UHI values are in the top 33%. In addition, quantiles can be divided into quartiles, which subset areas into four equal groups; quintiles, which separate areas into five groups; or deciles, which separate areas into 10 equal groups. The decision to subdivide quantiles is constrained by the number of units and by measurement error in obtaining the UHI (see Appendix C).

5. Creating an Index Plot of the Urban Health Index

Visualization of the UHI can be useful for understanding and reporting the state of health and its determinants for an urban area. One potentially useful visualization for examining the distribution of the UHI is an index plot. The index plot denotes UHI values against their ranked position in ascending order. Obtaining an index plot can be performed with statistical software or spreadsheet software. First, compute and record the ranked position of each area according to its UHI value. Most statistical software and spreadsheets programs have a function for computing the rank of a variable. Then, a scatterplot is obtained with UHI values along the y- (vertical) axis and the rank along the x- (horizontal) axis.³¹ When comparing plots, particularly when the number of areas represented differs across the plots, make sure that the size of the plots and metrics of the axes are identical to ensure that the plots are visually comparable.³² To illustrate, a sample UHI index plot is shown in Figure 7.

Figure 7: Example UHI Index Plot for Tokyo Prefecture, Japan³³



31 With certain software programs, it is sufficient to sort areas in ascending value and then request a line graph of the UHI.

32 The vertical axes should begin with 0 and end with 1. When the number of units varies, the horizontal axes (rank) can be standardized by dividing all ranks by n . For each plot, the unit with the smallest UHI value will have a horizontal axis value of $1/n$ and the unit with the largest UHI value will have a horizontal axis value of n/n or 1.

33 Ministry of Health, Labour and Welfare of Japan. 2003-2007 Vital statistics by local public health center and municipality, 2009 update [Japanese]. Incorporated Administrative Agency National Statistics Center (<http://www.e-stat.go.jp/>, accessed 11 September 2014). The unit of analysis was the smallest-level municipality (namely, non-designated cities, villages, towns, city wards of designated cities, and special wards). The indicators are sex-specific standardized mortality ratios (SMRs) for all-cause mortality and cause-specific mortality for cerebral vascular disease, heart disease, malignant neoplasm, and pneumonia.

An index plot can be enhanced by marking certain areas according to characteristics of the areas plotted. For example, the tails are indicative of greater inequities or disparities. The analyst may wish to identify the units belonging to the top and bottom deciles or by section of the city in which the unit is located, as shown in the index plot. Software programs will often have features to mark such points by color or shape. These distinctive marks can assist in direct identification of the geographic inequities or disparities. Steep plots with more pronounced tails depict greater disparity, as described in the next step.

6. Quantifying Geographic Disparities Using the Urban Health Index

The index plot provides visualization of the extent of geographic disparities for a particular urban area,³⁴ but it is often useful to have quantitative measures for a more precise analysis of geographic disparities. This step describes two measures: (1) the extent to which disparities exist among the units or areas occupying the central portion of the UHI distribution, and (2) the extent of disparities between the extremes of the UHI distribution.

Assessing Disparities Among Areas in the Middle Range of the UHI: The Disparity Gradient

Often, units occupying the midsection of the UHI distribution will display linear increments. That is, the difference between adjacently ranked units will be approximately equal for units not at either extreme of distribution. Given this, the slope of the line fitted through the UHI values arranged in ascending order can be used to quantify the UHI health disparities gradient among the non-extreme areas or midsection. Though the general approach is the same, the specific steps for calculating this quantity will depend on the software being used.

First, define the criterion for selecting the units or areas considered within the “middle” quantiles. One approach is to first define the criteria for areas at the extremes. For example, if the analyst defines the extremes as those areas belonging to the top or bottom deciles (top and bottom 10%) on the UHI, then the middle section of the distribution contains the central 80% of area. Then, use ordinary least squares (OLS) linear regression to fit the line through these points along the midsection and calculate the slope of this line.³⁵ In practice, this can be accomplished by using statistical software to regress the UHI values on their ascending rank. A spreadsheet function can also be used. For example, the =SLOPE function in Microsoft Excel can be used to obtain the value of the OLS slope by regressing the UHI variable on the rank variable.³⁶ The rank variable for the units in the midsection can be computed in Excel using the =RANK.AVG function.

The measurement scale of the slope when regressed on the rank variable represents an estimate of the average UHI difference between each adjacent unit and is affected by the number of units. However, it cannot be compared to a slope calculated for another area with a different number of units. Therefore, it is recommended to rescale the rank variable to be used in the regression by dividing each unit's rank by the total number of areas or units that constitute the midsection. The slope obtained from a regression of the UHI variable on the rescaled rank variable is the estimated difference between the two most extreme areas within the midsection. This rescaling of the rank for calculation of the health disparities slope or gradient facilitates comparison of slope values when the number of units varies (e.g., between two urban areas or over time). To get the actual change in UHI for a unit change in rank, convert back by multiplying the observed slope by the number of areas in the midsection.

34 Choropleth and other map types for the UHI are described in the next section, Mapping the Urban Health Index.

35 OLS regression will fit a line that minimizes the summed squared distance between the observed UHI values and the UHI values predicted by the linear function that defines the line.

36 Be certain that the regression is happening in the midsection and not in all areas. In Excel, use the =PERCENTILE.INC function to obtain the 10th and 90th percentiles. Then, create a new variable that copies the UHI values only for units with UHI values greater than the 10th percentile but less than the 90th percentile by using the IF and AND functions (e.g., =IF(AND(AK9>0.1,AK9<0.9),AH9,""). Then, this “trimmed” variable is regressed on the rank variable using the =SLOPE function. Statistical software such as SAS or SPSS contain functions for selecting units for regression analysis.

The slope of the midsection provides an estimate of the extent of heterogeneity in the overall grouping. A steep gradient suggests greater disparities across the group. A flat gradient, in contrast, suggests relative homogeneity in the midsection. To illustrate with the UHI data from Tokyo prefecture, Japan, partially shown in Figure 7, the UHI rank is computed for the middle 80% of values with UHI values $> .0907$ (10th percentile) and $< .3610$ (90th percentile). These points with the medium-blue hue are then rescaled as described above. The slope of the line fit through the data points with the medium-blue hue in Figure 7 was calculated as .25. This value indicates that the UHI values for this midsection spans approximately .25 UHI units. Reverse the effect of rescaling by multiplying this slope by $1/44$ equals .0057 (i.e., the average distance between each adjacent unit is .0057 UHI units). The gradient for the midsection of the distribution for the Tokyo prefecture might be considered relatively flat.

Assessing the Extent of Disparity between the Outer UHI Quantiles: The Disparity Difference and Ratio

The second measure of inequities and disparities captures the differences between the areas with highest UHI values versus those with the lowest UHI values. In the example shown in Figure 7, the upper 10% and the bottom 10% of areas are based on their UHI values. This measure is computed in one of two ways.

1. Calculate the Difference in the Mean Values of the Top and Bottom Quantiles of the UHI

This first method is the difference between the means of the upper and the bottom quantiles (equivalent to a risk difference). The difference of means, rather than a variant calculated using the medians (equal to the difference between the 95th and the 5th percentiles), is used to accentuate the difference between the extremes (the medians are less sensitive to extreme values).³⁷ Regardless of which central tendency measure is chosen, higher values reflect greater disparity between those with the best level of health determinants or health outcomes and those with the worst. To illustrate, see the data displayed in Figure 7. The means of the bottom 10% (light blue points) and the top 10% (dark blue points) are computed as .0593 and .3842, respectively, with a difference equal to .32. In comparison, the difference based on the medians is equal to .31.

2. Calculate the Ratio of the Mean Values of the Top and Bottom Quantiles of the UHI

The second method calculates the ratio of the mean of the upper part of the distribution to the mean of the lower part (equivalent to a risk ratio). This is a marker of the overall disparity between the best and the worst area units. A variant of this calculation would use the median rather than the mean. The ratio itself may be viewed as a relative risk comparing those exposed to “benefits” to those unexposed. To illustrate, see the data displayed in Figure 7. The ratio of the mean UHI for the top 10% to the mean UHI for the bottom 10% equals $.38/.059$ or 6.48. For comparison, the ratio computed based on the median of the top decile (.3833) to the median of the bottom decile (.0762) is equal to 5.02.

It is noteworthy that the ratio quantity is particularly sensitive to the mean of the lower extreme group and, thus, the choices for lower goalposts used when standardizing the indicators (see Standardization of the Indicators). To illustrate, consider if all units in Figure 7 improved by exactly .10 UHI units, such that the means of the bottom and top 10% were .159 and .484, respectively. Although the difference between these values remains the same, the ratio changes from 6.48 to 3.04, a decrease of more than 50%. If the scale of measurement for the UHI is on a ratio measurement scale, meaning that a value of zero has true meaning, then the ratio quantity for calculating health disparities or inequalities between the extreme best and worst area units is interpretable and may be preferred to the difference quantity. However, when the UHI does not possess ratio measurement scale properties, the health disparities ratio does not indicate the factor by which the top extreme group is better or worse (depending on the direction of the indicators), than the bottom extreme group. When an absolute meaning is difficult to interpret, a health disparity difference may be a better measure for comparing areas.

³⁷ However, if the analyst does not wish for one or a few very extreme outliers to exert a substantial influence on the difference, then the median can be used instead.

MAPPING THE URBAN HEALTH INDEX

Mapping the UHI provides the essential information needed by urban planners, public health professionals and decision makers to gain insight into the spatial patterns of health inequities.

Mapping the UHI requires a Geographic Information System (GIS) program and geospatial vector data of the boundaries of geopolitical or geo-statistical units used in the calculation of the UHI. A commonly used commercial GIS program is ArcGIS for Desktop® developed by Environmental Systems Research Institute (ESRI, Redlands, California). It currently works only with the Microsoft Windows operating systems. Users also can access free and open-source GIS programs. For instance, QGIS (<http://www.qgis.org/en/site/>) is compatible with both Windows and Mac operating systems. The common geospatial vector data format is the shapefile developed and regulated by ESRI and is mostly interoperable among ESRI and other GIS programs. Besides QGIS, a number of other free programs can be accessed such as GRASS GIS (<http://grass.osgeo.org/>), or MapWindow GIS (<http://www.mapwindow.org/>). See a comprehensive list at www.freegis.org. Even though the user interface of these free GIS programs may differ, the process of joining the UHI results in an Excel (or CSV) file with a geopolitical or geo-statistical boundary shapefile remains the same. The GIS dictionary at <http://support.esri.com/en/knowledgebase/Gisdictionary/browse> provides a comprehensive list of GIS terms, such as *shapefile* or *vector data*. Descriptions of comma-separated value (CSV) files are available at http://en.wikipedia.org/wiki/Comma-separated_values.

Mapping the UHI results in general will involve three steps (see Figure 8). (1) Make sure both the UHI data file and the attribute of the boundary shapefile have a common “key”. The key, which could include county name, zip code number, or other identifier must be unique and have the same format (e.g., either numeric or string). Because a shapefile has a database file (.dbf) to keep its attribute, Excel may be used to open the .dbf file to identify the common key. (2) Add both the UHI result data and the shapefile data to a GIS program, and, using the “Join” function, append the UHI result to the boundary shapefile for mapping. (3) The GIS program may be used to create a choropleth map based on UHI values. To be consistent with UHI ranking, a 10-group classification with quantile rule is used to identify the units that are in different decile groups (i.e., such as examining the lower 10% and the upper 10% of the units). The ColorBrewer is a free online cartographic program that allows the analyst to select color schemes for creating choropleth maps based on the number of classes and the nature of relationship among them (e.g., sequential, qualitative, or diverging).³⁸ See Appendix D for detailed instructions for using ArcGIS 10.1 and QGIS 2.2.0 in the Windows operating system.³⁹

To obtain the shapefile necessary for mapping the UHI, the analyst is encouraged to contact the administration having jurisdiction in the study area for any GIS boundary data pertaining to the boundaries that define the units for which the UHI was constructed. If data from the local administration is unavailable, a number of data depository sites may have boundary data available. A few are listed here:

- StatSilk: <http://www.statsilk.com/maps/download-free-shapefile-maps>
- DIVA-GIS: <http://www.diva-gis.org/gdata>
- ESRI Census boundaries in the United States: https://www.census.gov/geo/maps-data/data/prev_cartbndry_names.html

³⁸ Brewer C. Designing better maps: A guide for GIS users. Redlands (California): ESRI Press; 2005.

³⁹ Analysts may practice with these instructions using data files that can be downloaded from http://www.who.int/kobe_centre/publications/urban_health_index_toolkit/en/ (accessed 16 October 2014).

- GEOFABRIK: <http://www.geofabrik.de/data/shapefiles.html>
- OpenStreetMap: <http://wiki.openstreetmap.org/wiki/Shapefiles>
- Natural Earth: <http://www.naturalearthdata.com/tag/world-file/>
- MAPCRUZIN: <http://www.mapcruzin.com/download-free-arcgis-shapefiles.htm>
- Global Administrative Areas: <http://www.gadm.org/>

It is important to note that geopolitical or geo-statistical boundaries often experience changes, thus requiring users to pay particular attention to ensure the boundary shapefile obtained represents the boundaries applicable to the indicators. These changes are pronounced, especially in fast growing regions where small units such as towns or villages are absorbed into neighboring large cities to formulate new city boundaries. Therefore, make sure the boundary shapefile matches the years for the indicator data being used to construct the UHI. If the shapefile is unavailable, data pre-processing would be necessary to make the data congruent with an available shapefile. This could be facilitated by a GIS program mentioned earlier, or by a trained GIS consultant. Another challenge arising from changing boundaries is when different indicators are collected at different time periods during which the boundaries change. Analysts need to identify the indicators that are available in the various time periods. In order to accommodate the changes in boundaries, use the “Union” function in GIS geo-processing tools to compute a geometric union of the input boundary shapefile with the support of a GIS consultant. The “Overlay” function allows all features and their attributes to be written to an output shapefile. The output shapefile will have all geometries from the input boundaries.

Figure 8: Flowchart of Steps for Mapping the UHI

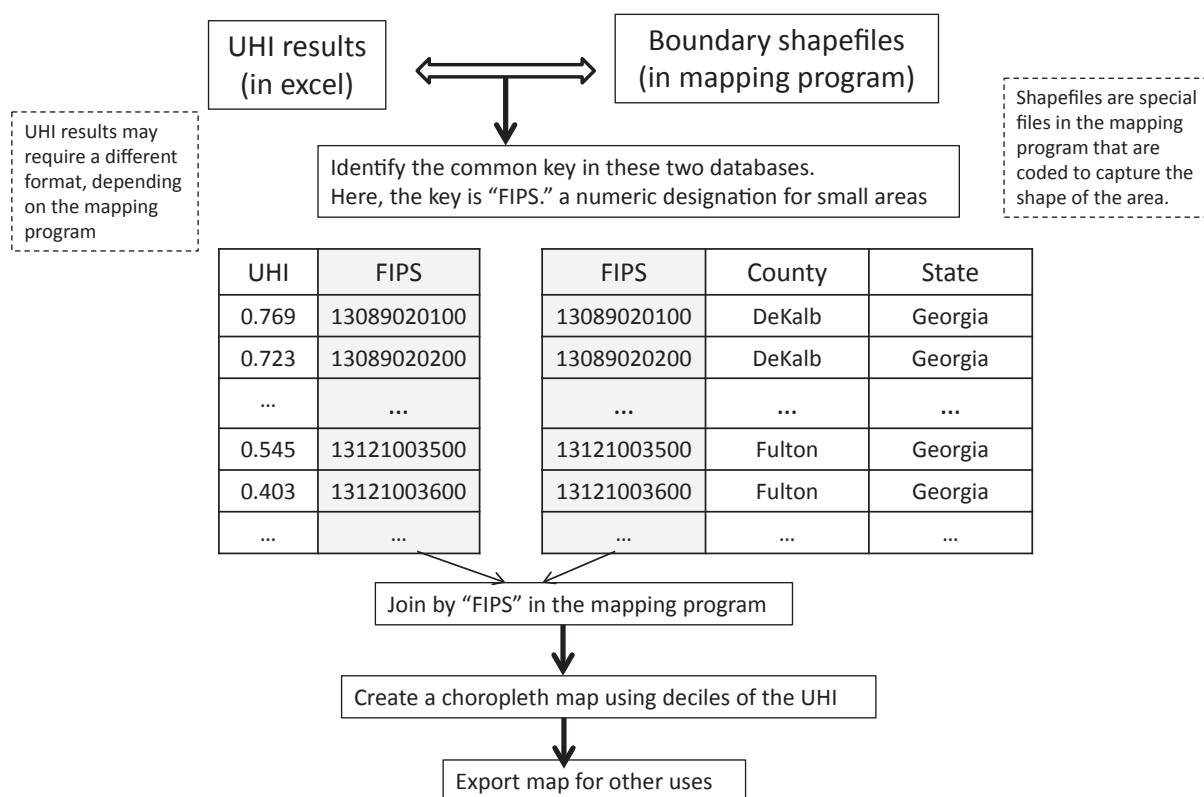
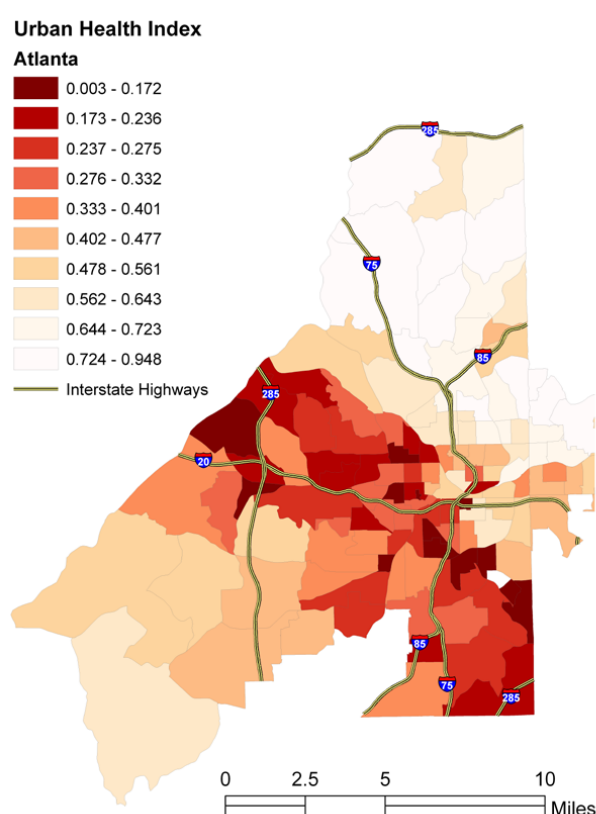


Figure 9 presents an example of mapping the UHI using health determinants at census-tract level in Atlanta, Georgia. Mapping provides the essential information needed to understand the health disparities in a spatial context. It reveals a pattern of geospatial inequity in the interested area, and allows urban planners, public health professionals, or policy makers to gain insights into the inequities in urban areas and pockets of deprivation (called *hot spots*). Such information may provide evidence and guidance into possible interventions toward control or elimination of those inequities. As seen in Figure 9, geographic disparity in health determinants is pronounced in Atlanta. The dark-red belt orientated from northwest to southeast are places where health determinants are severely disadvantaged. The legend indicates the disparities in the dark-red belt with UHI values below 0.2, compared to North Atlanta where UHI values are above 0.7.

Another advantage of using GIS to map the UHI is the ability to integrate information from other sources for decision mapping. For example, the analyst could, in tandem, utilize a crime map or a disease map overlaying the UHI map in Atlanta and reveal the spatial coincidence of those occurrences. A section of the deprived area, known as “The Bluff” in Northwest Atlanta, has struggled with profligate crime and violence, availability of drugs, high poverty rates, and is a reservoir of a variety of diseases, such as HIV. This area has become an area of focus for multiple improvement plans, including a number of parks and trails, as well as increased police presence.

