

GEOGRAPHIC INEQUITIES IN LOCAL HIGHER EDUCATION OPPORTUNITIES?
IDENTIFYING COLLEGE DESERTS WITH CRITICAL GIS

A Thesis

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

CHRISTOPHER KRAUSE

Dr. Timothy Matisziw, Thesis Supervisor

MAY 2017

The undersigned, appointed by the dean of the Graduate School, have examined the thesis
GEOGRAPHIC INEQUITIES IN LOCAL HIGHER EDUCATION OPPORTUNITIES?
IDENTIFYING COLLEGE DESERTS WITH CRITICAL GIS

presented by Christopher Krause,

candidate for the degree of master of arts

and hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Timothy Matisziw

Dr. Matthew Foulkes

Dr. Amalia Dache-Gerbino

Acknowledgements

I would like to thank my advisor, Dr. Timothy Matisziw, for his support throughout the thesis writing process. For helping me plan out my thesis in the very early drafting stages, I would like to thank my other committee members: Dr. Matthew Foulkes and Dr. Amalia Dache-Gerbino. Special thanks are due to Dr. Dache-Gerbino for continually inspiring me to utilize research for social justice.

Outside of my thesis committee, there are several other individuals who I would also like to acknowledge: my father for his many brainstorming sessions, Dena Lane-Bonds for being my ambassador to the Education, Leadership, and Policy Analysis Department, Julie Saperstein for helping me maintain sanity, and my wife, Cheyenne, for her constant encouragement during my sleepless days and nights of thesis work.

Table of Contents

Acknowledgements.....	ii
List of Figures.....	v
List of Tables.....	ix
List of Abbreviations.....	x
Abstract.....	xi
Chapter 1: Introduction.....	1
Chapter 2: Literature Review.....	3
2.1: Conceptual frameworks.....	3
2.2: Spatiality of conceptual frameworks.....	4
2.3: Parallels to food desert literature.....	5
2.4: Access and accessibility.....	6
2.5: GIS as a methodology.....	7
2.6: Critique of existing college access and college proximity literature.....	7
Chapter 3: Methodology.....	10
3.1: Measuring proximity.....	10
3.2: Measuring access.....	14
Chapter 4: Application of Methodology.....	16
4.1: Data Sources.....	17
4.1.1: College Scorecard.....	17
4.1.2: American Community Survey.....	18
4.1.3: Transportation Network.....	19
4.2: Study areas selection.....	20

4.2.1: Grouping of metropolitan statistical areas	20
4.2.2: Identifying similar metropolitan statistical areas	23
4.3: Measuring proximity.....	25
4.4: Measuring access	29
4.5: Measuring socio-economic status	30
Chapter 5: Results	31
5.1: Computational performance	31
5.2: Types of results	32
5.2.1: Maps of IHE access	32
5.2.2: Tables of IHE access by demography and SES	33
5.3: Dalton, GA MSA	34
5.4: Joplin, MO MSA.....	44
5.5: Vineland – Bridgeton, NJ MSA.....	54
5.6: Tuscaloosa, AL MSA	64
5.7: Riverside – San Bernardino – Ontario, CA MSA.....	74
5.8: Sacramento – Roseville – Arden – Arcade, CA MSA.....	84
Chapter 6: Conclusions	94
6.1: IHE access trends.....	94
6.2: Limitations of analysis.....	95
6.3: Future research considerations.....	96
Bibliography	98

List of Figures

Figure 1: Pictorialized service areas and a census unit.....	12
Figure 2: Pictorialized service areas, a census unit, and sub census units.....	13
Figure 3: Pictorialized intersection of service areas and census sub-units.....	14
Figure 4: IHEs within the 2014 – 2015 College Scorecard dataset.....	18
Figure 5: Pseudo F-statistic by number of clusters.....	22
Figure 6: Effectiveness of ACS variables to discrimination between MSA clusters.....	22
Figure 7: Socio-economic clusters of metropolitan statistical areas.....	23
Figure 8: Locations of study areas.....	25
Figure 9: Service areas for Missouri Southern State University, Joplin, MO.....	26
Figure 10: Joplin, MO MSA census block groups & service areas for MSSU.....	27
Figure 11: Census block population of a census block group in western Joplin, MO.....	28
Figure 12: Computation time for service area generation per IHE subset.....	31
Figure 13: Computation time for calculating the population-adjusted average travel time from a census block group to an IHE.....	32
Figure 14: Median commute time to work for the census block groups of the Dalton, GA MSA.....	36
Figure 15: Number of accessible IHEs (all types) for the census block groups of the Dalton, GA MSA.....	37
Figure 16: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Dalton, GA MSA.....	38
Figure 17: Number of accessible IHEs (predominantly associate’s-degree granting) for the census block groups of the Dalton, GA MSA.....	39
Figure 18: Number of accessible IHEs (predominantly bachelor’s degree granting) for the census block groups of the Dalton, GA MSA.....	40
Figure 19: Number of accessible IHEs (public) for the census block groups of the Dalton, GA MSA.....	41
Figure 20: Number of accessible IHEs (private non-profit) for the census block groups of the Dalton, GA MSA.....	42