Figure 9: An Example: Mapping the UHI in Atlanta, Georgia⁴⁰



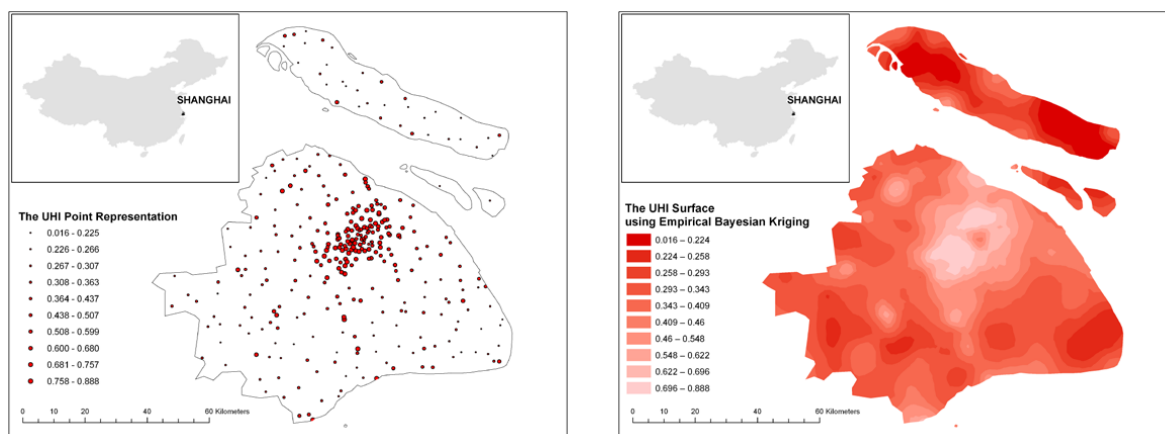
40 US Census Bureau. American Community Survey, American FactFinder [website]. (<http://factfinder2.census.gov>, accessed 12 September 2014). The analysis used the U.S. census tract and 5-year estimates of the following UHI indicators: percent employed, percent above the poverty level, percent with high school degree or higher, percent with bachelor's degree or higher, household mean income, household median income, and percent of households not headed by a single female with children under 18 years of age.

In addition to mapping, various spatial analysis techniques can be utilized to shed light on the geographic disparities in health using the UHI. One popular approach is to detect clusters (or hot spots) of either high or low UHI values using a cluster analysis method. In a multi-year UHI analysis, what areas are clustered with deteriorating or improving health determinants or outcomes? How do patterns and clusters of different health determinants or outcomes compare to one another from year to year? Two common cluster analysis methods are Anselin Local Moran's I or Getis-Ord G_i^* .^{41, 42, 43, 44} Explanations of the two methods are given in *The ESRI Guide to GIS Analysis*.⁴³ Both methods are implemented in ArcGIS, with the spatial statistics extension required, or in R software (<http://www.r-project.org/>).

When the boundary shapefile of geopolitical or geo-statistical units used in the calculation of the UHI is unavailable, collect a point shapefile, such as the geometric centers, to map the UHI or to conduct cluster analysis. In Figure 10, graduated or proportional symbols may be employed to visualize the UHI using different point sizes. Spatial smoothing techniques may be used to create a "surface" with a visual effect similar to a choropleth map. Smoothing techniques, such as the Inverse Distance Weighted method or Kriging methods, are available in both ArcGIS and QGIS. For detailed smoothing instructions in ArcGIS, refer to *Geostatistical Analyst Tutorial*.⁴⁵ *Spatial Statistical Data Analysis for GIS Users* may also be useful for further reading.⁴⁶

Figure 10 shows how the UHI in Shanghai, China, at a township level is transferred to a surface using the Empirical Bayesian Kriging technique in ArcGIS. Compared to the point map, the surface clearly shows the spatial pattern of geographic disparity and reveals the local pockets of poorer urban health determinants (indicated by darker colors) neighbored by better determinants (indicated by lighter colors).

Figure 10: The UHI Representation Using Point (left) and Surface (right) After Smoothing⁴⁷



41 Anselin L. Spatial econometrics: methods and models. Dordrecht (the Netherlands): Kluwer Academic; 1988.

42 Anselin L. Exploring spatial data with GeoDa: a workbook. Urbana: Centre for Spatially Integrated Social Science; 2005 (<https://geodacenter.asu.edu/system/files/geodaworkbook.pdf>, accessed 12 September 2014).

43 Mitchell A. The ESRI guide to GIS analysis, volume 2. Redlands (California): ESRI Press; 2002.

44 Getis A, Ord JK. Local spatial statistics: an overview. In: Longley P, Batty M, editors. Spatial analysis: modeling in a GIS environment. Cambridge: Wiley; 1996:261–267.

45 Geostatistical analyst tutorial. Redlands (California): ESRI Press; 1995–2010 (<http://help.arcgis.com/en/arcgisdesktop/10.0/pdf/geostatistical-analyst-tutorial.pdf>, accessed 12 September 2014).

46 Krivoruchko K. Spatial statistical data analysis for GIS users. Redlands (California): ESRI Press; 2011.

47 Shanghai, 2000 Population Census Data Assembly. Ann Arbor (Michigan): University of Michigan China Data Center (<http://china-datacenter.org>, accessed 12 September 2014). The UHI unit was town, and UHI indicators were: percent non-agricultural population, percent employed, percent professional and managerial jobs, percent with middle-school degree or higher, percent with high school degree or higher, percent with junior-college degree or higher, percent of household cooking using gas and electricity, percent of household with tap water, percent of household with own bath heating facilities, and percent of household with own lavatory.

ANALYSIS AND REPORT WRITING

For those reporting the results of a UHI to policy makers, they must take into consideration the context in which they work. An effective approach might be to prepare a one-page, easy-to-read summary, separate from the detailed report.

The person who does the analysis and report will probably not be the person or people who will act upon it. The policy makers or decision makers (DM) are likely to work in a vastly different arena from the analyst. They are likely to have different priorities and may not have the technical expertise to fully understand the report data. For the UHI to have the greatest effect, the report should be tailored to the time, knowledge, and interests of the DM. However, the details and methods will be important to other researchers or professionals; so those should also be provided. The solution is to write two reports. The first report should be one page or less and should be just a terse summary. The second will be a detailed monograph. The summary may be the only thing a DM will read, but the detailed monograph is available to the DM and anyone else who may find its contents of value. Please note that this opening page differs from an “executive summary,” which is a brief, but more comprehensive summary of major findings.

The One Page Summary

The first page of the Report should contain a heading with the title, author, and date. This is followed by a set of brief paragraphs that summarize the report’s main points:

TITLE: *The Distribution of an Index of Disparity*
Authors: -----
Date: ----

Purpose: *To determine whether the death rate and related determinants in [urban area] are evenly distributed or if there are some areas at higher risk.*

Method: *We used the Urban Health Index, a flexible metric that permits amalgamation of health indicators and is easily transferred to a map for visual depiction. We used [ward, census tract, postal code] data and combined indicators for overall mortality, infant mortality, maternal education, and access to clean water.*

Findings: **1.** *The map of [urban area] displays the intense concentration of high deprivation in the Southeast corner of the city, with lesser concentration in the Southwest and Central areas.*
2. *Some areas reflect better health at a scale 11 times greater than other areas, indicating enormous health disparities*
3. *For infant mortality alone, the Urban Health Index varies by as much as 20-fold.*

Implications: *These findings document the considerable disparity in health, living conditions, and opportunity for social advancement. They serve as a baseline to measure progress in affecting changes in people’s living conditions and health.*

Recommendations: *As demonstrated in other settings, improving maternal education may have the greatest single impact on these disparities, and can be achieved at relatively low cost.*

Comment: *In addition to programs for amelioration, investing in surveillance of disparities can bring attention to the problems of the poor, and create the basis for policy change.*

This brief summary is the most important part of the report. It should take the DM less than a minute to scan and pull from it the major points. Though it appears first, it may be helpful to write it last, when the important ideas have crystalized in the mind of the analyst. The language should be simple and only critical details should be included, such as the hardest hit areas reflect a disparity six times worse than the healthiest areas. A narrative writing style is preferred, but it may be useful to put the major findings in an outline format. However, do not put the whole summary in an outline format, as the impact of the findings can be vitiated.

Written recommendations may not be appropriate in some cases. Implications and comments should be rendered in a neutral, non-exhortative way. Editorializing may be appropriate in some circumstances, but a great deal depends on the administrative context and the relationship of those constructing the report to those reading it. A summary paragraph is required and should be tailored to local circumstances.

The Detailed Report

The detailed report is an opportunity to recreate the thinking and rationale behind the data choices, the process of analysis, and any major findings in the UHI. The report is a comprehensive recounting of the process used to produce the final graphs, tables, and narrative, as well as an evaluation of the results. The actual elements and content will vary with the availability of data and the social and political influences that generate the report. A template for a visual summary of the UHI is included in Appendix E. It can be included in the detailed report or used as a handout to accompany the one-page summary. A presentation template for communicating the results is included in Appendix F. The following outline is a suggested framework for such a report.

1.0 Background

- 1.1** Summarize the general circumstances of the urban area for analysis.
 - 1.1.1** Include dimensions, geophysical characteristics, population density, climate, heterogeneity, and recent growth.
 - 1.1.2** List public services, such as water, roads, transportation, sanitation, sewage, and power access.
 - 1.1.3** Relate social infrastructure by per-capita spending, such as education, medical care, and legal services.
 - 1.1.4** Provide information on economic setting, such as median income, employment by age and sex group, and any known disparities.
- 1.2** Include a concise statement of the elements motivating this investigation of health levels, health determinants, and health disparities.
- 1.3** Present the foremost issues and most pressing problems the report hopes to address. For example, the area's access to clean water or to health care, or substandard housing.

2.0 Data sources

- 2.1** Provide a detailed discussion of the data sources available, including the details of access (manner and dates of access) and the years the indicators were measured.
 - 2.1.1** For each data source (i.e., census, ad hoc surveys, DHS surveys, etc.), include information for available:
 - 2.1.1.1** Health indicators
 - 2.1.1.2** Health determinants
 - 2.1.1.3** Health outcomes

2.1.1.4 Environmental indicators

2.1.2 Present an evaluation of the area level.

2.1.3 Provide ancillary resources that can assist in deciphering the data.

2.1.3.1 “Upstream” data on general economic and social conditions.

2.1.3.2 Include methods for applying larger area data for smaller area estimation.

2.2 Indicate the rationale for indicator selection:

2.2.1 Include the reasons for creating the set(s) of indicators.

2.2.2 Relate the rationale for inclusion of indicators from different domains.

3.0 Methods for collecting and assembling data

3.1 Provide a general background on acquisition of the data.

3.1.1 Discuss the data that could not be obtained.

3.2 List the procedures required for editing and formatting data for use in the construction of the UHI(s).

3.3 Note the methods for constructing the UHI:

3.3.1 Use the description in this handbook, if necessary.

3.3.2 Describe any deviations from the standard method and stress how they were necessitated by the availability and format of the data.

3.3.3 Provide a rationale for specific choices made in assembling the UHI.

4.0 Results

4.1 Provide the following for each of the UHIs calculated:

4.1.1 Graph of the UHI (y-axis) against the rank ordering of the area (x-axis).

4.1.2 Calculation of the disparity ratio and/or difference.

4.1.3 Calculation of the disparity slope.

4.1.4 Map of the area depicting deciles of UHI, with hues from lightest (best) to darkest (worst).

4.2 For each UHI calculated, provide commentary on interpretation of the four elements in 4.1.

4.2.1 Thinking of the disparity ratio as a relative risk, provide a qualitative assessment of its size.⁴⁸

4.2.2 Thinking of the disparity slope as a measure of heterogeneity, provide a qualitative assessment.

4.2.3 Provide a detailed description of the worst areas and contrast these against the best.

4.2.3.1 Return to section 1.1 and characterize the worst areas with regard to indicators and variables.

4.3 Calculate target values (goals) for indicators and determinants that would reduce the observed disparities by some measurable amount.

4.4 For UHI variables in which the joint effect of indicators and determinants are assessed, describe the implications of the interaction of health outcomes and health determinants. For example, if access to clean water and infant mortality are included, note the potential implication that can be drawn.

5.0 Discussion

5.1 Discuss the consistency and intensity of the findings. For example, discuss the worst areas with regard to all the other factors examined.

5.2 Discuss the implications of the findings for social policy.

5.2.1 Consider human suffering, economic deprivation, the social cost of the loss of human capital, the barriers to developing an educated and healthy work force, etc.

5.2.2 Discuss the feasibility of reaching the goals (see 4.3) that would ameliorate the disparities

⁴⁸ The disparity difference can be interpreted with respect to historical values of the disparity difference.

6.0 Recommendations (This section is optional and depends on local administrative circumstances)

- 6.1** Recommend the constellation of target data required to refine the assessment of disparities.
- 6.2** Recommend priority areas for consideration for amelioration of disparities.
- 6.3** Recommend target values for indicators that could improve disparities by a given percentage.
- 6.4** Recommend enhancements to current surveillance and analysis that would improve the ability to assess the problems identified.

As noted, this outline is only a framework for a report. The details will vary considerably depending on the local administrative circumstances, the availability of data, and the relationship of the analytic activity to the decision making process. Strive for an even-handed, non-advocacy tone. Let the data and findings speak for themselves. It is highly likely that all persons concerned with this process are well aware of the considerable disparities that exist in their urban areas. This analysis can quantify those disparities, display them, provide mechanisms for setting goals for its amelioration (see footnote 53), and indicate a process for monitoring those goals. Acceptance of this process may be the first step in the formation of public policy that may address any given issue.

LIMITATIONS AND CONDITIONS FOR INTERPRETATION OF THE URBAN HEALTH INDEX

The UHI's strength is not only its simplicity but also its flexibility for adaptation to the local urban context or for use across multiple urban areas. This is vital for its usefulness to local urban planners and policy makers. In this section, factors that limit or condition the interpretation of the UHI are given.

The limitations and conditions described in this section are not specific to the UHI. Rather, they are pertinent to many, if not most, existing health indices that are currently in use. Despite its general relevance, its application to the UHI will facilitate appropriate explanation of indicators, data sources, and interpretation.

Reliability and Validity of Indicators

To an extent, the UHI can be quite forgiving of the shortcomings of data to be used as indicators for the UHI. The index, as a composite of several indicators, may contain less errors than the indicators themselves. The presence of well-measured indicators may compensate for other less reliable indicators which are correlated. Select indicators and data sources carefully (see Selection of Indicators: A framework to draw from), as the interpretation and usefulness of the index could be undermined by indicators with poor validity or reliability. The potential threats to reliability and validity will often be indicator specific. While too numerous to mention all of them here, a few are mentioned given their relevance to the UHI.

When intra-urban units are of interest, as will often be the case for the UHI, these suburban areas will often be characterized by relatively small populations (<5,000). When data are not based on a census, samples from these subareas may be too small for reliable measurement (sampling error). In some instances, government or governmental entities may even restrict access to small-area estimates to protect against misuse. An example is the 1-year U.S. American Community Survey (ACS) data. These data are not readily available for areas with populations less than 65,000 owing to poor precision at smaller population levels. Even when data are collected from the entire population, as in a census, measures of uncommon or infrequent events may render those estimates unreliable. For example, certain cause-specific mortality events. In these instances, there are possible solutions, although none are without potential drawbacks:

- Employ sophisticated empirical Bayesian smoothing methods to obtain more precise area-level estimates. These methods “borrow” on information from adjacent areas to obtain more reliable estimates for a particular area. However, this method may obscure the extent to which there is heterogeneity or differences in health outcomes and/or indicators. When using such methods, it is important to understand and communicate that the UHI may underrepresent the extent of differences and health disparities within or among urban areas.
- Use data that have been aggregated over multiple years. For example, in contrast to the 1-year ACS, the 3-year ACS are available for areas with population of at least 20,000, and 5-year ACS data are available for all areas enumerated by the U.S. Census Bureau. The limitations of this solution

are that multiple-year data are less recent and relevant and multiple-year data may not always be available.

- Aggregate adjacent areas into larger areas where sufficiently precise estimates are available. The challenges with implementing this approach are that there are a myriad of options and methods for aggregation, and there are no clear standards for determining which method is best. In addition, the aggregated units are likely to be less homogenous and may be less relevant for urban planners deciding health policies.

Sampling error is not the only threat to the reliability and validity of indicators. It may even be a less pernicious threat because it is often measureable through statistical methods; it is random (non-systematic); and there are methods, such as those described earlier, for addressing it. Non-random sources of error are often more problematic because they are more difficult to identify and to suppress, and, by their nature, bias area-level indicator estimates. For area-level estimates, non-random error may occur when certain subpopulations with a non-random, spatial distribution are unintentionally under- or over-sampled. For example, when a census experiences under-participation by certain recent immigrant populations. The homeless, recent immigrants, poor, and other marginalized populations are likely to be missed in surveys. Areas with fewer resources may have more systematic biases in the measurement of indicators than resources of more affluent areas.

Another situation associated with non-random error is if an analyst employs a proxy measure for an indicator. For example, when an analyst uses data on anti-retroviral (ARV) use as a proxy for HIV prevalence. The data are likely to be biased due to confounding factors such as HIV awareness and access to ARVs. Low-income areas may have higher HIV prevalence, but the pattern may be underestimated because these persons are less aware of HIV status and may have less access to ARVs. While there is no single solution for this issue, a point of practical advice is to incorporate multiple indicators within a domain in hopes that the biases and weaknesses of some indicators will balance against those of other indicators when aggregated into the UHI. There is no guarantee that such a strategy will work. There may even be the potential for exacerbating the issue. But, when the selection of multiple indicators within a domain is conducted with thoughtful consideration of each candidate indicator's potential limitations, the resultant UHI will be more valid and reliable than its components.

Once candidate indicators are chosen and aggregated into the UHI, document the data sources and describe the limitations of the indicators. These limitations may necessitate that any interpretative statement involving the UHI be appropriately qualified. One approach to qualifying statements made of the UHI is to calculate and provide standard error or confidence interval estimates. When variance or standard error estimates are available for each indicator, it is possible to calculate the variance (see Appendix C). The misinterpretation of the UHI can be prevented by providing confidence intervals in reports about the UHI with descriptions of the confidence intervals. Descriptions should be written with policy makers in mind.⁴⁹

It is anticipated that many analysts and policy makers will be tempted to rank different geographic areas according to their UHI values. Rankings are often intuitive (sometimes deceptively so) and can easily be communicated to various stakeholders. However, it should be kept in mind that UHI ranks are also subject to the same unreliability and biases that affect the UHI. Estimating error variance of ranks is a rather intractable statistical issue. Some practical guidance is provided to minimize misinterpretation

49 Keep in mind that such confidence intervals are only able to reflect the sources of random error that are captured by the indicator error variance estimates. When sources of error are not captured in these estimates, they cannot be reflected in the confidence intervals.

of rankings. Comparing areas with adjacent rankings should be avoided, or interpreted with extreme caution. For example, to say that area X, ranked 30th on the UHI, has less-desirable health than area Y, ranked 29th, ignores the uncertainty in these measured rankings. Thus, such statements should not be made unless they reflect areas whose UHI confidence intervals do not overlap.

In addition, avoid over-interpreting shifts in rankings over time. For example, avoid assumptions when area X is ranked 30th at time t , but is later ranked 31st at time $t+1$. Not only could area X have improved in its health indicators (just not as much as other areas with adjacent ranks at time t) it may also be that the change in UHI was not statistically reliable. Perhaps more useful than rankings of the UHI, and possibly less prone to misuse, is the practice of grouping areas based on their UHI scores. Record these groupings in the report. For example, areas could be grouped into tertiles consisting of the top-third, middle-third, and bottom-third of the index. If larger numbers of areas and more reliable indicators are used, finer groupings, such as quartiles, quintiles, or deciles are possible.

Consistency of Indicators

Another important consideration will be the consistency of indicators and their measurements over time and across space. In many cases, indicators available for one urban area or at one point in time are not measured or available in another urban area or at another point in time. Consider a common measure of educational achievement, the percent having graduated from high school – that might be chosen as an indicator for a UHI of health determinants. However, measures of high school graduation rates may vary across urban areas or within large urban areas with autonomous educational institutions. For example, one area may base its high school graduation rate on those who graduated within 4 years of entering high school, whereas another area may base it on those who graduated within 5 years.⁵⁰

Similarly, such official measures may change over time, or the methods of measurement may change (e.g., when paper-based surveys change to telephone-based surveys). Such changes may confound interpretation of differences among areas or changes over time. When measurement methods or definitions do change or vary, it is important to make note of these changes in any report and to appropriately qualify or condition interpretations. Where it is possible to calculate rates on new data using older definitions, test the impact of such changes. However, such calculations are often not possible or even feasible.

Though it is difficult to generalize, results of indicator exchangeability tests indicate that when the UHI is composed of several highly correlated indicators, changing measurement definitions or methods may only have a minimal systematic impact on the UHI and/or the UHI ranking order. In relation, changes or differences may occur in calculating the UHI, specifically during the step of standardizing the indicators. When the UHI is being used to make comparisons over time or across areas, it is critical that the indicators be standardized with the same goalposts (see Standardization of the Indicators). Failure to do so may result in the appearance of differences or changes that are entirely artifacts of the different goalposts. Fortunately, it is easy to re-calculate the UHI with new goalposts.