Figure 21: Number of accessible IHEs (private for-profit) for the census block groups of the Dalton, GA MSA	43
Figure 22: Median commute time to work for the census block groups of the Joplin, MO MSA.....	46
Figure 23: Number of accessible IHEs (all types) for the census block groups of the Joplin, MO MSA.....	47
Figure 24: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Joplin, MO MSA.....	48
Figure 25: Number of accessible IHEs (predominantly associate’s-degree granting) for the census block groups of the Joplin, MO MSA	49
Figure 26: Number of accessible IHEs (predominantly bachelor’s-degree granting) for the census block groups of the Joplin, MO MSA.....	50
Figure 27: Number of accessible IHEs (public) for the census block groups of the Joplin, MO MSA	51
Figure 28: Number of accessible IHEs (private non-profit) for the census block groups of the Joplin, MO MSA.....	52
Figure 29: Number of accessible IHEs (private for-profit) for the census block groups of the Joplin, MO MSA.....	53
Figure 30: Median commute time to work for the census block groups of the Vineland – Bridgeton, NJ MSA	56
Figure 31: Number of accessible IHEs (all types) for the census block groups of the Vineland – Bridgeton, NJ MSA.....	57
Figure 32: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Vineland – Bridgeton, NJ MSA	58
Figure 33: Number of accessible IHEs (predominantly associate’s-degree granting) for the census block groups of the Vineland – Bridgeton, NJ MSA	59
Figure 34: Number of accessible IHEs (predominantly bachelor’s-degree granting) for the census block groups of the Vineland – Bridgeton, NJ MSA	60
Figure 35: Number of accessible IHEs (public) for the census block groups of the Vineland – Bridgeton, NJ MSA.....	61
Figure 36: Number of accessible IHEs (private non-profit) for the census block groups of the Vineland – Bridgeton, NJ MSA.....	62

Figure 37: Number of accessible IHEs (private for-profit) for the census block groups of the Vineland – Bridgeton, NJ MSA.....	63
Figure 38: Median commute time to work for the census block groups of the Tuscaloosa, AL MSA.....	66
Figure 39: Number of accessible IHEs (all types) for the census block groups of the Tuscaloosa, AL MSA.....	67
Figure 40: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Tuscaloosa, AL MSA	68
Figure 41: Number of accessible IHEs (predominantly associate’s-degree granting) for the census block groups of the Tuscaloosa, AL MSA	69
Figure 42: Number of accessible IHEs (predominantly bachelor’s-degree granting) for the census block groups of the Tuscaloosa, AL MSA	70
Figure 43: Number of accessible IHEs (public) for the census block groups of the Tuscaloosa, AL MSA.....	71
Figure 44: Number of accessible IHEs (private non-profit) for the census block groups of the Tuscaloosa, AL MSA.....	72
Figure 45: Number of accessible IHEs (private for-profit) for the census block groups of the Tuscaloosa, AL MSA.....	73
Figure 46: Median commute time to work for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA	76
Figure 47: Number of accessible IHEs (all types) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA.....	77
Figure 48: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA	78
Figure 49: Number of accessible IHEs (predominantly associate’s-degree granting) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA	79
Figure 50: Number of accessible IHEs (predominantly bachelor’s-degree granting) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA	80
Figure 51: Number of accessible IHEs (public) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA.....	81
Figure 52: Number of accessible IHEs (private non-profit) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA.....	82

Figure 53: Number of accessible IHEs (private for-profit) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA..... 83

Figure 54: Median commute time to work for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA 86

Figure 55: Number of accessible IHEs (all types) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA 87

Figure 56: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA88

Figure 57: Number of accessible IHEs (predominantly associate’s-degree granting) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA..... 89