50 Areas may vary in whether they include students who move out of the jurisdiction and whose graduation status is unknown in the denominator and/or numerator when calculating the rates.

UHI: Relative versus Absolute Interpretations

Because indicators for the UHI are standardized to an area's position relative to lower and upper goalposts, which may be determined based upon the areas at the maximum and minimum, the UHI must be interpreted relative to these goalposts. Without explicit reference to the goalposts, interpreting a UHI value of .5 as being indicative of poor health or good health is difficult. Interpretation of a particular UHI score or range is relative to these goalposts and should be relayed in the report. The more meaningful and relevant the goalposts, the more easily the UHI values beyond just relative comparisons of areas can be interpreted.

Also consider the measurement scale of the UHI. Often, indicators will be measured on a ratio scale, where statements such as, "A mortality rate of 40 per 100,000 is twice as large as a rate of 20 per 100,000" can be made. Indicators are on a ratio scale when they have a true zero value where a value of zero indicates the absence of a quantity being measured.⁵¹ In contrast, indicators measured on an interval scale cannot have statements such as, "A value of 40 on this metric is twice as large as 20." Many subjective ratings or psychological constructs are measured on an interval (or ordinal) scale. When all indicators are measured on a ratio scale, the UHI will also have ratio scale properties, if and only if the standardization process uses zero as the lower goalpost. When this condition applies, the statement that, "The UHI equal to .6 is twice as healthy or unhealthy (depending on the direction of the indicators) as an area with a UHI equal to .3."⁵² When the UHI does not have ratio scale properties, such statements are not appropriate.

Avoid Extrapolating beyond the Indicators

As noted in Selecting Indicators for the Urban Health Index, the UHI is defined by the indicators that are chosen for its construction. There is a risk of generalizing beyond those indicators. This risk is perhaps more so among those who did not actually construct the UHI and thus less familiar with why certain indicators were chosen or omitted for its construction. To guard against overgeneralization, reports based on the UHI should be explicit about the indicators chosen and those omitted, as well as the domains. Qualifying interpretations and acknowledging that a different set of indicators could have generated different interpretations may avoid this pitfall.

51 In some cases, an alternate lower goalpost choice might result in a standardized indicator that could be considered on a ratio scale if interpretation shifted to a different indicator. For example, consider life expectancy, a ratio scale measure, as an indicator for the Human Development Index (HDI). The developers of the HDI standardized this variable using a lower goalpost of 20 years, which was considered a subsistence value. A transformation of a ratio scale variable that subtracts or adds a constant (such as a non-zero lower goal post value) would result in a variable that no longer has ratio scale properties. However, for the standardized life expectancy variable, a zero represents a subsistence level, and thus could still be considered a meaningful zero value if interpretation shifted from life expectancy to a life expectancy subsistence metric.

52 Of course, "healthy" and "unhealthy" are defined by indicators chosen to comprise the index.

INTERACTIONS WITH PLANNERS AND DECISION MAKERS —USING THE URBAN HEALTH INDEX TO SET PRIORITIES AND GOALS FOR POLICY CHANGE

People in positions to influence policies have competing priorities. The UHI analyst may be disappointed by the reception of the data and recommendations. It is important to present information on the status of urban dwellers with appropriate goals and expectations. The process of social change, in particular the betterment of persons who are less affluent, is a tortuous one, and a long-range view is needed. In fact, adopting a long-term approach may be beneficial in presenting recommendations for the amelioration of urban health and urban health disparities.

This index, and others like it, is of little use if it does not provide decision makers with solid justification for the policies they may pursue. The UHI has the advantage of built-in recognition of the multifactorial nature of health disparities. Since it draws on a variety of rubrics, domains, and indicators, the UHI can point decision makers to broad-based overlapping social ills. They can, in turn, respond with policies that are comprehensive and inclusive. Small steps are certainly warranted, but an approach that includes education, health care access, unemployment, built environment, clean water, and clean air, is likely to have far more impact. The UHI has an intrinsic structure that permits mixing elements from different domains, and it can support a broad vision and justify policies of public support for social goods.

Convincing planners and decision makers can be fraught with frustration. Support for health and education, for example, requires public investment, often looked upon as an income transfer, a term that is anathema to those more affluent, and, thus, a political danger to decision makers. Yet, despite its solid justification, measures like the UHI may find its utility limited, since inequalities and disparities have built-in mechanisms for their own persistence. The spiral of increasing disparity and increasing resistance to change is important to bear in mind when

A question likely to be raised is how can or should policy makers use the UHI to set goals or targets, and, once targets are set, how can progress be measured. Targets can be set for relative or absolute improvements in UHI values for particular units or areas within a city (e.g., zip code S will experience X% improvement in its UHI value over the next Y years).

Alternatively, targets can be set for the whole city. These targets could reference the average or median UHI, the UHI disparities ratio, difference, or slope (e.g., the UHI Ratio will decrease by X amount over the next Y years).

Although such targets and goals are often arbitrary, they can be useful for directing resources and attention to an issue. Goals should have wide consensus, be well-defined and be given a concrete, realistic timeline for achievement. Benchmarks for progress should be identified by local stakeholders. These benchmarks can be from baseline values or comparative data from neighboring areas.

setting goals for policy change⁵³, and in addressing the concerns of policy makers. There are several guiding principles that may be of use:

1. Amelioration of disparities should be viewed as having a positive effect on society as a whole, not just a matter for “helping the poor.” The argument for social justice is a critical one, but it may not resonate with those who deal with day-to-day politics or those who take a pejorative view of the poor. Rather, it may be more effective to stress that a healthy and educated body politic is the basis for an innovative society and a competent workforce. These positive characteristics can further the economic goals of an entire urban center. An industrial and entrepreneurial enterprise cannot flourish in an environment of extreme disparity, and real economic gains will not occur with a large impoverished society. Thus, a case can be made that those currently affluent will do even better with a less disparate society.
2. A long-range view is required. The UHI presents a current cross-section and would thus seem to be a victim of short-term thinking. Use of historical data, to the extent available, and amalgamating data from various sources, should be a goal of the UHI. The context and perspective that a long term view provides is critical for planning. Admittedly, many political systems are not amenable to long-term planning. This is one of the major obstacles analysts and decision makers themselves face.
3. Stress to the decision makers that the opposite of inequality is not equality. The goal is not to create a homogeneous society, but rather to reduce disparities among the population. Raising and flattening the disparity slope places more persons in a position of opportunity and permits society to make greater use of its human capital. Having persons at the upper end of such a distribution is consistent with social development.
4. Seek the broadest possible base for support. People who can actually effect change are often in the upper echelons of an organization. The analyst should seek the support of those who work in the appropriate agencies by attempting to build relationships with those in charge and develop a base of consensus on the findings and the actions that need to be taken. It cannot be emphasized enough that success in UHI presentation is often determined by pre-established personal relationships.
5. The best presentation strategy is understatement. It is probably most effective for an audience to discover the message in what you say, rather than to overemphasize the message. The words of the Swiss developmental psychologist Jean Piaget are relevant: “Children have real understanding only of that which they invent themselves, and each time that we try to teach them something too quickly, we keep them from reinventing it themselves.”

53 When targets are linked to the UHI ratio or difference, be cognizant that “improvements” signaled by a ratio decrease could be obtained by *either* improvements in health for the least healthy areas or declines in health by the most healthy areas. The former situation being more desirable than the latter. Another consideration when setting targets is that the UHI is a *relative* index. If the lower and upper goalposts are pegged to the areas that mark the minimum or maximum (see Standardization of the Indicators), then changes in indicator values for those areas will change the UHI values for all other areas, even when none of the other areas experience any change in their indicator values. This is because the goalposts have changed. This is not inherently problematic, but it must be recognized when setting targets and tracking their progress.

APPENDIX

Appendix A: Guide to an Excel-based Tool (v. 1) for Calculation and Visualization of the Urban Health Index

A Microsoft Excel-based tool has been developed to aid the analyst in calculating the Urban Health Index (UHI)⁵⁴. This appendix will describe the functions and capabilities of this tool and provide guidance on its use.

Tools and Capabilities

The Excel UHI tool provides several options for constructing the UHI. Current and planned user tools and options include:

- A dedicated spreadsheet where the analyst can enter data to be used in constructing the UHI.
- A dialog box for selecting indicators and a geographic identifier variable.
- An automatic calculation of statistics for selected indicators.
- An automatic identification of whether missing data (represented by blank cells) is present for the selected indicators.⁵⁵
- The option to change the default settings for indicator weights (see Amalgamation of the Standardized Indicators and Appendix B).
- The option to change default settings for upper and lower goalpost (see Standardization of the Indicators).
- The option to choose the percentile thresholds for calculating the health disparities ratio, difference, and gradient.
- An automatic calculation of the UHI, UHI ranking, and UHI percentiles for each area.
- An automatic calculation of UHI summary statistics, including the UHI disparities ratio, disparities difference, and disparities gradient (see Calculating Summary Statistics for the UHI).⁵⁶
- An automatic generation of the UHI Index Plot.
- An automatic creation of a worksheet containing the selected indicators and geographic identifier, calculated UHI, UHI ranking, and UHI percentiles for easy export to another program, such as a Geographic Information Systems (GIS) program.
- An option to clear the results.

The output of the current version of the Excel UHI tool, provides:

- The number of selected indicators and observations.
- Default settings for upper and lower goalpost values used to standardize indicators and the indicator weights that will be used when calculating the UHI (with an option to change these default settings).
- Summary statistics for each indicator, including the mean, standard deviation, minimum and maximum values, 10th percentile, 50th percentile or median, 90th percentile, and bivariate correlations.

54 Both this handbook and the Excel-based calculation tool are available for download from the WHO Centre for Health Development's website at http://www.who.int/kobe_centre/publications/urban_health_index_toolkit/en/ (accessed 16 October 2014).

55 Identification of missing data will only occur if a data cell is empty (blank). See Preparing the Data. If there are blank cells for one or more selected indicators, the analyst will be alerted, and the UHI will be calculated only for areas with non-blank cells.

56 The analyst will be prompted to enter a percentage that defines the upper and lower extremes.

- Summary statistics for the UHI, including mean, standard deviation, minimum and maximum, and select percentiles.
- UHI disparity and inequality statistics, including the UHI health disparities ratio and difference statistics, along with the mean UHI for the best and worst quantiles, and the UHI disparities gradient.
- The UHI plotted.

Instructions for Using the Tool

1. Open the Tool with Microsoft Excel

In order to use the tool, the analyst will need to have a licensed copy of the Microsoft Excel spreadsheet software. The tool was developed and tested on Microsoft Excel 2010 and 2013 using Microsoft Windows 7 and 8. It is not currently known whether the functions of the tool are all fully compatible with older versions of Microsoft Excel, particularly versions older than Microsoft Excel 2007 or versions designed for other operating systems.

After saving the tool as an .xlsm file to a computer or network drive and opening it in Microsoft Excel, the user must enable macros in order for the tool to function properly. If macros are disabled by default, the user should see a Security Warning display and an option to Enable Content. Select the option by clicking on the button. Any other prompts to enable editing of the contents of the tool should be selected. After enabling the macros, three worksheet tabs will appear at the bottom: **Instructions**, **Data**, and **Options**. The **Instructions** sheet provides an outline of the instructions provided here. Indicator data are entered on the **Data** sheet (see next step). The **Options** sheet shows what indicators can be selected and computed statistics for the selected indicators. It is also where default settings for indicator weights and goalposts can be viewed and changed, and is where calculation of the UHI is initiated.

2. Enter Data

To enter data, click on the **Data** sheet. The sheet should be empty if it is the first time this tool is accessed. Enter the data for the indicators. Data can be either typed manually or copied and pasted from another spreadsheet file or program. The data should be arranged as shown in Figure 3. Rows define the geographic areas, and columns define the indicators and other variables, such as an area identifier. The top row (row 1) should be reserved for the variable names. The name for the first variable will begin with cell A1 and data for the first area will begin in cell A2.

3. Select Indicators

To select indicators for the UHI, click on the **Options** sheet. Then, click the button Select Data. A dialog box will open and display where indicators can be selected. Select from the list of variable names obtained from Row 1 of the **Data** sheet by selecting and holding the Ctrl or Shift keys to select multiple indicators. Another option will also allow selection of an area or location identifier variable. This is recommended so the UHI values computed can be linked to their respective areas. Once all indicators are selected, click the Select Data button. This action will create several tables (shown on the **Options** sheet) and the **Data for Export** sheet (described below).⁵⁷

The number of selected indicators and observations without missing data are displayed in cells B2 and B3 on the **Options** page. It is a good idea to check these values to confirm whether or not they are as expected. *Table 2* contains the summary statistics for the selected indicators: mean,

⁵⁷ If any selected indicators has missing data, represented by one or more empty or blank cells, a warning will display. Blank cells are highlighted on the Data page. The UHI will be calculated only for areas that have occupied cells for the selected indicators after Continue is selected. Cancel, which cancels the operation, can also be selected.

standard deviation, minimum and maximum, 10th percentile, 90th percentile, and the median (50th percentile). *Table 3* contains the correlation matrix for the selected indicators. This table will provide information that will be helpful in determining whether indicators are well-suited for aggregation. For example, a strong negative correlation of an indicator with other indicators may suggest that the indicator needs to be reversed, or perhaps there is a complicated relationship with the other indicators and therefore should be omitted from the index. High correlations among subsets of indicators may suggest that indicators within such a subset are redundant or could be grouped together. This observation may alert the analyst to information on the selection of weights. *Table 1* contains the default settings for lower and upper goalpost values. These will be used in standardizing the selected indicators (see Standardization of the Indicators). In addition, the table contains the default weights that will be applied to the indicators when computing the UHI (see Amalgamation of the Standardized Indicators). The default upper goalpost value for any indicator is the maximum for that indicator. The default lower goalpost value is the minimum value, minus a small constant, for that indicator. The default allows each standardized indicator to be given equal (specifically, unit) weights when calculating the UHI. If alternative goalpost values and/or weights are desired, enter the desired settings in *Table 1*. However, if a goalpost value is manually changed, a lower goalpost value that is less than the minimum for that indicator and an upper goalpost value that is equal to or greater than the maximum for that indicator must be chosen. This is necessary for the standardized indicators to be bounded by 0 and 1. Not following this requirement will result in either an error message or an improper standardization.

4. Calculate the Urban Health Index

After verifying the selected indicators and desired settings for the goalposts and weights, obtain the UHI values for each area by clicking the Calculate UHI button on the **Options** sheet. A prompt is given to define the proportion of cases at the extremes of the UHI distribution that will be included in the extreme groups for calculation of the UHI health disparities ratio and difference, and for the UHI health disparities gradient, or slope. The default is .2. This default defines 20% of the areas as extreme areas, or the bottom and top 10%. Keep the default value or enter a new value greater than 0 but less than 1 and click OK. This will lead to the generation of two new sheets: **UHI Statistics** and **UHI Index Plot**. The **UHI Statistics** sheet will contain the UHI, UHI ranking, and UHI percentile for each area, as well as the summary statistics for the UHI and UHI disparity statistics. If a geographic identifier variable is selected, it will also be displayed. The UHI summary statistics include the mean, standard deviation, minimum and maximum, 10th percentile, median, and 90th percentile. The UHI disparity statistics are the UHI health disparities ratio, the health disparities difference, and the slope or gradient (see Calculating Summary Statistics for the UHI). The **UHI Index Plot** sheet contains an index plot of the UHI (UHI plotted against its rank, in ascending fashion) (see Creating an Index Plot of the UHI). The index plot can be formatted and copied and pasted into another document, such as a report.

5. Export Data

If the UHI data are exported to another program, such as a GIS program (see Appendix D), or merged with other data, all data can be copied and pasted into the other program or can be directly imported or accessed from this sheet. To facilitate exporting of the UHI data, the **Data for Export** sheet contains the selected indicators, calculated UHI variables, and the geographic or area identifier variable (if selected).

6. Clear the Results

To reset or clear the results, click the Clear results button on the **Options** sheet. This will delete the tables in the **Options** sheet, as well as the other sheets created (**UHI Statistics**, **Data for Export** and **UHI Index Plot**). It will not delete the **Data** sheet or any data entered manually.

Appendix B: Guidance for a Weighted Urban Health Index

Consider the formulae for the Urban Health Index provided earlier,

$$UHI = \left(\prod_{i=1}^n I_i^S \right)^{\frac{1}{j}}$$

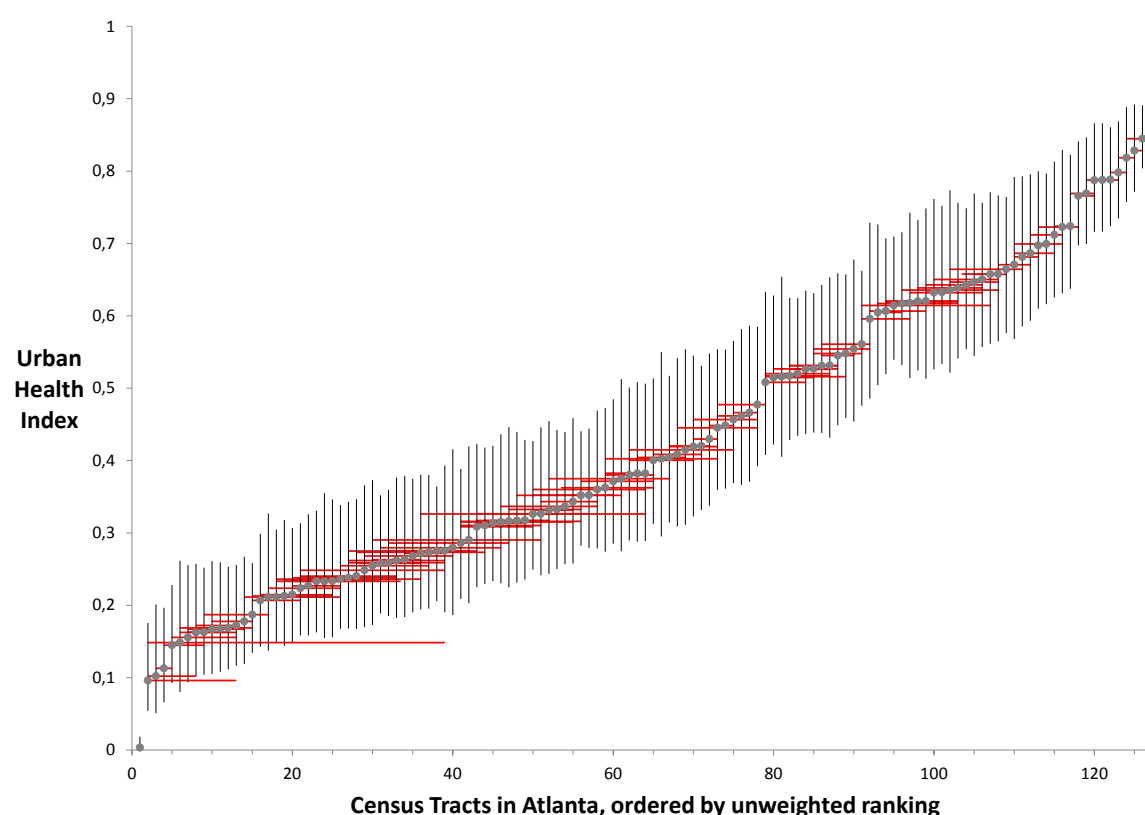
This formula gives equal weight to each indicator so that each indicator contributes equally to the UHI. However, in some contexts, some indicators may be given greater weight or other indicators given less weight in computing the UHI. For example, perhaps some health determinant indicators are known to have a stronger relationship to health outcomes than other health determinant indicators. In this situation, the UHI could reflect this by giving more weight to the stronger predictors.

Consider a situation when indicators fall into different domains and the desire is to give each domain equal weight or a weight that is not proportional to the number of indicators. For example, consider a UHI for health determinants that includes two education indicators (percent graduating from secondary school and percent graduating with a post-secondary degree), two economic indicators (median household income and percent below poverty level), and just a single environmental indicator. If using the aforementioned equation for the geometric mean, each indicator would be given equal weight, but the result is most driven by the economic and education domains, where each is represented by two of the five indicators and is least influenced by the environmental domain (20%), which contributes only a single indicator. If the desire was to weight each domain equally, so that, for example, the two education indicators collectively carried as much weight as the environmental indicator, then a weighted geometric mean can be calculated:

$$UHI = \left(\prod_{i=1}^j S I_i^{w_i} \right)^{1/\sum_{i=1}^j w_i}$$

where w_i are the weights for the indicators. When all weights are equal to 1 (referred to as unit weighted or unweighted), then this equation reduces to the previous equation for the geometric mean. To calculate a weighted UHI, first, the standardized indicators are raised to a power equal to their respective weights. Then, multiply these values. Finally, this product is raised to the $1/\sum_{i=1}^j w_i$ power. When the weights are scaled to the sum of the number of indicators, $1/\sum_{i=1}^j w_i$ simply becomes $1/j$. Consider further that weighting the economic and education indicators *each* .5 and the environmental indicator 1 would result in an index with equal weights for each domain. Ultimately, the decision to weight or not should be determined by the desired emphasis and interpretation of the index and as to whether the weighting scheme is defensible on those grounds. We have conducted a simulated weight analysis on a fixed set of indicators. The results suggested that the UHI and rank positions remained fairly stable across a broad range of weights, as shown in Figure 11.