Figure 58: Number of accessible IHEs (predominantly bachelor’s-degree granting) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA90

Figure 59: Number of accessible IHEs (public) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA 91

Figure 60: Number of accessible IHEs (private non-profit) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA..... 92

Figure 61: Number of accessible IHEs (private for-profit) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA..... 93

List of Tables

Table 1: Socio-economic variation of MSA clusters	23
Table 2: Socio-economic similarity and education attainment dissimilarity of study areas	24
Table 3: Example from Tuscaloosa, AL of the count of any accessible IHE for different demographic and economic factors	33
Table 4: Median number of accessible IHEs (by type) for demographic and SES factors of the Dalton, GA MSA	35
Table 5: Average number of accessible IHEs (by type) for demographic and SES factors of the Dalton, GA MSA	35
Table 6: Median number of accessible IHEs (by type) for demographic and SES factors of the Joplin, MO MSA	45
Table 7: Average number of accessible IHEs (by type) for demographic and SES factors of the Joplin, MO MSA	45
Table 8: Median number of accessible IHEs (by type) for demographic and SES factors of the Vineland – Bridgeton, NJ MSA.....	55
Table 9: Average number of accessible IHEs (by type) for demographic and SES factors of the Vineland – Bridgeton, NJ MSA.....	55
Table 10: Median number of accessible IHEs (by type) for demographic and SES factors of the Tuscaloosa, AL MSA	65
Table 11: Average number of accessible IHEs (by type) for demographic and SES factors of the Tuscaloosa, AL MSA	65
Table 12: Median number of accessible IHEs (by type) for demographic and SES factors of the Riverside – San Bernardino – Ontario, CA MSA.....	75
Table 13: Average number of accessible IHEs (by type) for demographic and SES factors of the Riverside – San Bernardino – Ontario, CA MSA.....	75
Table 14: Median number of accessible IHEs (by type) for demographic and SES factors of the Sacramento – Roseville – Arden - Arcade, CA MSA	85
Table 15: Average number of accessible IHEs (by type) for demographic and SES factors of the Sacramento – Roseville – Arden - Arcade, CA MSA	85

List of Abbreviations

ACS	American Community Survey
CPU	Central Processing Unit
CSV	Comma-Separated Values
FIPS	Federal Information Processing Standard
GB	Gigabyte
GIS	Geographic Information System
IHE	Institute of Higher Education
IPEDS	Integrated Postsecondary Education Data System
MSA	Metropolitan Statistical Area
MMSU	Missouri Southern State University
NELS	National Education Longitudinal Study
NHGIS	National Historical Geographic Information System
NSLDS	National Student Loan Data System
OSM	OpenStreetMaps
PC	Personal Computer
SES	Socio-Economic Status
SQL	Structured Query Language
SSD	Solid-State Drive

GEOGRAPHIC INEQUITIES IN LOCAL HIGHER EDUCATION OPPORTUNITIES?
IDENTIFYING COLLEGE DESERTS WITH CRITICAL GIS

Christopher Krause

Dr. Timothy Matisziw, Thesis Supervisor

Abstract

In a world where one's future is heavily impacted by having postsecondary education, access to college is a pertinent research topic. Access is a widely researched topic, but only recently has college access been studied specifically. This study proposes a geographic information systems based methodology for quantifying college access at multiple spatial scales. This methodology was implemented with the Python programming language and ArcGIS. A sample of six metropolitan statistical areas were identified and analyzed using the developed methodology. Within this sample, college access varied primarily by socio-economic status although some variation between race/ethnicity was identified. Further research is needed to assess whether these trends are generalizable. Quantification of college access will aid policy-makers to prepare reforms to reduce the inequity of college access.

Chapter 1: Introduction

Worldwide, the number one predictor of “thriving” in life is having a “good job” (Clifton 2013, 167). Employee’s job prospects are increased dramatically from post-secondary education. Besides opening up access to higher paying occupations, higher-educated employees usually earn much more than their less-educated coworkers in the same occupation. Obtaining a two- or four-year degree provides a significant economic return over a high school degree: 32% and 74% increase in average yearly income respectively (Carnevale, Rose, and Cheah 2013). In fact, “the difference in earning between those who go to college and those who don’t is growing – meaning that postsecondary education is more important than ever” (2013, 20).

Besides the economic benefits of postsecondary education at the individual level, there are economic benefits for society at large. College educated individuals with higher incomes, pay more taxes, are less likely to draw upon social support programs, and are less likely to commit crimes, reducing additional spending on incarceration (Carroll and Erkut 2009). College educated individuals have been shown to be more active citizens, live healthier lifestyles, engage more in their children’s education, and more likely to move up the socioeconomic ladder (Ma, Pender, and Welch 2016). Therefore, having access to institutions of higher education (IHEs) seems crucial for the future success of this nation both economically and socially.

Does geographic proximity affect access to a college education? That is the central question of this study. This study proposes a new methodology for measuring geographic access to IHEs at a very local scale, the U.S. census block group reporting units, using geographic information systems (GIS). Census block groups can be enriched

with many others types of data including demographic and socio-economic. Using this enriched data, measures of college access can be determined not only for a given area but for additional factors such as race or socio-economic status (SES).

Past research on access has found many communities of lower socio-economic status (SES) lacking in access to sources of quality food (i.e. grocery stores) (Alwitt and Donley 1997). These areas of degraded access to food, dubbed “food deserts,” have inspired recent research on a dichotomy of “college deserts” and “college oases” (Dache-Gerbino 2016). Research on college deserts has typically been focused on identifying where college access is lacking and who may be impacted.

In a highly mobile society, it could be argued that college deserts are only a minor consideration because individuals desiring a postsecondary education could easily relocate if their local area lacked access to IHEs. However, Americans have been relocating at a historically low rate and Millennials, in particular, are thought to be less mobile than prior generations of young adults (US Census Bureau 2016; Fry 2017). This lack of mobility combined with the importance of local college access in the college choice process (Turley 2009), frames local access to IHEs as particularly relevant.

The stakes are high and controversial in this research due to the advent of for-profit colleges that have been accused of preying on low-income minority communities often leaving them with high debts and incomplete or unmarketable degrees (Lynch, Engle, and Cruz 2010). Thus, a clear and defensible research methodology is the first step in any analysis on this subject, one rife with public policy implications.

This study proposes an efficient methodology to investigate college access at a local scale for large geographic areas. To demonstrate this methodology, six

metropolitan statistical areas (MSAs) with relatively minimal residential geographic mobility were analyzed with respect to access to IHEs.

Chapter 2: Literature Review

2.1: Conceptual frameworks

Any quantification of college access necessitates some conceptually grounding. Conceptual frameworks provide the theoretical and practical justification for the proposed analysis. Understanding a prospective student's reasons for applying to and/or enrolling in an IHE is therefore central to measuring access to IHE.

Fortunately, the decision-making process of prospective college students is a widely researched topic. A highly cited and frequently used conceptual model in this respect was created by Hossler and Gallagher (1987). Their conceptual model involves a three-stage temporal process:

1. *The predisposition stage*: This stage occurs during primary and secondary education and relates to activities that influence a student's attitude towards higher education. Examples of these types of activities include coaching from guidance counselors, taking college preparatory classes, or studying for entrance exams.
2. *The search stage*: This stage typically occurs during the latter years of high school and involves the identification of prospective institutions, compilation of application requirements, and following through with the application process.
3. *The choice stage*: This stage typically occurs during the student's final semester of high school. Around this time, the student receives acceptance notices and can better assess the availability of financial aid. At this point, the student must make the decision to commit to a particular institution or perhaps hold off on attending an IHE.

2.2: Spatiality of conceptual frameworks

The largest weakness of many of the conceptual models investigated to date is their lack of consideration of spatial relationships that may influence access to IHE. The geographic context of prospective students is not considered despite research indicating that the ability to attend college close to home is a very important factor during the college choice process, particularly amongst minorities and those who are socioeconomically disadvantaged (Turley 2009).

However, each component of the previously identified conceptual frameworks *is* indeed heavily influenced by geographic relationships:

- *The predisposition stage:* A student living near a college may have an increased aspiration to attend. Turley found “the number of colleges within commuting distance is associated with higher odds of applying to college” although no significant difference was found between the number of colleges within commuting distance and the students’ choice in enroll in college (2009, 138).
- *The search stage:* The internet has made it much easier to find relevant information about colleges such as programs offered, admission requirements, and application procedures. However, research has shown students and their parents, particularly those from low-income communities, rely heavily on face-to-face communication to “interpret” this information (Brown, Wohn, and Ellison 2016, 115). Therefore, it is possible that students further from a college may be less capable of acquiring sufficient information about the colleges and apply/enroll.