Figure 11: Results of the Weighting Simulations



Note: Vertical bars represent the 10th and 90th percentiles of the distribution of UHIs for each census tract. Horizontal bars depict the range of ranks for each census tract.

For at least a few units, the ordered ranks and UHI did vary. It is important to choose weights that are defensible and have consensus. In absence of a justifiable reason for differential weighting, it is advisable to default to an unweighted UHI. The Excel-based UHI tool will default to such an analysis; however, desired weights can be manually entered (see Select Indicators), which can automatically be scaled to the sum of the number of indicators. Using the 2000 census data from towns in Shanghai, China, Figure 12 illustrates the calculation of the weighted UHI. In this illustration, there are five indicators: percent with senior middle school graduation, percent employed in a professional occupation, percent employed, percent of households with gas/electric service, and percent of households with own lavatory. These indicators represent three hypothetical domains: education, employment, and household/sanitation. If the desire was to equally weight the domains, weights of 1, .5, .5, .5, .5, respectively could be applied. This scheme would give employment or household/sanitation indicators half the influence (weight) as given to the education indicators when calculating the UHI.⁵⁸

58 The Excel function used to calculate the weighted UHI value in cell M3 of the table shown in Figure 12 is =POWER(PRODUCT(POWER(G3,1),POWER(H3,.5),POWER(I3,.5),POWER(J3,.5),POWER(K3,.5)),1/3).

Figure 12: Illustrated Calculation of the Weighted Geometric Mean

	A	G	H	I	J	K	L	M
1		Education	Employment		Household/Sanitation			
2	GBTOWN	I_{educat}^S	$I_{prof_empl}^S$	I_{employ}^S	$I_{gas/elect}^S$	$I_{lavatory}^S$	$\left(\prod_{i=1}^j I_i^{w_i} \right)^{1/\sum_{i=1}^j w_i} =$	Weighted UHI
3	310101001	0.754	0.441	0.380	0.987	0.356	$\left(.754^{.5} *.441^{.5} *...*.356^{.5} \right)^{\frac{1}{.5}}$	0.568
330	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
331	310230458	0.003	0.018	1.000	0.005	1.000	$\left(.003^{.5} *.018^{.5} *...*1.00^{.5} \right)^{\frac{1}{.5}}$	0.029
332	310230459	0.011	0.002	1.000	0.806	1.000	$\left(.011^{.5} *.002^{.5} *...*1.00^{.5} \right)^{\frac{1}{.5}}$	0.076

Appendix C: Derivation of the Variance for the Urban Health Index⁵⁹

Derivation of the Variance for UHI: Without Assuming Normality

The derivation of the variance for UHI is based on the following equations where the moments of a smooth function f of a random variable X are approximated by using the Taylor expansion. Let μ_X and σ_X^2 be the mean and variance for X , respectively. Using the second-degree Taylor approximation:

$$E[f(X)] \approx f(\mu_X) + \frac{1}{2} f''(\mu_X) \sigma_X^2, \quad (0.1)$$

$$Var(f(X)) \approx [f'(\mu_X)]^2 \sigma_X^2. \quad (0.2)$$

Let Y be another random variable with mean μ_Y and variance σ_Y^2 . Based on the Taylor approximation:

$$Cov(f(X), f(Y)) \approx f'(\mu_X) f'(\mu_Y) Cov(X, Y) \quad (0.3)$$

To get the variance for UHI, we write the UHI in an equivalent form:

$$UHI = \exp\left(\frac{1}{j} \sum_{i=1}^j \ln(I_i^S)\right) \equiv \exp(Z),$$

Where

$$Z = \ln(UHI) = \frac{1}{j} \sum_{i=1}^j \ln(I_i^S)$$

$$\text{Let } \mu_i = E[I_i^S] = \frac{E[I_i] - \min^*(I)}{\max(I) - \min^*(I)} \quad \text{and} \quad \sigma_i^2 = Var[I_i^S] = \frac{Var[I_i]}{[\max(I) - \min^*(I)]^2}$$

⁵⁹ This appendix provides the derivation of the variance estimator for the UHI when the indicators are unit-weighted (which is the method described in this Handbook). When non-unit weights are used, modifications are necessary. It is also assumed that lower and upper goalposts, $\max(I)$ and $\min^*(I)$, are constants (not random).

denote the mean and variance of the standardized indicator I_i^S , respectively.

Let

$$\sigma_{ii'} = \text{Cov}(I_i^S, I_{i'}^S) = \frac{\text{Cov}(I_i, I_{i'})}{[\max(I) - \min^*(I)]^2}$$

be the covariance between the standardized indicators I_i^S and $I_{i'}^S$.
Using equations (0.1), (0.2) and (0.3) with $f(x) = \ln(x)$:

$$E[\ln(I_i^S)] \approx \ln(\mu_i) - \frac{\sigma_i^2}{2\mu_i^2},$$

$$\text{Var}[\ln(I_i^S)] \approx \frac{\sigma_i^2}{\mu_i^2},$$

$$\text{Cov}(\ln(I_i^S), \ln(I_{i'}^S)) \approx \frac{\sigma_{ii'}}{\mu_i \mu_{i'}}.$$

Then

$$E[Z] = E[\ln(UHI)] = E\left[\frac{1}{j} \sum_{i=1}^j \ln(I_i^S)\right] \approx \frac{1}{j} \sum_{i=1}^j \left(\ln(\mu_i) - \frac{\sigma_i^2}{2\mu_i^2} \right),$$

$$\text{Var}[Z] = \text{Var}[\ln(UHI)] = \text{Var}\left[\frac{1}{j} \sum_{i=1}^j \ln(I_i^S)\right] \approx \frac{1}{j^2} \left\{ \sum_{i=1}^j \left(\frac{\sigma_i^2}{\mu_i^2} \right) + 2 \sum_{i < i'} \frac{\sigma_{ii'}}{\mu_i \mu_{i'}} \right\}.$$

Applying the equations (0.1), (0.2) and (0.3) with $f(x) = \exp(x)$:

$$E[UHI] = E[\exp(Z)] \approx \exp(E[Z]) + \frac{1}{2} \exp(E[Z]) \text{Var}[Z]$$

$$\approx \left(\prod_{i=1}^j \mu_i \right)^{1/j} \exp\left(-\frac{1}{j} \sum_{i=1}^j \left(\frac{\sigma_i^2}{2\mu_i^2} \right)\right) \left[1 + \frac{\text{Var}[Z]}{2} \right],$$

and

$$\text{Var}[UHI] = \text{Var}[\exp(Z)] \approx \exp(2E[Z]) \text{Var}[Z]$$

$$\approx \left(\prod_{i=1}^j \mu_i \right)^{2/j} \exp\left(-\frac{1}{j} \sum_{i=1}^j \left(\frac{\sigma_i^2}{\mu_i^2} \right)\right) \text{Var}[Z].$$

Derivation of the Variance for UHI: With Assumed Normality

When indicators having a joint normal distribution is assumed, a higher precision in the approximation of variance for UHI can be achieved. If X has a normal distribution with mean μ_X and variance σ_X^2 , the equation is $E[(X - \mu_X)^3] = 0$ and $E[(X - \mu_X)^4] = 3\sigma_X^4$. For a function f that is sufficiently differentiable, based on the fourth degree Taylor approximation:

$$E[f(X)] \approx E\left[f(\mu_x) + f'(\mu_x)(X - \mu_x) + \frac{1}{2} f''(\mu_x)(X - \mu_x)^2 + \frac{1}{6} f'''(\mu_x)(X - \mu_x)^3 + \frac{1}{24} f^{(4)}(\mu_x)(X - \mu_x)^4 \right]$$

$$= f(\mu_X) + \frac{1}{2} f''(\mu_X) \sigma_X^2 + \frac{1}{8} f^{(4)}(\mu_X) \sigma_X^4, \quad (0.4)$$

$$\begin{aligned} E[f(X)^2] &\approx E\left[\left\{f(\mu_X) + f'(\mu_X)(X - \mu_X) + \frac{1}{2} f''(\mu_X)(X - \mu_X)^2 + \frac{1}{6} f'''(\mu_X)(X - \mu_X)^3 + \frac{1}{24} f^{(4)}(\mu_X)(X - \mu_X)^4\right\}^2\right] \\ &\approx f(\mu_X)^2 + \{[f'(\mu_X)]^2 + f(\mu_X) f''(\mu_X)\} \sigma_X^2 + 3\left\{\frac{1}{4} [f''(\mu_X)]^2 + \frac{1}{3} f'(\mu_X) f'''(\mu_X) + \frac{1}{12} f(\mu_X) f^{(4)}(\mu_X)\right\} \sigma_X^4, \end{aligned}$$

So

$$\begin{aligned} Var[f(X)] &= E[f(X)^2] - \{E[f(X)]\}^2 \\ &\approx [f'(\mu_X)]^2 \sigma_X^2 + \left\{\frac{1}{2} [f''(\mu_X)]^2 + f'(\mu_X) f'''(\mu_X)\right\} \sigma_X^4. \end{aligned} \quad (0.5)$$

Let Y be another normal random variable with mean μ_Y and variance σ_Y^2 . Assuming that the joint distribution of X and Y is normal, then

$$\begin{aligned} E[(X - \mu_X)(Y - \mu_Y)^2] &= E\{E[(X - \mu_X)(Y - \mu_Y)^2 | Y]\} = E\{E[(X - \mu_X) | Y](Y - \mu_Y)^2\} \\ &= E\{[Cov(X, Y) \sigma_Y^{-2} (Y - \mu_Y)](Y - \mu_Y)^2\} = 0, \\ E[(X - \mu_X)(Y - \mu_Y)^3] &= E\{[Cov(X, Y) \sigma_Y^{-2} (Y - \mu_Y)](Y - \mu_Y)^3\} \\ &= Cov(X, Y) \sigma_Y^{-2} E[(Y - \mu_Y)^4] = 3Cov(X, Y) \sigma_Y^2. \end{aligned}$$

Let

$$\mu_{X|Y} = E(X | Y) = \mu_X + Cov(X, Y) \sigma_Y^{-2} (Y - \mu_Y), \text{ in addition:}$$

$$\begin{aligned} E[(X - \mu_X)^2 (Y - \mu_Y)^2] &= E\{E[(X - \mu_X)^2 (Y - \mu_Y)^2 | Y]\} = E\{E[(X - \mu_X)^2 | Y](Y - \mu_Y)^2\} \\ &= E\{E[(X - \mu_{X|Y})^2 + 2(\mu_{X|Y} - \mu_X)(X - \mu_{X|Y}) + (\mu_{X|Y} - \mu_X)^2 | Y](Y - \mu_Y)^2\} \\ &= E\{[\sigma_X^2 - [Cov(X, Y)]^2 \sigma_Y^{-2} + [Cov(X, Y)]^2 \sigma_Y^{-4} (Y - \mu_Y)^2](Y - \mu_Y)^2\} \\ &= [\sigma_X^2 - [Cov(X, Y)]^2 \sigma_Y^{-2}] \sigma_Y^2 + [Cov(X, Y)]^2 \sigma_Y^{-4} 3\sigma_Y^4 \\ &= \sigma_X^2 \sigma_Y^2 + 2Cov(X, Y). \end{aligned}$$

Using these results and the fourth degree Taylor approximation:

$$\begin{aligned} E[f(X)f(Y)] &\approx E\left[\left(\sum_{i=0}^4 \frac{1}{i!} f^{(i)}(\mu_X)(X - \mu_X)^i\right)\left(\sum_{i=0}^4 \frac{1}{i!} f^{(i)}(\mu_Y)(Y - \mu_Y)^i\right)\right] \\ &\approx E[f(\mu_X)f(\mu_Y) + \frac{1}{2} f(\mu_X)f''(\mu_Y)(Y - \mu_Y)^2 + \frac{1}{24} f(\mu_X)f^{(4)}(\mu_Y)(Y - \mu_Y)^4] \end{aligned}$$

$$\begin{aligned}
 & +f'(\mu_X)f'(\mu_Y)(X-\mu_X)(Y-\mu_Y)+\frac{1}{6}f'(\mu_X)f'''(\mu_Y)(X-\mu_X)(Y-\mu_Y)^3 \\
 & +\frac{1}{2}f''(\mu_X)f(\mu_Y)(X-\mu_X)^2+\frac{1}{4}f''(\mu_X)f''(\mu_Y)(X-\mu_X)^2(Y-\mu_Y)^2 \\
 & +\frac{1}{6}f'''(\mu_X)f'(\mu_Y)(X-\mu_X)^3(Y-\mu_Y)+\frac{1}{24}f^{(4)}(\mu_X)f(\mu_Y)(X-\mu_X)^4] \\
 & =f(\mu_X)f(\mu_Y)+\frac{1}{2}f(\mu_X)f''(\mu_Y)\sigma_Y^2+\frac{1}{8}f(\mu_X)f^{(4)}(\mu_Y)\sigma_Y^4+f'(\mu_X)f'(\mu_Y)Cov(X,Y) \\
 & +\frac{1}{2}f'(\mu_X)f'''(\mu_Y)Cov(X,Y)\sigma_Y^2+\frac{1}{2}f''(\mu_X)f(\mu_Y)\sigma_X^2+\frac{1}{4}f''(\mu_X)f''(\mu_Y)[\sigma_X^2\sigma_Y^2+2Cov(X,Y)] \\
 & +\frac{1}{2}f'''(\mu_X)f'(\mu_Y)Cov(X,Y)\sigma_X^2+\frac{1}{8}f^{(4)}(\mu_X)f(\mu_Y)\sigma_X^4
 \end{aligned}$$

Then $Cov(f(X), f(Y)) = E[f(X)f(Y)] - E[f(X)]E[f(Y)]$

$$\approx \left[f'(\mu_X)f'(\mu_Y) + \frac{1}{2}f'(\mu_X)f'''(\mu_Y)\sigma_Y^2 + \frac{1}{2}f''(\mu_X)f''(\mu_Y)Cov(X,Y) + \frac{1}{2}f'''(\mu_X)f'(\mu_Y)\sigma_X^2 \right] Cov(X,Y). \quad (0.6)$$

Using the same notation applied to the case without normal assumption, using equations (0.4), (0.5) and (0.6) with

$$f(x) = \ln(x), \quad f''(x) = -\frac{1}{x^2} \quad \text{and} \quad f^{(4)}(x) = -\frac{6}{x^4} :$$

$$E[\ln(I_i^S)] = \ln(\mu_i) - \frac{\sigma_i^2}{2\mu_i^2} - \frac{3\sigma_i^4}{4\mu_i^4},$$

$$Var(\ln(I_i^S)) = \frac{\sigma_i^2}{\mu_i^2} + \frac{5\sigma_i^4}{2\mu_i^4},$$

$$Cov(\ln(I_i^S), \ln(I_{i'}^S)) = \frac{\sigma_{ii'}}{\mu_i\mu_{i'}} \left(1 + \frac{\sigma_i^2}{\mu_i^2} + \frac{\sigma_{ii'}}{2\mu_i\mu_{i'}} + \frac{\sigma_{i'}^2}{\mu_{i'}^2} \right).$$

Then

$$E[Z] = E[\ln(UHI)] = E\left[\frac{1}{j} \sum_{i=1}^j \ln(I_i^S) \right] = \frac{1}{j} \sum_{i=1}^j \left(\ln(\mu_i) - \frac{\sigma_i^2}{2\mu_i^2} - \frac{3\sigma_i^4}{4\mu_i^4} \right),$$

$$Var[Z] = Var[\ln(UHI)] = Var\left[\frac{1}{j} \sum_{i=1}^j \ln(I_i^S) \right]$$

$$= \frac{1}{j^2} \left\{ \sum_{i=1}^j \left(\frac{\sigma_i^2}{\mu_i^2} + \frac{5\sigma_i^4}{2\mu_i^4} \right) + 2 \sum_{i < i'} \frac{\sigma_{ii'}}{\mu_i\mu_{i'}} \left(1 + \frac{\sigma_i^2}{\mu_i^2} + \frac{\sigma_{ii'}}{2\mu_i\mu_{i'}} + \frac{\sigma_{i'}^2}{\mu_{i'}^2} \right) \right\}.$$

Applying the equations (0.4), (0.5) and (0.6) with $f(x) = \exp(x)$:

$$E[UHI] = \left(\prod_i \mu_i \right)^{1/j} \exp \left(-\frac{1}{j} \sum_{i=1}^j \left(\frac{\sigma_i^2}{2\mu_i^2} + \frac{3\sigma_i^4}{4\mu_i^4} \right) \right) \left[1 + \frac{Var[Z]}{2} \right],$$

and

$$Var(UHI) = \left(\prod_i \mu_i \right)^{2/j} \exp \left(-\frac{2}{j} \sum_{i=1}^j \left(\frac{\sigma_i^2}{2\mu_i^2} + \frac{3\sigma_i^4}{4\mu_i^4} \right) \right) Var[Z].$$

Appendix D: Guidance and Illustration of Mapping the Urban Health Index with QGIS and ArcGIS

Using ArcGIS for Mapping the UHI

Mapping the UHI requires a Geographic Information System (GIS) program and geospatial vector data of the boundaries of geopolitical or geo-statistical units used in the calculation of the UHI. A commonly used GIS program is ArcGIS for Desktop, developed by Environmental Systems Research Institute (ESRI, Redlands, CA). Currently it works only with Windows systems. The common geospatial vector data format is shapefile, developed and regulated by ESRI, and is mostly interoperable among ESRI and other GIS programs.



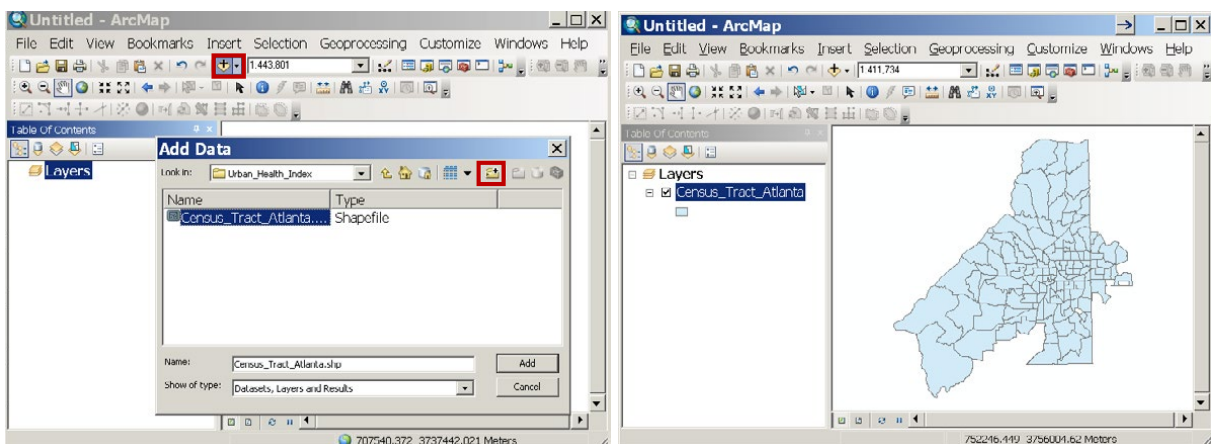
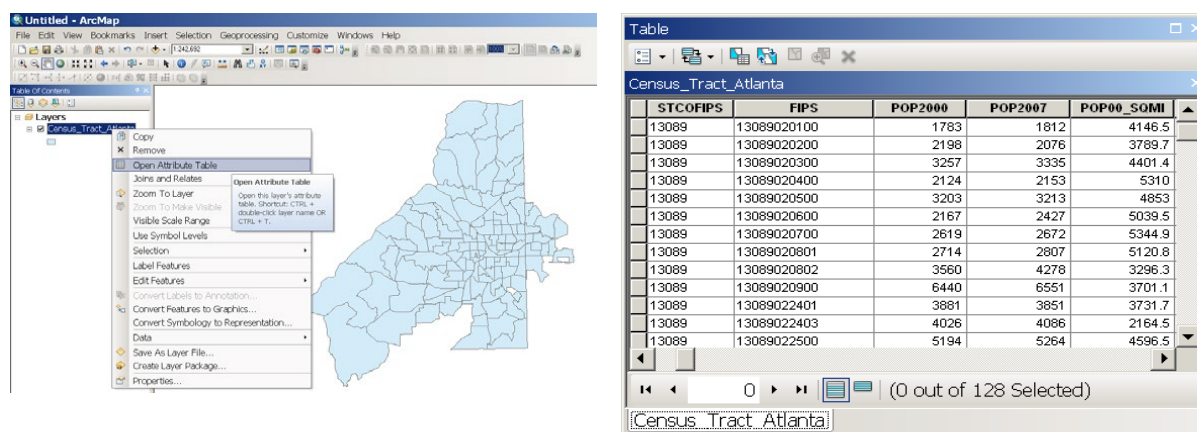
1. An initial step of mapping the UHI is to ensure that the GIS shapefile data and the UHI data are ready to use. It is important to check the attribute of the GIS shapefile when there is a field storing the geolocation ID. This ID field will be used as the “key” to append the UHI result calculated in the Excel toolkit. The ID may be a census tract ID, ZIP code, or township name, but each shall be unique without any redundancy in this field. To check the redundancy, start ArcMap in the ArcGIS package, and click **Add**  to add the shapefile to ArcMap (see Figure 13). In the **Add Data** window, an icon  may be used to navigate to the folder where the unit boundary shapefile is located. Once the shapefile is identified, click **Add** and the map will be shown in ArcMap.