- *The choice stage*: There comes a time when a prospective student has to make the final decision if/where they are going to attend college. As mentioned earlier, proximity is a major factor for many students (Turley 2009).

Despite living near an IHE, acceptance/enrollment may not be an option if it is too expensive or the admission requirements are too high relative to one's educational background. If a student's eligibility (financially and academically) precludes them from attending this nearby university, they may choose not to apply for college at all or attempt to commute a further distance to a more suitable college (Kipp, Wohlford, and Price 2002). However, it would be a gross oversimplification to suggest affordability is "a simple function of income alone...It is more likely to accrue through a combination of socio-economic processes" (Singleton 2012, 49) including at a minimum "financial capital, human capital, and social capital" (2012, 36).

2.3: Parallels to food desert literature

While there has been a considerable amount of research investigating college access, there has been a relatively less focus on investigating the geographic relationship between college proximity and college access. However, geographic access has been studied more extensively in other contexts, such as the access to food for example.

Access to food across the United States is very heterogeneous over multiple scales. For example, at the national level, low-income neighborhoods have 30% fewer grocery stores compared to the highest income neighborhoods. In Philadelphia, the disparity is much wider; the highest income neighborhoods were found to have 156% more grocery stores than the low-income neighborhoods (Weinberg 1995). Racial differences in access to supermarkets have been observed as well. For example, predominately Black neighborhoods have been found to have fewer supermarkets than

predominately White neighborhoods. Despite controlling for covariates, such as neighborhood income, Black neighborhoods were found to have only 52% as many chain supermarkets compared to similar White neighborhoods (Powell et al. 2007).

2.4: Access and accessibility

Various attempts to define “food desert” have led to debate about whether the phenomena being labelled “food desert” actually exists (Walker, Keane, and Burke 2010). Consider these proposed definitions: “urban areas with 10 or fewer stores and no stores more than 20 employees” (Hendrickson, Smith, and Eikenberry 2006) and “poor urban areas, where residents cannot buy affordable, healthy food” (Cummins and Macintyre 2002). Assuming data is available or can be collected, the former definition is easily quantified. The latter definition lacks an operational definition making it very difficult to quantify despite arguably being a *better* definition.

This highlights the fundamental dilemma of access research: how does one define access or accessibility? “Accessibility...is a slippery notion...one of those common terms that everyone uses until faced with the problem of defining and measuring it” (Gould 1969, 64). Accessibility is not a physical object which can be measured directly, rather it is a concept or perception which can be evaluated or judged differently depending on the scope and context of the research (Scott 2000). Although widely used interchangeably, in transportation geography these terms cannot be used interchangeably: “access generally refers to the opportunity (and level of effort) associated with entering a transportation system” whereas accessibility is the extent to which destinations can be reached with a combination of transportation modes (Matisziw and Grubestic 2010, 712).

2.5: GIS as a methodology

A literature review of 31 peer-reviewed journal articles researching food deserts in the United States found that GIS-based methods to be the most utilized analysis approach (Walker, Keane, and Burke 2010). A literature review of 29 peer-reviewed journal articles specifically using GIS-based measures of the food environment found two common approaches to measuring access: density and proximity. Studies using a density approach measured the number of food outlets in a defined zone (for example, a zip code). Studies using a proximity approach measured the distance between a consumer and the closest food outlet using Euclidean or network distance. Several studies combined aspects of both density and proximity. Several studies even considered travel time for different transportation modes (Charreire et al. 2010).

2.6: Critique of existing college access and college proximity literature

Methodologically, many studies analyzing access to IHEs are rooted in data collected through longitudinal studies. In these types of studies, demographic and geographic information regarding the participants at the end of high school can be obtained. The decision of the individual to attend (or not attend) an IHE is also typically noted. Provided the home location and college location of individuals is available, measuring how far students travelled to college is then nothing more than a trivial calculation (Turley 2009; Gibbons and Vignoles 2012). The implications of such research are far from trivial, however:

The aggregate behaviors of people as recorded in empirical analysis of large datasets are related to the differentiation in common attitudes that exist within socio-spatial classification groups...For educational markets, this is a highly relevant topic, as it relates directly to our ability to identify those groups in society that may miss out on certain life chances through a restricted ability to compete for the advantages sustained from a quality education. (Singleton 2012, 39)

Higher education is a rapidly evolving field. For example, today's students are more likely to start college with credits earned during high school and online classes are becoming more prevalent. For these reasons and more, the relevance of these longitudinal studies diminishes because they are so outdated. For example, Turley (2009) used the National Education Longitudinal Study (NELS) of 1988. In her study, participants were assessed from eighth grade to the point where they were making decisions to attend IHE in the early 1990s, almost two decades prior to publication of her study.