Figure 13: The User Interface of ArcMap to add a Shapefile (left) and the Resulting Screen After the Shapefile is Added (right)



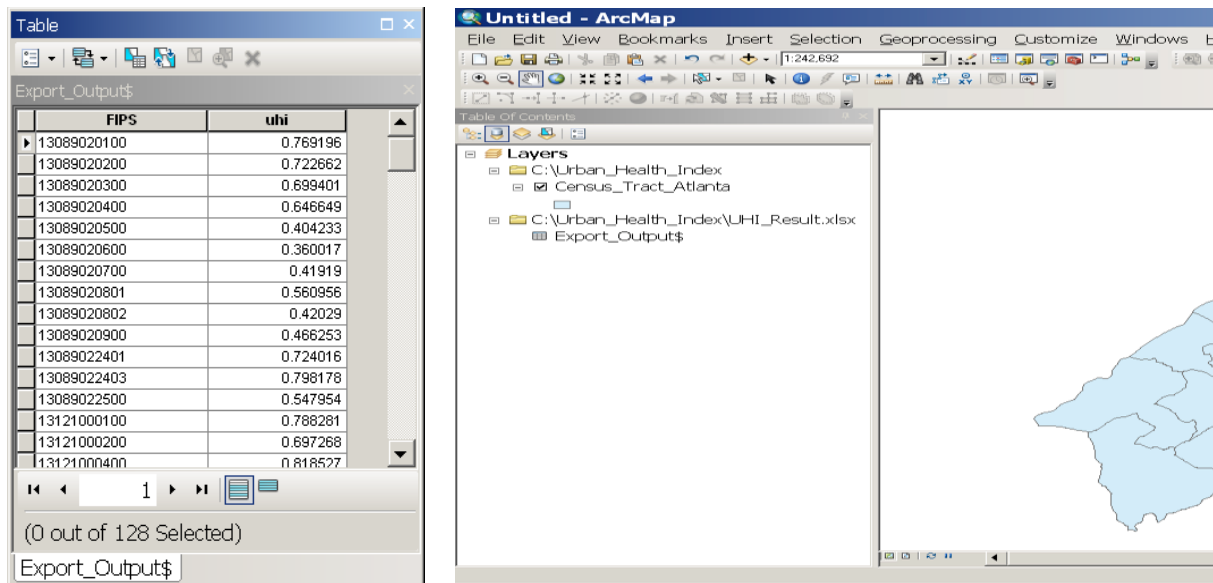
To check the field of geolocation ID, right-click the shapefile in **Table Of Contents** of ArcMap (Figure 14 left), and left-click **Open Attribute Table**. An attribute table similar to an Excel spreadsheet will pop up. In Figure 14, the FIPS field (geolocation ID) is unique but the STCOFIPS field is not (Figure 14 right).

Figure 14: The User Interface to Open an Attribute Table (left) and the Resulting Attribute Table (right)



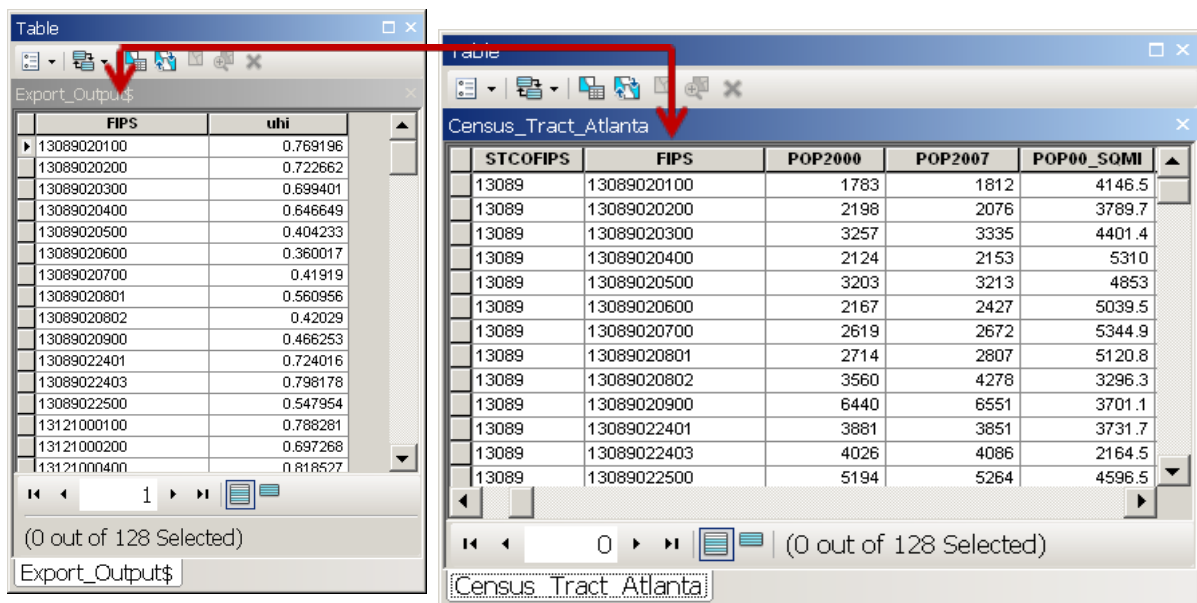
- The second step will be to prepare the UHI result in ArcMap. Ensure the UHI result in Excel has the fields of geolocation ID and the UHI result. The field name, or column title, should be 10 characters or less and contain only a letter, number, or underscore with no other non-letter characters such as !, @ # or space. Because ArcMap has a maximum field name length of 10 characters, field names in Excel longer than 10 characters will be truncated in ArcMap later on. It is highly recommended that the UHI spreadsheet to be added into ArcMap has only two columns: the field of geolocation ID and the UHI. Data must begin at the most upper-left cell with no graphs or charts in the spreadsheet.
- The third step will be to add the UHI result in ArcMap and append to the geopolitical or geo-statistical boundary shapefile. Close the UHI result Excel file and go to ArcMap to repeat the Add Data process presented earlier in order to add the UHI Excel spreadsheet to ArcMap. Because an Excel file may have multiple sheets, designate which sheet to import into ArcMap. After the UHI spreadsheet is added into ArcMap, open the attribute table by right-clicking the UHI spreadsheet in the **Table of Contents** of ArcMap to ensure that the UHI result is imported properly (see Figure 15). The UHI field needs to be in numerical format, which is aligned to the right (see the UHI field in Figure 15 left). On the contrary, ArcMap aligns string fields to the left (see the FIPS field in Figure 15 left). If the UHI spreadsheet cannot be shown in ArcMap or with only empty cells, go to **Table of Contents** in ArcMap, right-click the UHI result spreadsheet, and click **Remove**. Use the Excel program to save the UHI result to a comma-separated valued (CSV) file. Close the CSV file in Excel and add the UHI result file in CSV format to ArcMap. Eventually, the **Table of Contents** in ArcMap will have two files: the geopolitical or geo-statistical boundary shapefile and the UHI result file (see Figure 15 right).

Figure 15: The UHI Excel file Imported into ArcMap (left) and the ArcMap Table of Contents (right)



- The fourth step will be to append the UHI result to the geopolitical or geo-statistical boundary shapefile. This is implemented through the “join” function in ArcMap that appends the fields of the UHI result to the attribute of boundary shapefile through the field common to both which is the “key.” This “key” is the geolocation ID in this handbook (see Figure 16).

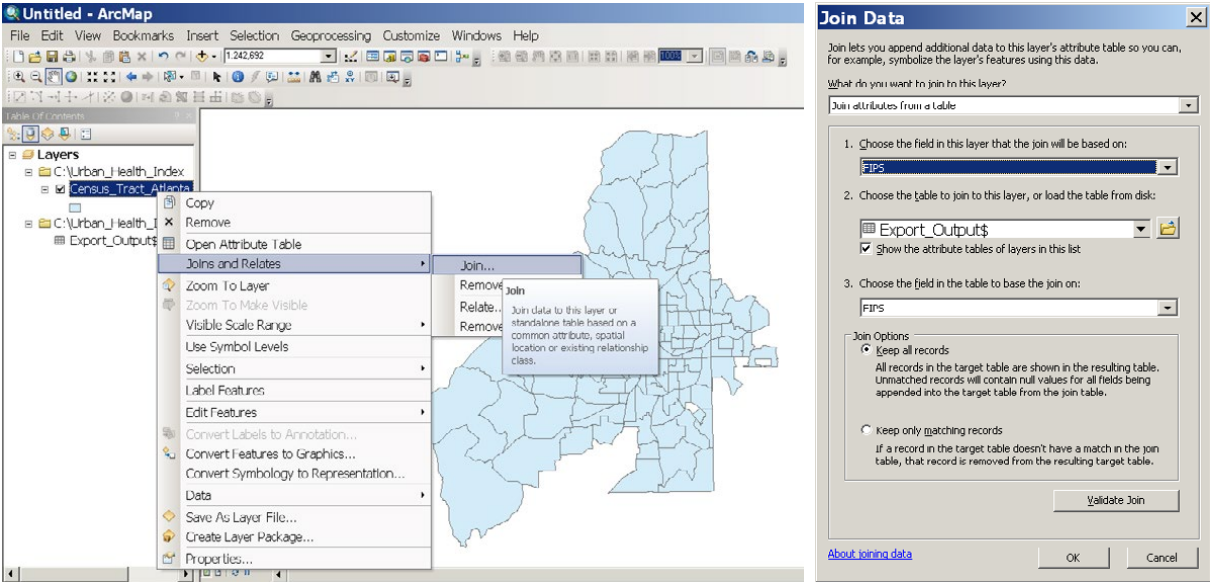
Figure 16: Concept Model of the Join Function in ArcMap



Go to **Table of Contents** in ArcMap, right-click the boundary shapefile, point to **Joins and Relates**, and click **Join** (see Figure 17, left). A **Join Data** window will open. In this window, refer to the parameter setting shown in Figure 17, right. Specifically, select **Join attributes from a table** in the first dropdown menu. In 1 of **Join Data** window, choose the field name of the geolocation ID in the attribute of the boundary shapefile. In 2 of **Join Data** window, choose the UHI result file. In 3 of **Join Data** window,

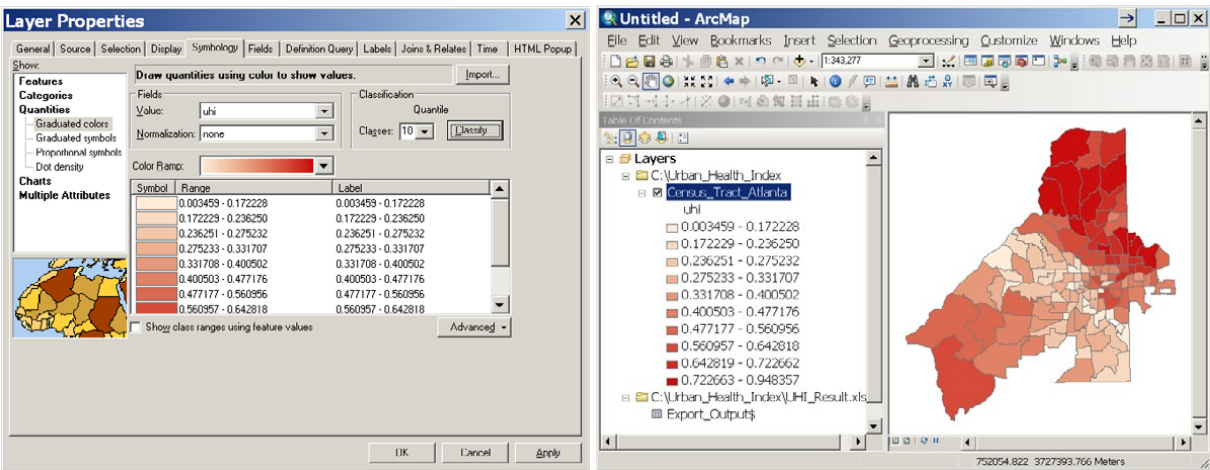
choose the field name of the geolocation ID in the UHI result file. Finally, choose **Keep all records**, click **Validate Join** button to validate the operation, and then click **OK** to dismiss the window.

Figure 17: Process to Select the Join Function (left), and the Join Data Window (right)



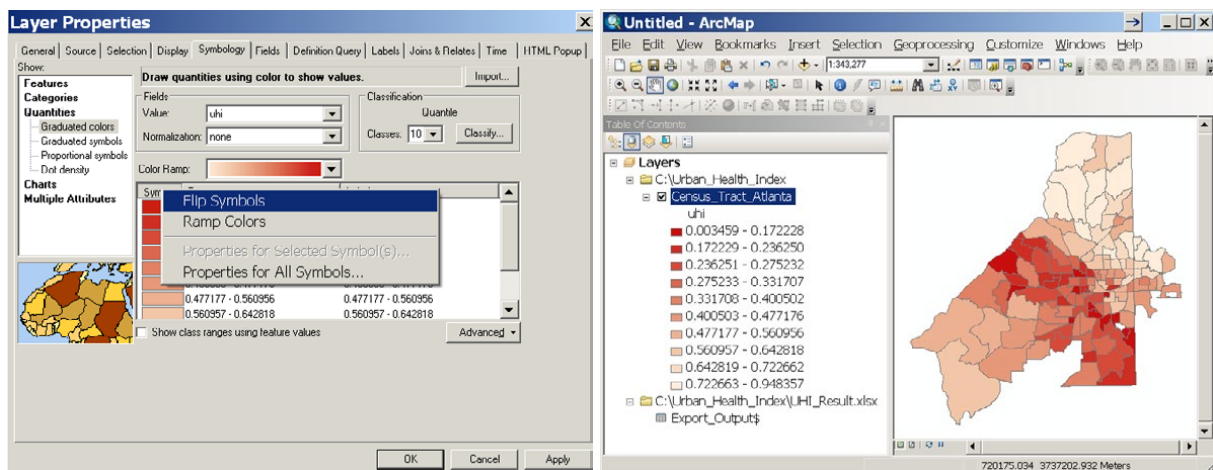
5. The fifth step is to visualize the UHI result. Right-click the geopolitical or geo-statistical boundary shapefile and click **Properties**. In the **Layer Properties** window, go to the **Symbology** tab, click **Quantities**, and click **Graduated Colors**. In the **Field Values**, select the numerical UHI result field and desirable color ramp. Click the **Classify** button and choose **Quantile** as **Classification Method** and 10 classes. This will create deciles. Click **OK** to dismiss the **Classification Window** (see Figure 18, left) and another **OK** to dismiss the **Layer Properties** window. The UHI result will be displayed in ArcMap (see Figure 18, right). The selection of 10 classes using **Quantile** is recommended as it is convenient to isolate the top and bottom 10%. The number of classes or another classification scheme can be adjusted by using the **Classify** button, if necessary.

Figure 18: Symbology Setting (left) and the Choropleth Map (right)



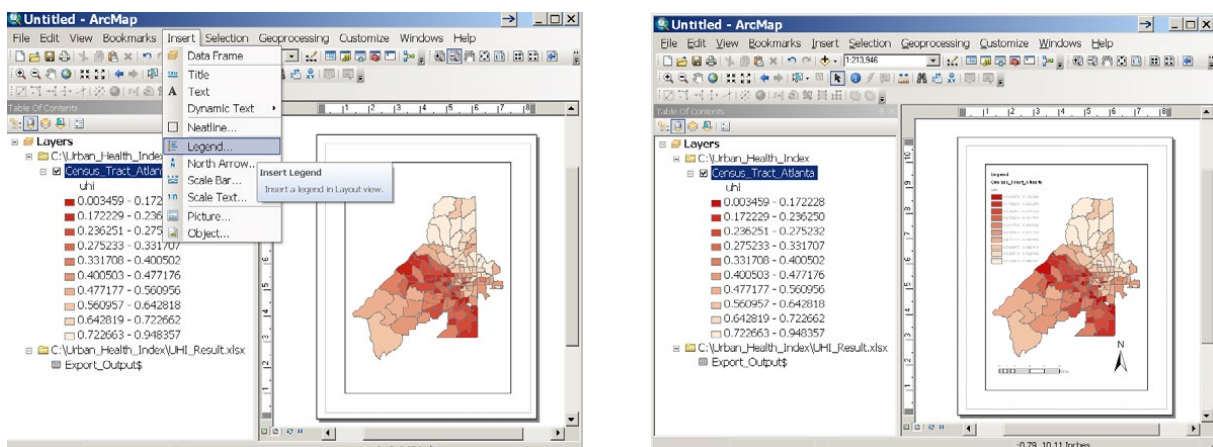
In general, higher UHI values should indicate better health. By default, both ArcGIS and QGIS set the higher (or better) UHI values with darker colors. A better visual effect may be achieved by assigning darker colors to lower (or worse) UHI values to highlight disadvantaged areas. Reverse the color ramp by going to **Layer Properties** window. Click **Symbol** below **Color Ramp** and point to **Flip Symbols** (See Figure 19, left). Click **OK** to dismiss the **Layer Properties** window and the color ramp will be flipped (See Figure 19, right).

Figure 19: Flip Option in Layer Properties Window (left) and the Map after Flipping Map Colors (right)



- The final step is to export the map as an image or as a printable file. Click the **Layout View** on the **View** menu. A set of map elements including legend, scale bar, title are available in **Insert** on the menu once the **Layout View** is activated (see Figure 20, left). In addition, ArcMap allows a portrait or landscape orientation for the map by going to **File** on the menu and clicking **Page and Print Setup**. Use to adjust the size and location of the map element, and use to zoom in, zoom out, pan, and zoom. Once the map design is complete (see Figure 20, right), it can be exported by going to **File** and **Export Map**. The format (i.e., TIFF, EPS, or JPEG) can be selected, and the resolution of the image changed. These images can be directly inserted into document files or presentation files. The user may also print the map by going to **File** on the menu and **Print**.

Figure 20: Insert Menu to Add Map Elements (left) and the Final Map (right)

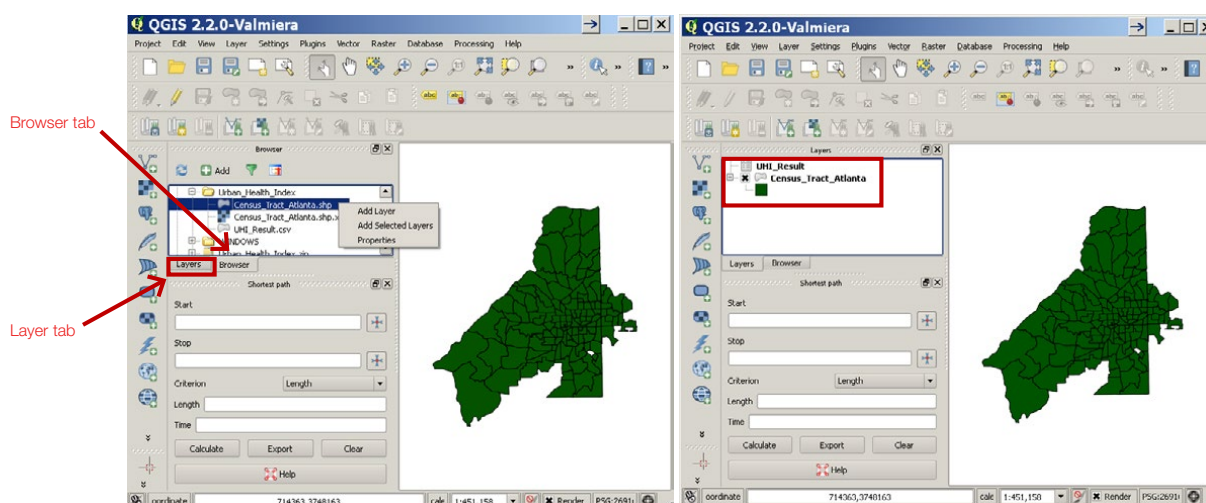


Using QGIS for Mapping the Urban Health Index

This guide instructs how to map the UHI using QGIS, a free, cross-platform and open source desktop GIS program, with a goal of providing flexibility to those who have no access to ArcGIS. It uses the same example data from Using ArcGIS for Mapping the UHI. The version of QGIS used in this instruction is QGIS Valmiera Version 2.2.0 (<http://www.qgis.org/en/site/>). Other QGIS versions will be similar to the Version 2.2.0, and the directions will remain largely the same. Access to the internet is required to download QGIS and required plug-in tools. Online resources, such as the user guide, tips, and tutorials are available for the use of QGIS (http://docs.qgis.org/2.2/en/docs/user_manual/; <http://www.qgistutorials.com/en/>).

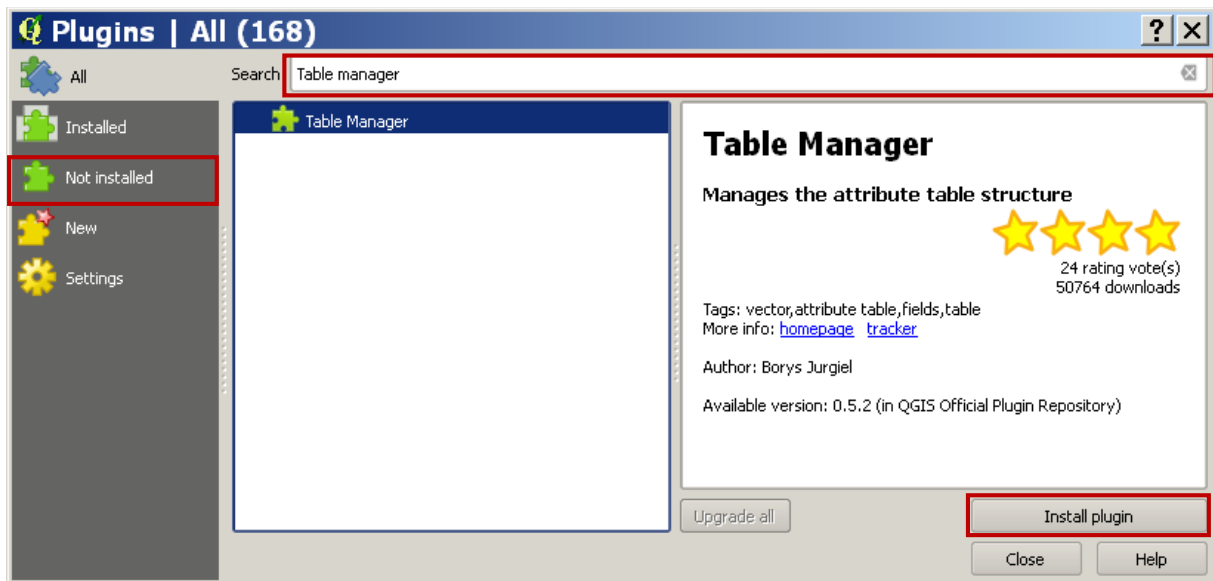
1. **Convert the UHI Excel file to CSV format.** Unlike ArcGIS, QGIS is currently incompatible with Excel, so converting the Excel file to a CSV file is necessary. Go to **File** and click **Save As** and then select **CSV (Comma delimited) (*.csv)** as the save type. Click **Save** to execute and select **Yes** to dismiss the **Warning Window**. Close the Excel file.
2. **Add the geopolitical or geo-statistical boundary shapefile and the UHI data in CSV file to QGIS.**
 - Start QGIS Desktop 2.2.0. Click Browser tab under the **Layers** window (see Figure 21).
 - Navigate to the folder containing the UHI and boundary shapefile.
 - Select the shapefile.
 - Right-click the highlighted shapefile.
 - Select **Add layer** from the drop-down list.
 - Navigate to the folder containing the UHI result in CSV format.
 - Right-click the highlighted CSV file and select **Add layer**.
 - Click **Layers** tab to check that both the boundary shapefile and the CSV file (UHI result) were added (see Figure 21, left).

Figure 21: QGIS Interface of Adding Data (left) and the Browser Tree (right)



3. **Update plugins.** A plugin named **Table Manager** is not included by default installation but is needed for mapping the UHI. Click **Vector** on the main menu and check if **Table Manager** appears in the drop-down list. If **Table Manager** is not present, install (or reinstall) the plugin. Click the **Plugins** tag and click **Install** and **Manage Plugins**. In the **Plugins** window, click the **Not Installed** tag from the list at the left side and then type in "table manager" in the search bar (Figure 22).

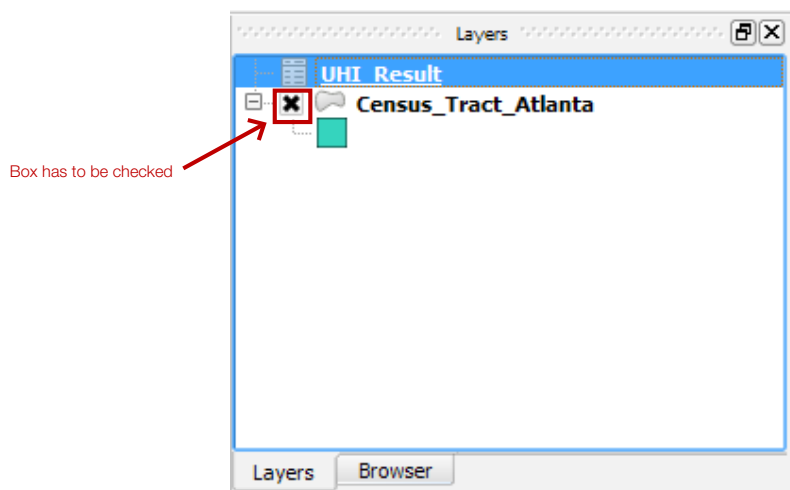
Figure 22: Plugins Interface



Select **Table Manager** from the search results and then click the **Install plugin** button (Figure 22) to install. If it is necessary to reinstall an installed plugin, click the **Installed** tag, select the plugin to reinstall, and then click the **Reinstall** button.

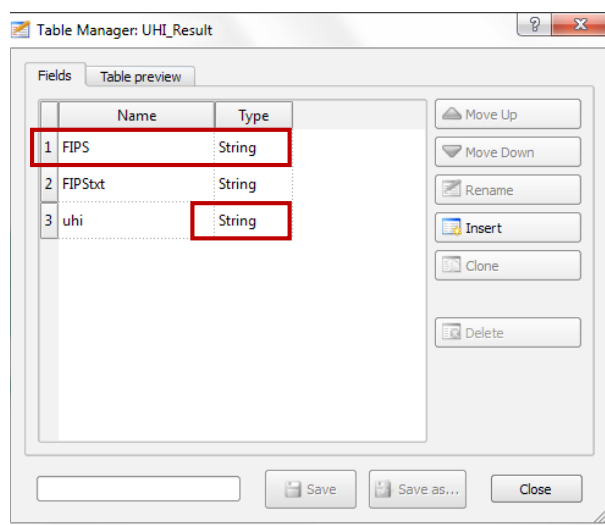
Manage the attribute table. Fields can be added, deleted or just examined by using the **Table Manager**. At the **Layers Window**, select the CSV file by clicking it (Figure 23).

Figure 23: Interface for the Layers Window



Click **Vector** in the main menu, go to **Table Manager** and then click **Table Manager**. In the **Table manager** window, view the fields of the attribute table of the CSV file. Determine which is the key field used to join the CSV file to the geopolitical or geo-statistical boundary shapefile.

Figure 24: Checking the Data Type for UHI Values When Joining UHI Data to the Shapefile

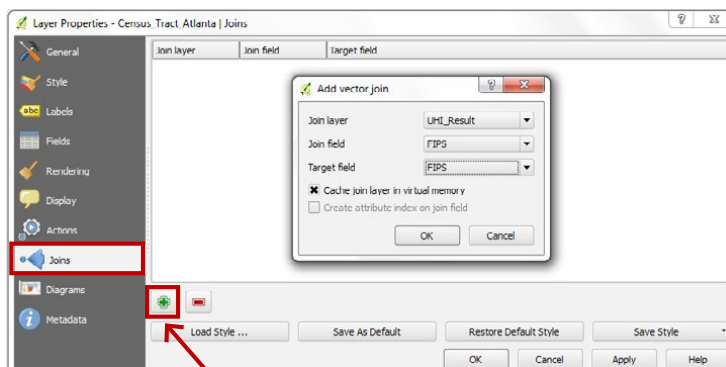


It is necessary to check the data type of the field of the UHI values (Figure 24). During the conversion from Excel to CSV, the numerical UHI values may convert to the string format (see the UHI field in Figure 24). Make a numerical field using the **Double** format in order to map the UHI result. Double format refers to double-precision floating point numbers (http://en.wikipedia.org/wiki/Double-precision_floating-point_format). It stores fractional numbers with more precision. Therefore, create a new field and use field calculation to obtain the UHI result in the **Double** format. In order to do so, first, join the CSV file to the geopolitical or geo-statistical boundary shapefile.

4. Join the CSV file to the boundary shapefile (see Figure 25).

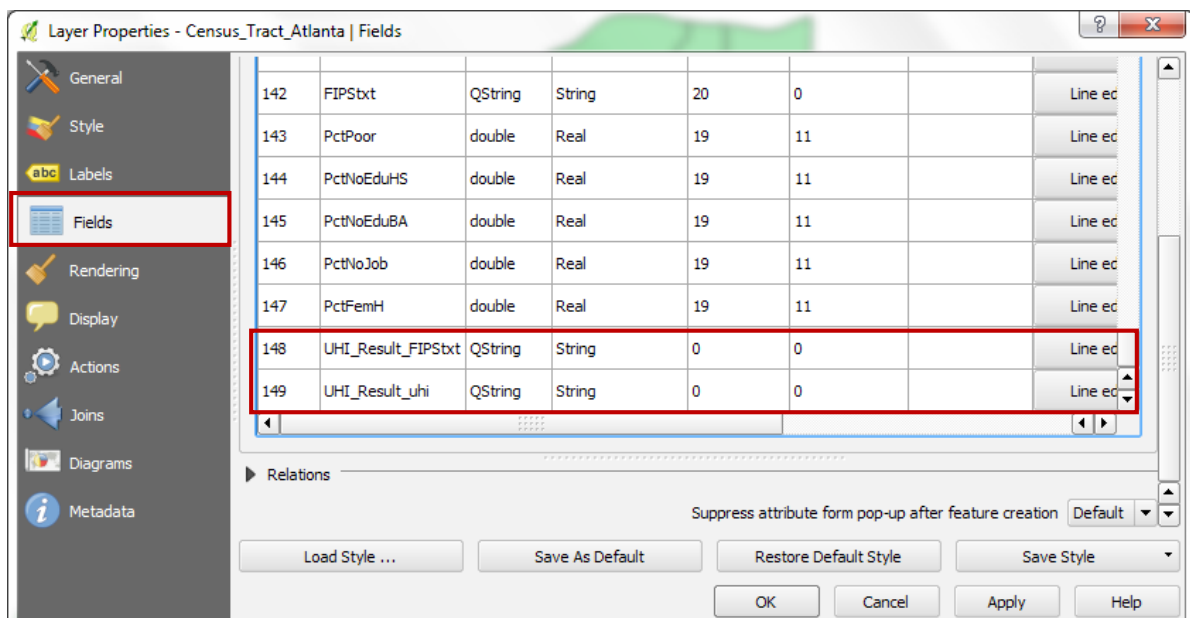
- In the **Layers Window** at left bottom of screen, right-click the boundary shapefile.
- Select **Properties**.
- In the **Layer Properties** window, click **Joins** tag on the left side.
- Click the **green plus** button to add a new join.
- In the **Add vector join** window, select the CSV file for **Join layer**, select the key field in the CSV file which the join will be based on for **Join field**, and select the key field in the boundary shapefile for **Target field**.
- Click **OK** to dismiss the **Add vector join** window.

Figure 25: Joining the UHI File in CSV Format to a Boundary Shapefile



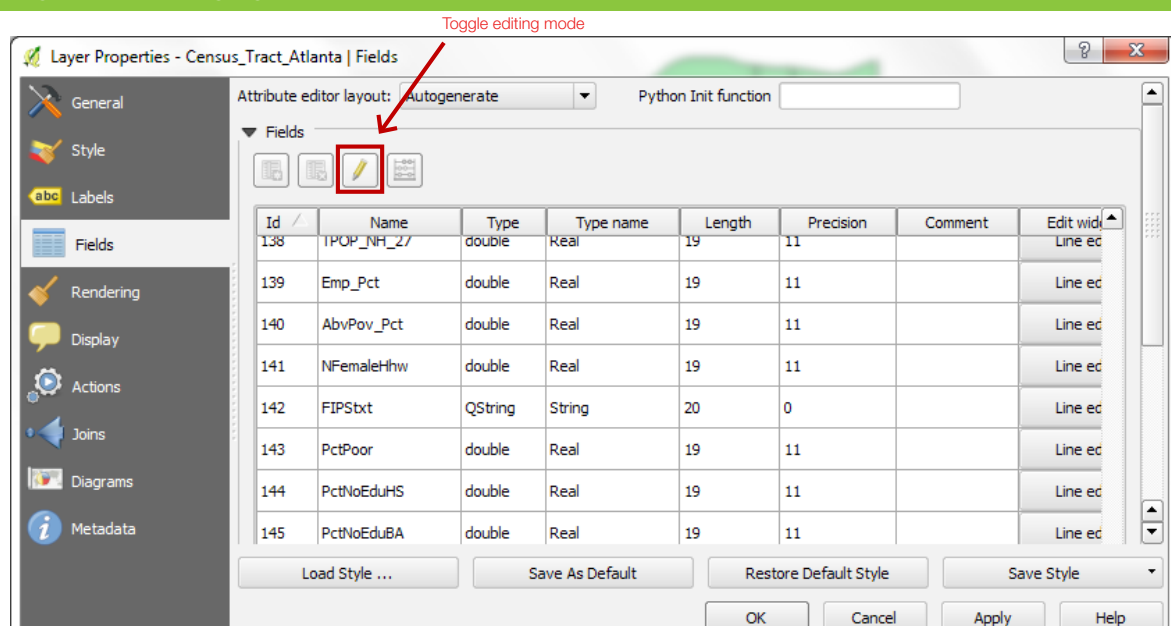
In the **Layer Properties** window (Figure 26), the user can click the **Fields** tag on the left, scroll down to the bottom of the attribute table, and check whether the fields in the CSV file are attached. Because the UHI result field might be a string format, one will need to change it to a double format.

Figure 26: Viewing Fields in the Layer Properties Window



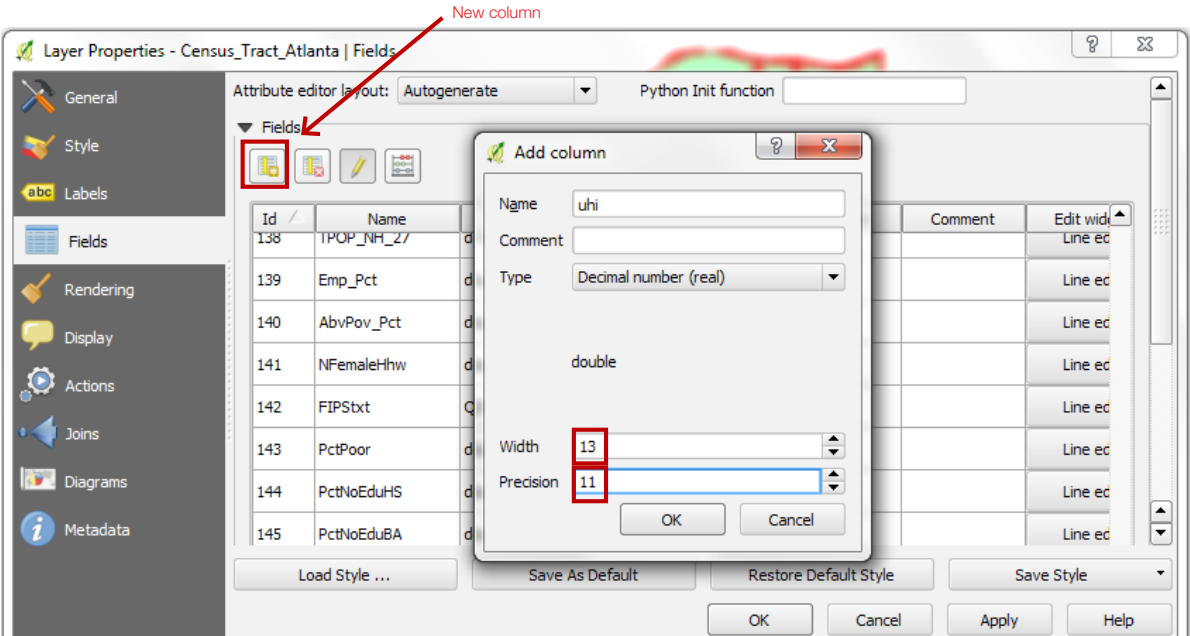
- 5. Change String to Double using field calculation.** In the **Fields** tag of the **Layer Properties** window, the user may click the **toggle editing mode** button (see Figure 27). The click will trigger the feature editing mode. Under this mode, geometries of the shapefile or the content of its attributes can be edited. In Figure 27, values in the attribute table are changed.

Figure 27: Changing Layer Properties



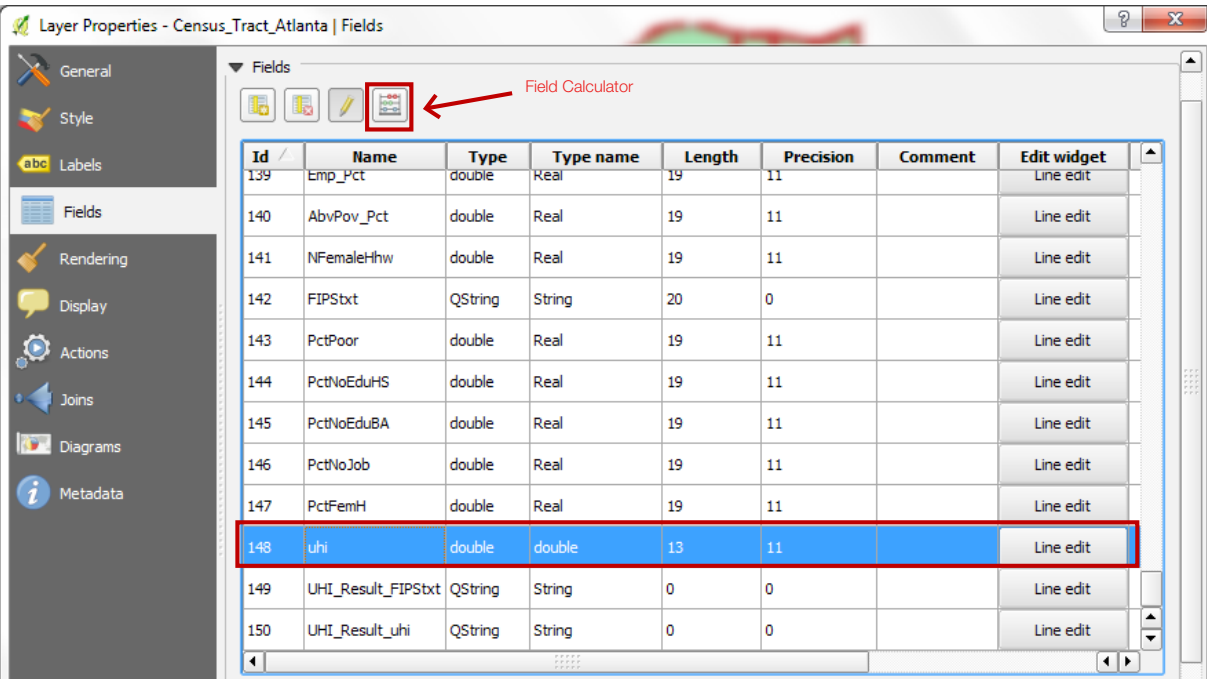
Click the **New Column** button on the **Layer Properties** window (Figure 28). In the **Add column** window, give a name to this new column for the UHI result and select **Decimal number (real)** for **Type**. Type in 13 for **Width** and 11 for the **Precision**. Click **OK** to dismiss the **Add column** window.

Figure 28: Add Column Window



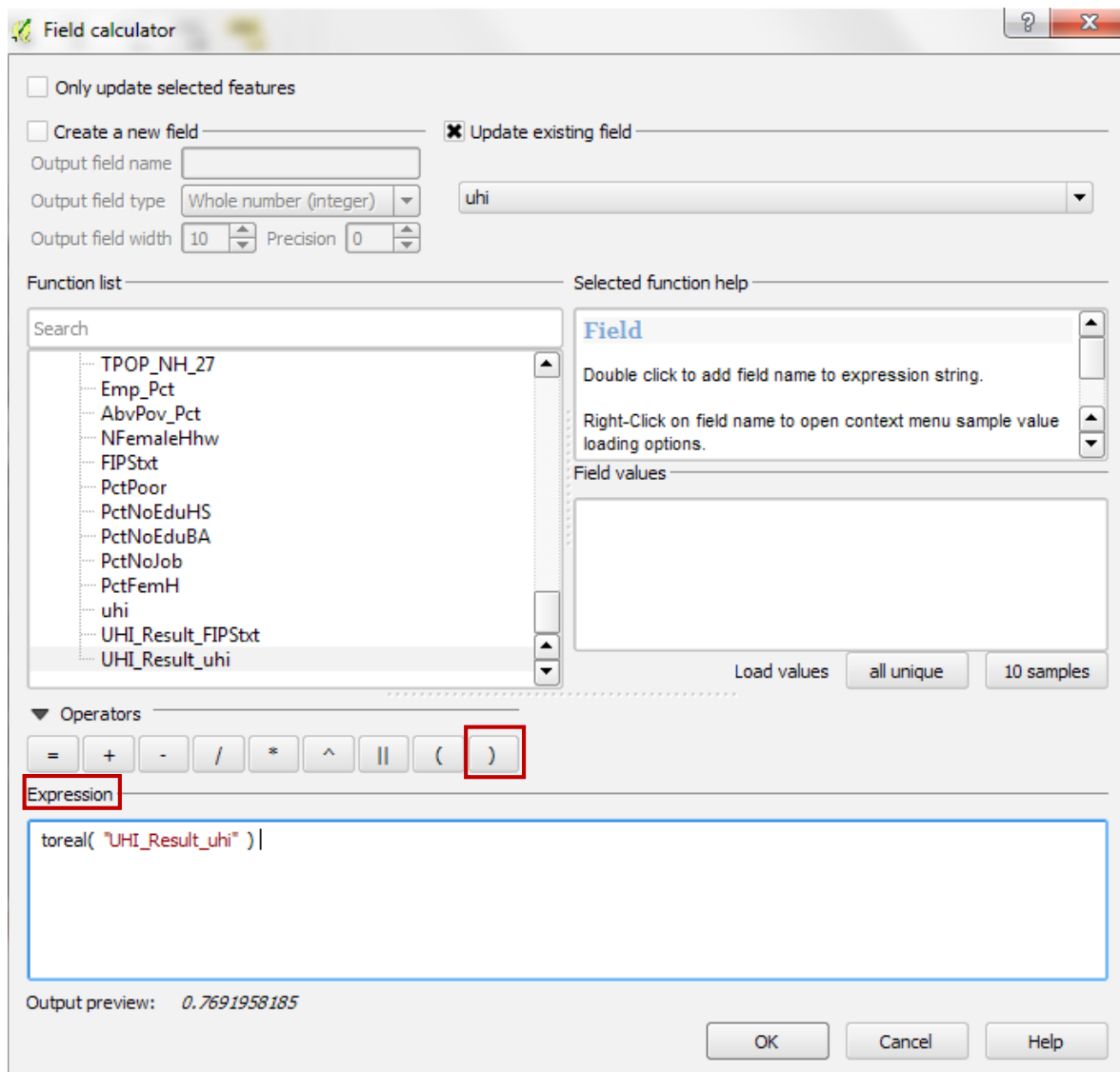
Verify whether the new numerical field is added (Figure 29). Now, update the values for this field. Click the **Field calculator** button (Figure 29).

Figure 29: Layer Properties Window after the New Column is Added



In the **Field calculator** window (Figure 30), set parameters as shown in Figure 30. Check the box besides Update existing field and then select the numerical field recording the UHI values from the drop-down list. In the **Function** list, double click **Conversions** and double click **Toreal** (the toreal function will convert a string into a real number with a fractional part). In the **Function** list, scroll down to the bottom, double click **Fields and values**, and double click the string field recording the UHI values. Single click the right bracket button, and click **OK** to dismiss the **Field calculator** window.

Figure 30: Field Calculator Window



6. **Complete the calculation and check the calculation result.** Go back to the **Layer Properties** window, click the **toggle editing mode** button to turn off the editing mode, and click **Save** for the **Stop editing** warning window to save the field calculation. Click **OK** to dismiss the **Layer Properties** window. Right click the boundary shapefile from the **Layers** window and select **Open Attribute Table**. Scroll to the right side and examine the UHI values of the numerical field (Figure 31). The values of the UHI results in the string format (being aligned left) now transferred to the numerical field (being aligned right) can be seen. Close the attribute table of the boundary shapefile.

Figure 31: Attribute Table Window UHI Data

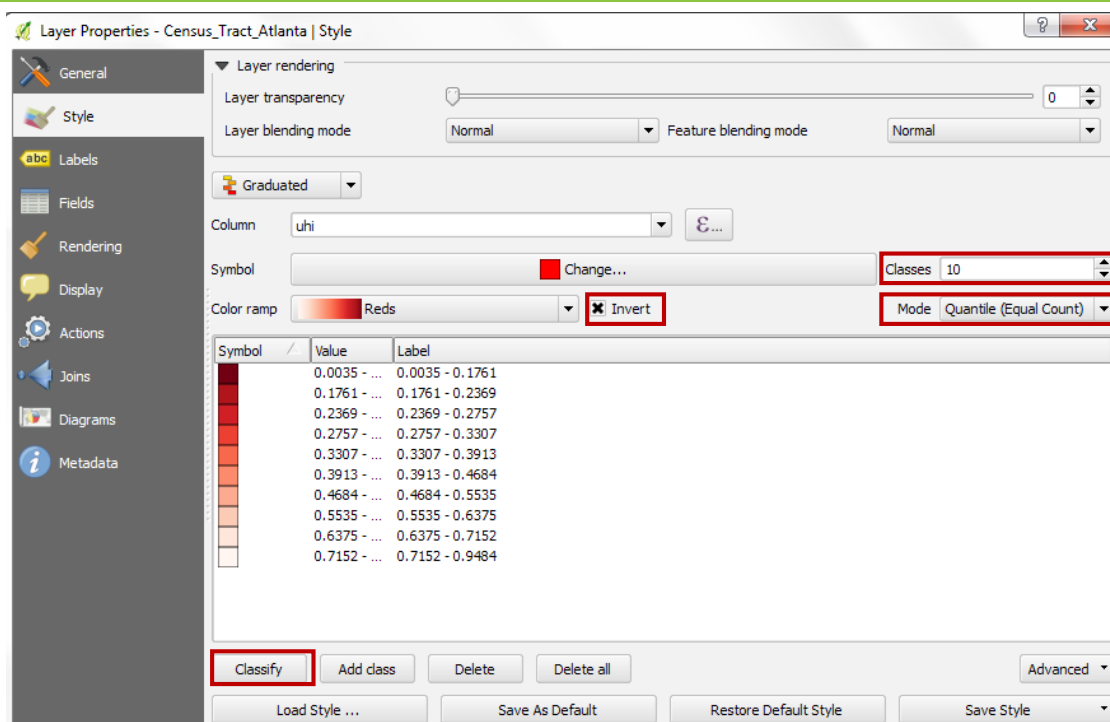
Attribute table - Census_Tract_Atlanta :: Features total: 128, filtered: 128, selected: 0

xt	PctPoor	PctNoEduHS	PctNoEduBA	PctNoJob	PctFemH	uhi	UHI_Result_FIPStxt	UHI_Result_uhi
0	2.8000000000	5.7000000000	22.3000000000	1.0000000000	4.2000000000	0.76919581850	13089020100	0.76919581850
1	0.0000000000	2.2000000000	29.7000000000	0.0000000000	0.0000000000	0.72266211210	13089020200	0.72266211210
2	0.9000000000	5.5000000000	20.8000000000	7.3000000000	6.7000000000	0.69940131220	13089020300	0.69940131220
3	0.0000000000	0.7000000000	25.9000000000	5.5000000000	2.6000000000	0.64664926170	13089020400	0.64664926170
4	19.2000000000	26.7000000000	68.8000000000	7.1000000000	9.5000000000	0.40423299950	13089020500	0.40423299950
5	51.9000000000	12.7000000000	64.8000000000	16.8000000000	21.6000000000	0.36001688060	13089020600	0.36001688060
6	6.5000000000	26.2000000000	69.1000000000	4.3000000000	9.6000000000	0.41919020020	13089020700	0.41919020020
7	6.4000000000	10.5000000000	60.2000000000	9.8000000000	3.0000000000	0.56095622210	13089020801	0.56095622210
8	29.3000000000	13.1000000000	66.2000000000	13.8000000000	12.7000000000	0.42029005830	13089020802	0.42029005830
9	15.6000000000	18.3000000000	61.1000000000	6.1000000000	7.3000000000	0.46625258330	13089020900	0.46625258330

Show All Features

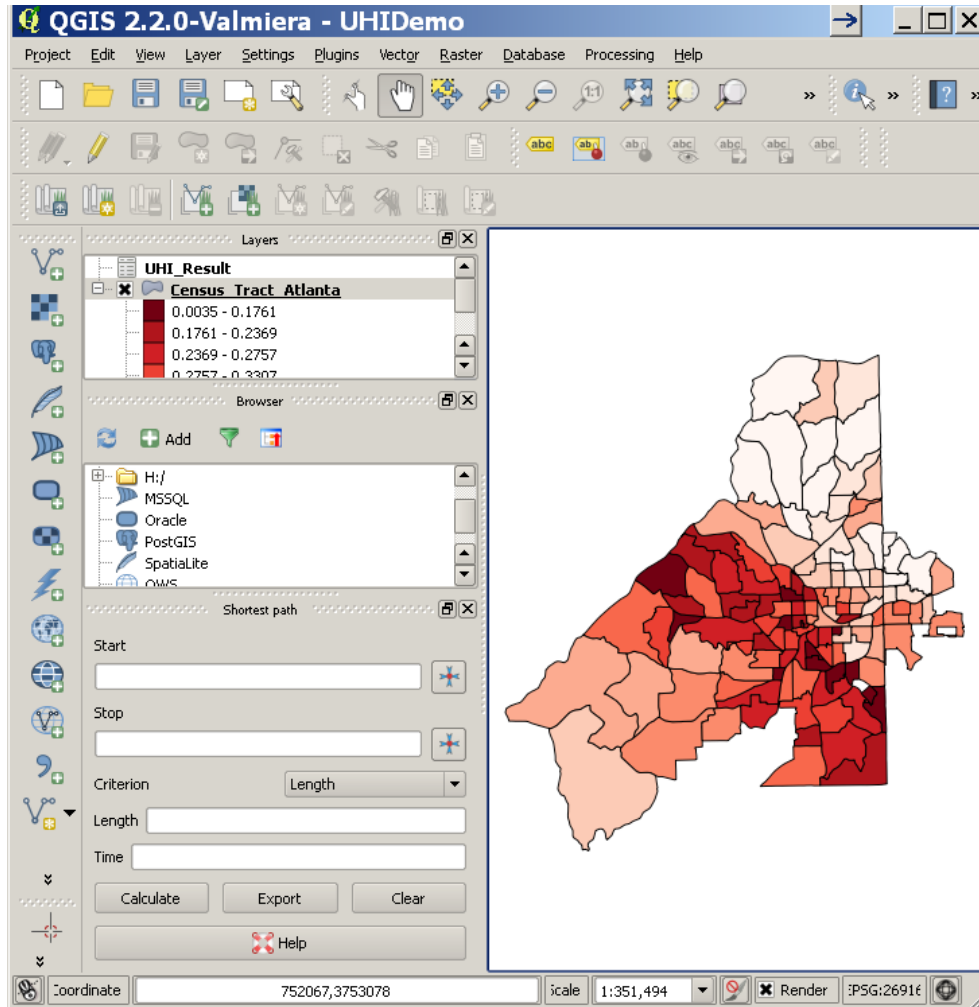
7. **Create a choropleth map of the UHI.** Now, make a choropleth map of the UHI. In the **Layers** tab, right click the boundary shapefile and select **Properties**. In the **Layer Properties** window, click the **Style** tag, then set the parameters as shown in Figure 32. Click **Single** symbol to activate the drop-down list and select **Graduated**. For **Column**, select the numerical field recording the UHI, and determine the desired classification method, number of classes, and the color scheme. The demonstration in Figure 32 increases the number of classes from 5 (default) to 10 and changes mode to **Quantile (Equal Count)**. It also changes the **Symbol** to a red color and selects **Reds** for the **Color ramp**. If flipping the order of the color ramp is necessary, check the box beside **Invert** and click the **Classify** button to invert the order or direction of the color ramp. Click **OK** to dismiss the **Layer Properties** window.

Figure 32: Creating a Choropleth Map in the Layer Properties Window



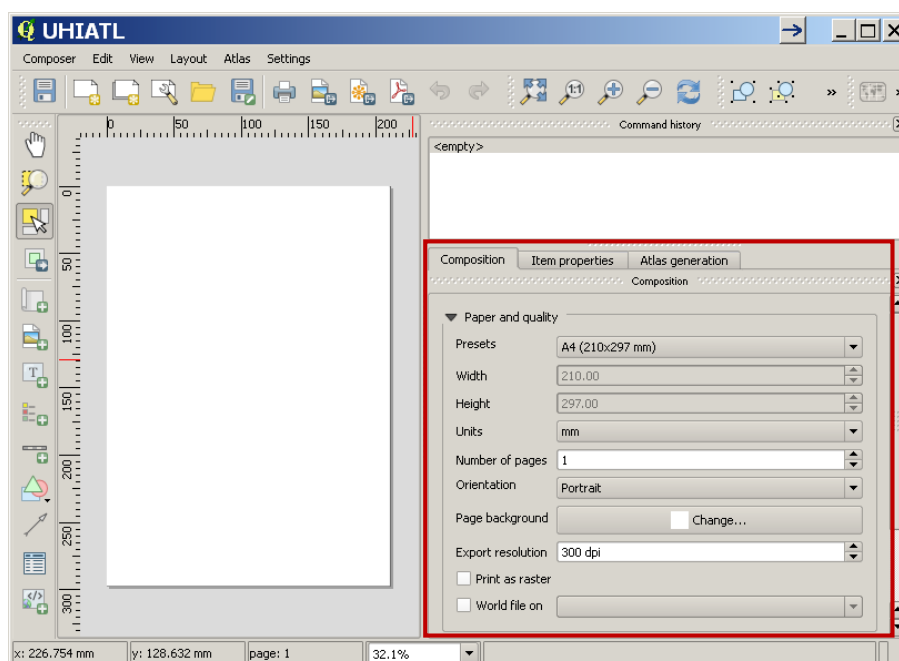
The UHI result can now be seen using a choropleth map in QGIS (Figure 33).

Figure 33: UHI Results in a Choropleth Map



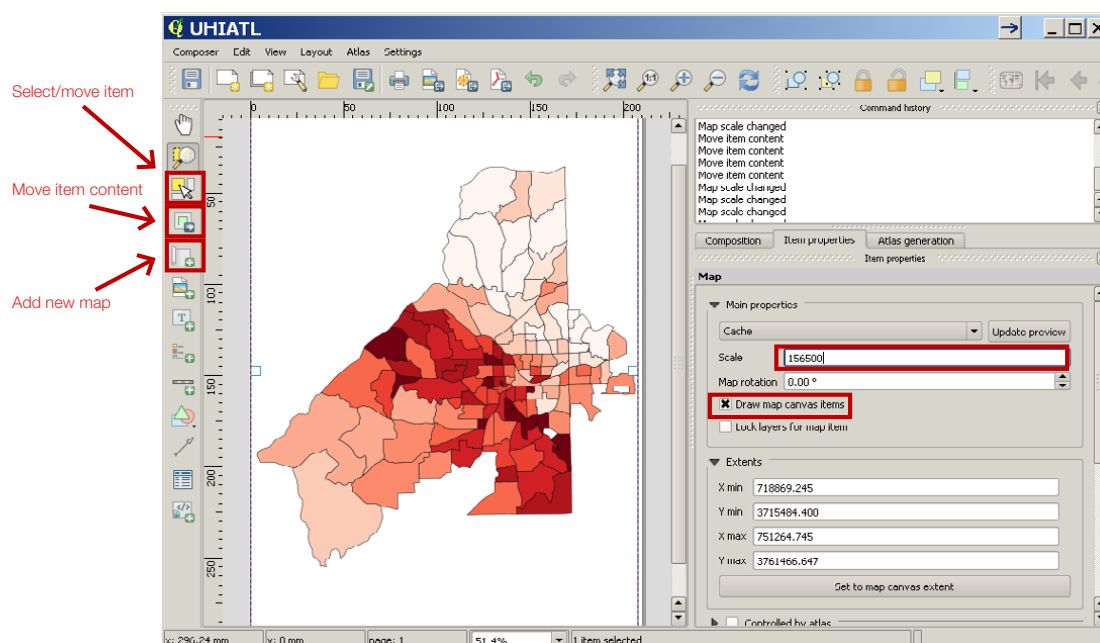
8. **Export the UHI map.** If the UHI map needs to be exported, click **Project** from the main menu and click **New Print Composer**. In the **Composer Title** window, name the composer and click **OK** to enter the **Print Composer** window for map design (Figure 34). The **Print Composer** window consists of three tabs: **Composition**, **Item properties**, and **Atlas generation**. The **Composition** tab allows one to set the paper size, orientation, and export resolution. The resolution is set at 300 dpi or higher to achieve best print quality.

Figure 34: Paper and Quality Set-Up in Print Composer Window



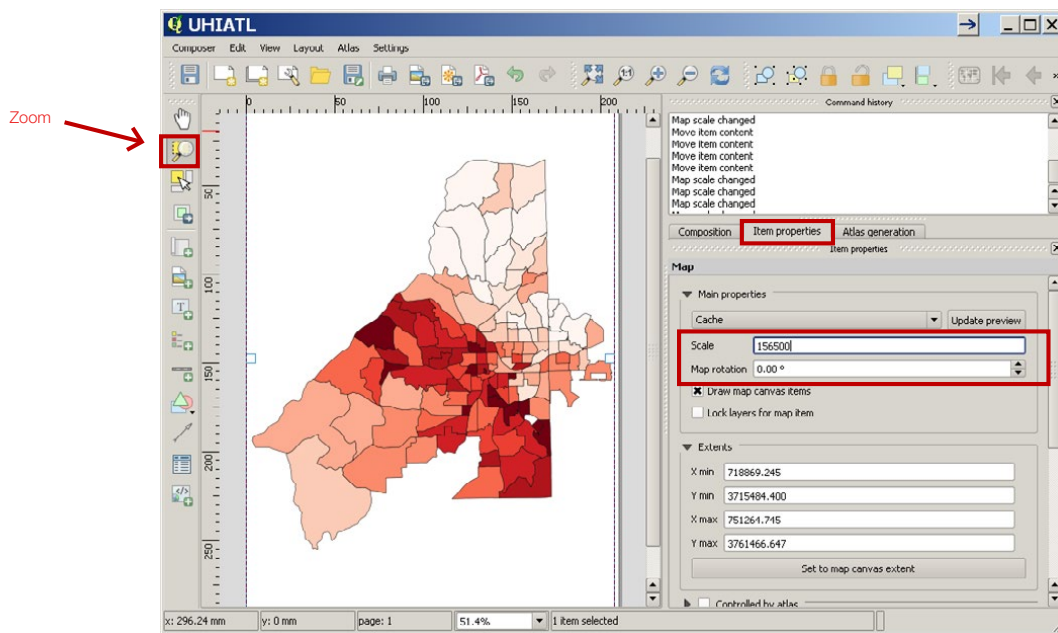
- Click **Add new map** button from the tool bar on the left side to activate tool.
- With this tool activated, click and drag a box on the blank page.
- The UHI map is now added to the print composer (Figure 35).

Figure 35: Add, select, and Move a Map in the Print Composer Window



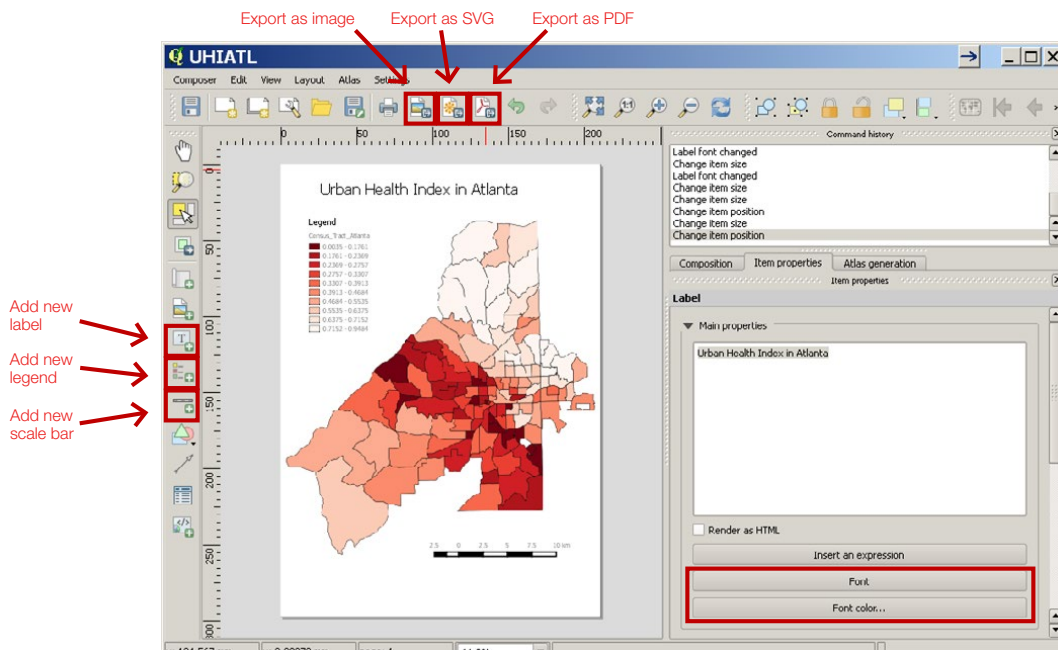
The Item properties tab (Figure 36) on the right side allows a user to adjust the scale of the map or the map rotation.

Figure 36: Item Properties in the Print Composer Window



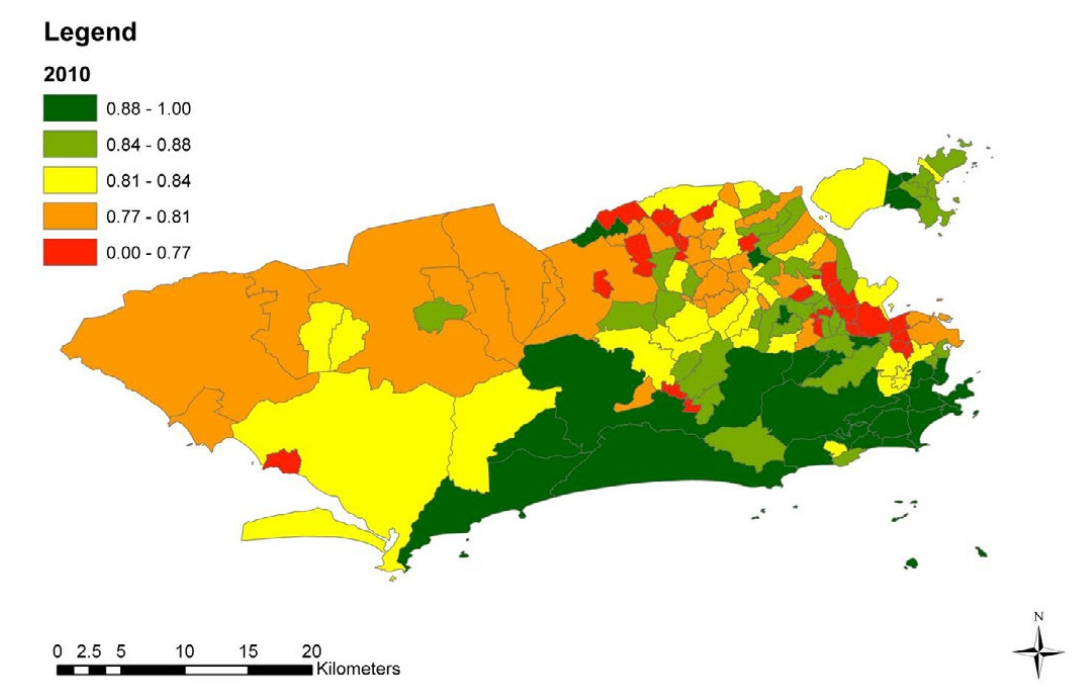
Then, use the **Add new label**, **Add new legend**, and **Add new scale bar** icons (Figure 37) to add map elements such as title, legend, and scale bar to the print composer. Click an icon, then click the desired location on the map to insert the element. Adjust the size of each map element by clicking and dragging its corner. For text elements on the map, use the options of **Font** and **Font color** (Figure 37) to make changes to the text. When the map design is complete, use icons at the top to export files to image, scalable vector graphics (SVG; http://en.wikipedia.org/wiki/Scalable_Vector_Graphics), or PDF file formats. SVG files are often used in Adobe Illustrator.

Figure 37: Selecting Map Elements to Print in the Composer Window



Appendix E: A Visual Summary for Communicating the Urban Health Index⁶⁰

Figure 1. Map displaying Urban Health Index in 142 neighborhoods of Rio de Janeiro Municipality, 2010.



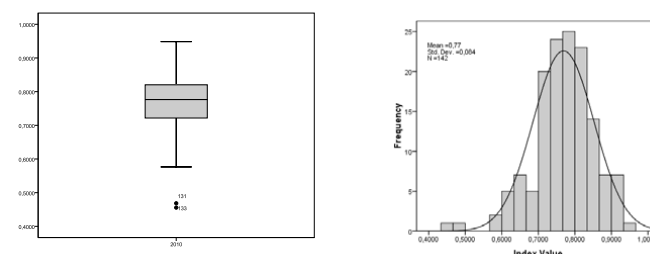
Index

Disparity Ratio	1,35
Disparity Gradient	0,18
SE (MR)	7,94

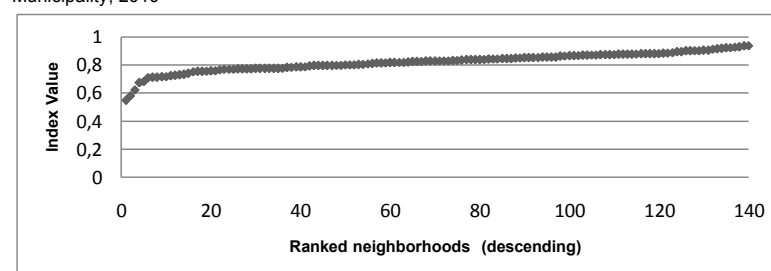
UHI Distribution

Mean	0,82
Standard Deviation	0,07
Minimum	0,55
Maximum	0,96
Range	0,41
Median	0,83
10th Percentile	0,74
90th Percentile	0,90

Graph 1a, 1b. Boxplot and Histogram with distribution of UHI values in 142 neighborhoods of Rio de Janeiro Municipality, 2010.

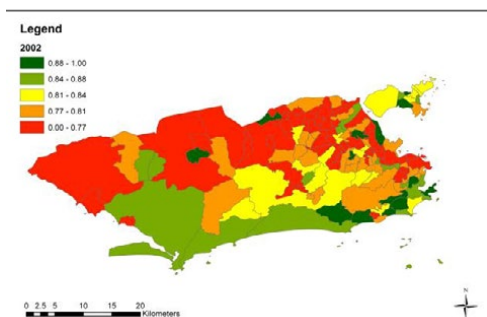


Graph 2. Index Plot of UHI values in 142 neighborhoods of Rio de Janeiro Municipality, 2010



⁶⁰ Template courtesy of Martin Bortz. Based on research reported in: Bortz M, Ramroth H, Barcellos C, Weaver SR, Rothenberg R, Magalhaes M, Kano M. Disaggregating health inequalities within Rio de Janeiro, 2002-2010, by applying an Urban Health Inequality Index. Cadernos de Saúde Pública. In press. Indicators: mortality rates for diabetes, ischemic heart disease, breast and cervix cancer, HIV, tuberculosis, infant deaths, traffic accident, and homicide for the 160 wards of Rio de Janeiro, Brazil. Data obtained from DATASUS/TABNET.

Figure 2. Map displaying the UHI in Rio de Janeiro Municipality, 2002.



Graph 3. Ranked distribution of UHI values in Rio de Janeiro Municipality, 2002 and 2010.

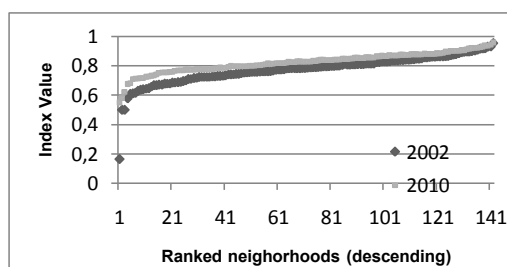
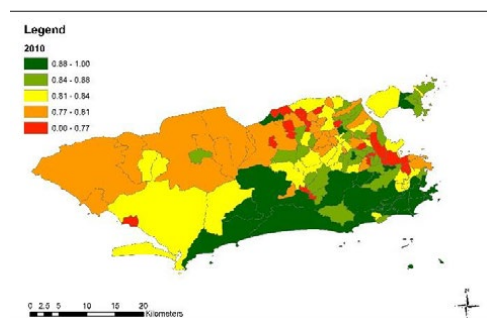


Figure 3. Map displaying UHI in Rio de Janeiro Municipality, 2010.



Graph 4. Boxplot with distribution of UHI values in Rio de Janeiro Municipality, 2002 and 2010.

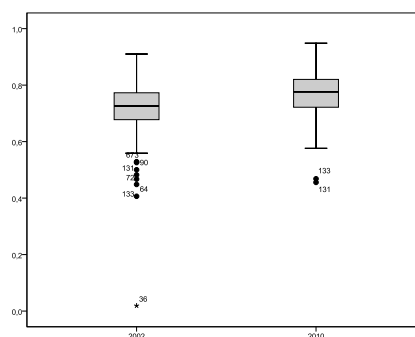
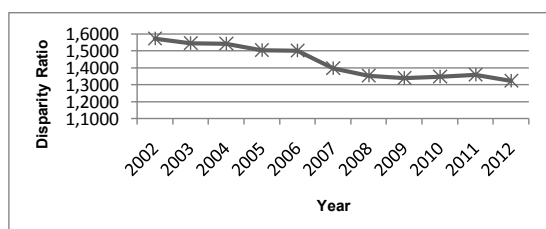


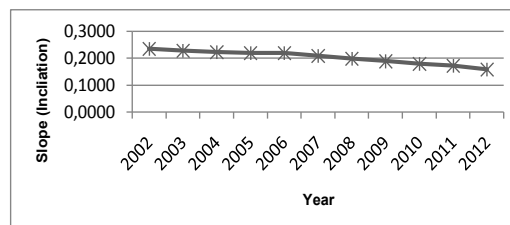
Table 1. Descriptive statistics for UHI Rio de Janeiro Municipality, 2002-2012.

Index	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Disparity Ratio	1.57	1.55	1.54	1.50	1.50	1.40	1.35	1.34	1.35	1.36	1.32
Disparity Gradient	0.23	0.23	0.22	0.22	0.22	0.21	0.20	0.19	0.18	0.17	0.16
Standard Error	9.03	8.73	8.65	8.66	8.56	8.43	8.21	7.99	7.94	7.87	7.20
Index Values	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Mean	0.77	0.78	0.79	0.79	0.79	0.80	0.81	0.82	0.82	0.83	0.85
Standard Dev	0.10	0.10	0.10	0.09	0.09	0.08	0.07	0.07	0.07	0.07	0.07
Minimum	0.16	0.17	0.12	0.19	0.22	0.45	0.56	0.56	0.55	0.53	0.52
Maximum	0.95	0.95	0.94	0.93	0.92	0.93	0.94	0.95	0.96	0.96	0.96
Range	0.79	0.78	0.81	0.74	0.70	0.48	0.38	0.39	0.41	0.43	0.45
Median	0.79	0.79	0.79	0.80	0.80	0.82	0.82	0.83	0.83	0.84	0.86
10th Percentile	0.67	0.69	0.69	0.68	0.69	0.70	0.72	0.73	0.74	0.75	0.78
90th Percentile	0.88	0.88	0.90	0.90	0.89	0.89	0.89	0.90	0.90	0.90	0.93

Graph 5. Disparity Ratio for UHI in Janeiro Municipality, 2002-2012.



Graph 6. Disparity Gradient for UHI in Rio de Janeiro Municipality, 2002-2012.



Appendix F: A Presentation Template for Communicating the Urban Health Index Results

Slide #1

Title of Presentation

Topic of Presentation

Presenter’s Name, Organization, Date

Slide #2

Motivation

Purpose of Presentation:

Graphic (optional)

Major Issues Discussed

Expected Outcome:

Slide #3

Background

This section is highly variable and must be tailored to the audience. For example:

- (1) An international presentation:
 - a) Describe the urban area in general (population size, density, land mass).
 - b) Describe the small area jurisdiction that is available.
 - c) Describe what is known about level of health and health disparity.
 - d) Provide greater detail on the particular area of observation.

- (2) A presentation to a decision maker:
 - a) Brief history of the problem.
 - b) State area of focus and rationale behind the choice.
 - c) Explain urgent issues requiring resolution.
 - d) Explain recent development of flexible methods to assess the problem.

Map of Area

Slide #4

Calculating The Urban Health Index

1. Provide brief background on its development:
 - a) Responds to need for a flexible approach, rather than a fixed Index.
 - b) Can be adopted to a variety of circumstances.
 - c) Can accommodate different data sources.
 - d) Can combine data from different domains.

2. Construction of the UHI

- a) Standardize the indicators:

$$\text{Standard Value (SV)} = \frac{\text{Value}_i - \text{min}}{\text{max} - \text{min}}$$

Note: Minimum (min) and maximum (max) refer to lower and upper goalposts. The choice of minimum and maximum values is complicated and may require in depth explanation, depending on the audience and setting.

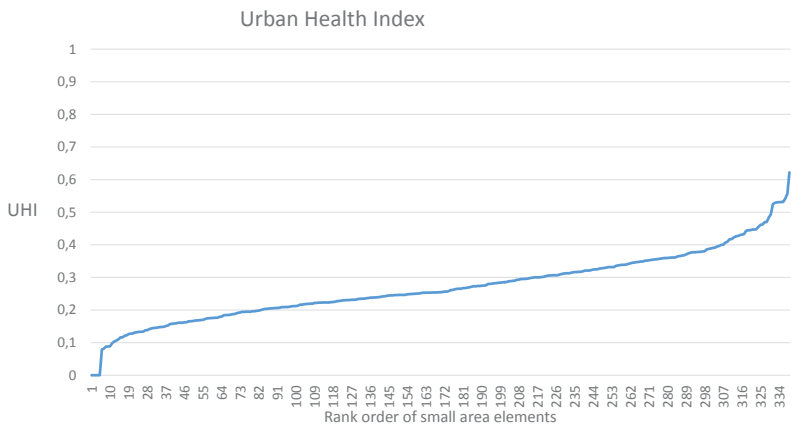
- b) Combine the indicators:

$$\left(\prod_{i=1}^n I_i \right)^{\frac{1}{n}} = \overbrace{(I_1)(I_2)(I_3)}^{\text{for 3 indicators}}^{\frac{1}{3}} = \text{UHI for a given area}$$

Slide #5

Graphing the Urban Health Index

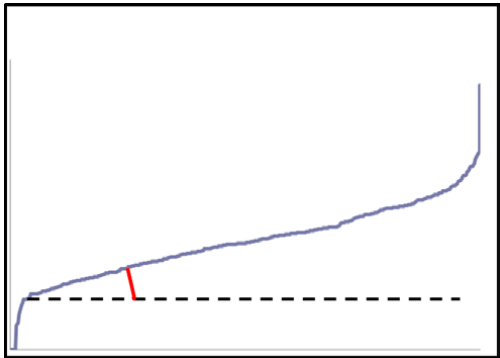
The UHIs for each area are arranged in rank order and the rank ordering is graphed against the UHI values



The graph is linear in its middle portion but veers dramatically up and down at the high and low end.

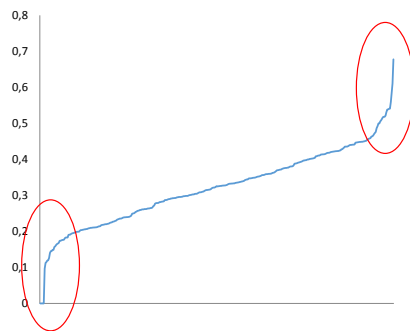
Slide #6

The Health Disparity Slope



Slide #7

The Health Disparity Ratio



The health disparity ratio is the mean of the highest decile divided by the mean of the lowest decile.

A high ratio indicates marked disparity between the highest and lowest groupings.

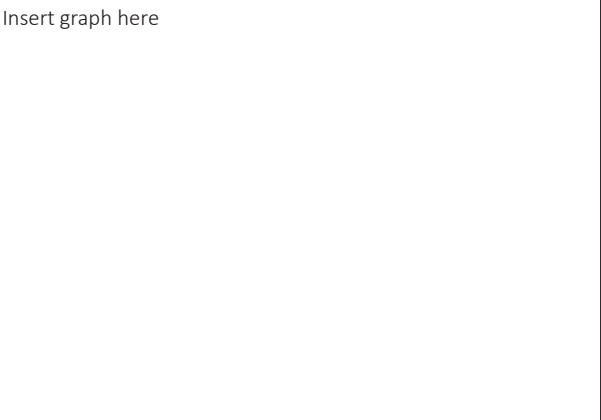
Slide #8

The Data Sources we used for this analysis

1. List the major sources of data used for analysis:
 - a)
 - b)
 - c)
 - d)
2. Discuss reasons for choosing the particular indicators.
3. Discuss strengths and weakness.

Slide #9

Graph of the UHI for this area



Slide #10

The Health Disparity Slope and Ratio for this area:

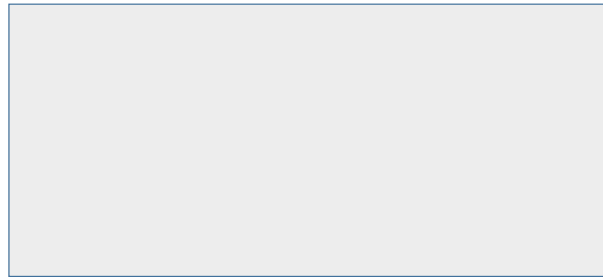
Report the slope and discuss its meaning for this area

Report the ratio and discuss its meaning for this area

Slide #11

Map of the UHI for this area

Insert a map of the area displaying the gradient of UHI in this community.



Comment on the meaning and interpretation of the map.

Slide #12

Other Findings and Analyses

Other analyses might include:

- Other combinations of indicators that you used.
- Different maxima and minima employed and their effect on the rankings and the map.
- Other analytic techniques (i.e., correlation, regression) to look at specific issues.
- Ancillary data outside this framework that illuminate these issues.
- Comparison with prior work.
- Analyses in program.

Slide #13

Conclusions

A concise summary of the major findings and their implications.

Slide #14

Recommendations

Suggest goals and targets based on UHI results.

Understand not all recommendations will be required in all circumstances.

If appropriate, give careful thought to the presentation context and the level of recommendation considered.

This handbook is written primarily for public health workers and policy analysts who require a method for quantifying urban health disparities and presenting such information to policy planners and decision makers. Research and academic communities may also have an interest. The handbook guides the reader through the step-by-step construction and calculation of the **Urban Health Index** - a single, customizable metric that can be used to measure and map the disparities in health determinants and outcomes in urban areas. Actual applications of the method in Atlanta (USA), Tokyo (Japan), Shanghai (China), and Rio de Janeiro (Brazil) are provided as illustrative examples. The technical appendices describe an accompanying Microsoft Excel-based tool for calculating the Urban Health Index.

ISBN 978 92 4 150780 6



9 789241 507806