Another potential methodological limitation of the Turley (2009) study relates to how distances were calculated. For each participant in the NELS, the zip code of their home and college were recorded. The Euclidean (straight-line) distance was then computed from the centroid of their home zip code to the centroid of their college zip code (Turley 2009). This measurement technique does not account for the road network, therefore it may or may not accurately reflect the commuting distance or travel time. Within urban areas, Euclidean distance highly correlates with commuting distance. Particularly in rural areas, zip codes cover a greater geographic extent, so the centroid of the zip code may not be a good approximation of the student's home location or the college's location (Jones et al. 2010).

Unlike many European studies, studies of college access in the United States must account for many vast rural areas. For example, in England approximately 1.9% of the population lives further than 80 km from a college (Gibbons and Vignoles 2012). In Canada, that same statistic is 20.1% (Frenette 2006). One study even omitted the entire state of Wyoming since it has only one college (Hillman and Weichman 2016). Granted

the population of Wyoming is small, but there are still individuals without close geographic access to IHEs who are not be represented by their findings.

Another issue with the existing literature is the lack of spatial resolution. Hillman and Weichman, for example, defined college deserts as “places with either of the following two conditions: 1. Zero colleges or universities are located nearby, or 2. One community college is the only public broad-access institution nearby” (2016, 4). At face value, this definition may seem sufficient, but the definition of “nearby” is defined as these intuitions being located within a commuting zones, micropolitan statistical areas, and metropolitan statistical areas. They argue that these aggregations of counties are a sufficient definition of “local” (2016), however in cases of limited mobility looking at a more local neighborhood level is likely more appropriate (Walker, Keane, and Burke 2010; Dache-Gerbino 2016). Limited mobility can be caused by a myriad of factors such as: not owning a personal vehicle, living somewhere without public transportation, or having other commitments during the day such as working or caring for family.

From a fundamental perspective, college deserts can be assessed primarily in two ways: looking at the service area of the institutes (Hillman and Weichman 2016) or looking at the areas in the most need (Dache-Gerbino 2016). Without delving into the details about how the methodology is different between these approaches, it can be fairly said that they inherently are trying to answer different questions. The former questions how optimally colleges are located under the assumption that college proximity is a highly relevant factor in a student’s decision to pursue college. The latter asserts that certain groups of people, such as minorities, are underserved by colleges and attempt to describe the spatial injustice in comparison to groups better served by colleges. The

service area methodology can be used to delineate spatial injustices by analyzing the individuals within the service area of a college and comparing those individuals to those not within the service area of a college. This can be done without first identifying particular groups that the researcher subjectively believes may be marginalized.

Determining which variables to include in such a model to delineate college deserts is inherently subjective and introduces uncertainty. However, if several models are created each using different parameters, the results can be compared through a sensitivity analysis (Chen et al. 2009). The validity of the model can be determined by the consistency of model output with various parameters. If models do not agree, that indicates greater uncertainty in the validity of the model's parameters.

Although there is no established methodology for delineating college deserts, by working within a conceptual framework and adapting/improving upon the methodologies of others, this thesis project strives to propose an adaptable methodological framework which can be used as a decision support system for policymakers.

Chapter 3: Methodology

3.1: Measuring proximity

Measuring proximity from a point to another point is typically straightforward. For example, with a home address and the address of an IHE, once geocoded, the Euclidean distance could be measured or the driving time between them could be computed. Euclidean distance, commonly referred to as “the distance as the bird flies” is not always directly relatable to travel time (Boscoe, Henry, and Zdeb 2012) because no regard to availability of roads or hierarchy of available roads is considered with this methodology. However, rather than computing the proximity from a point to another point, in this study, the proximity from a polygon to a point needs to be computed.

The simplest measure of proximity between a polygon and a point would be the Euclidean distance. Commonly, the Euclidean distance is either measured from the point to the nearest vertex on the polygon or the centroid of the polygon. Not only do the same methodological considerations about using Euclidean distance as a measure of proximity apply here, but the spatial extent of the polygon must be reduced to a single point. To compute driving times, most GIS software either requires the user to convert the polygon into a point or the GIS software converts the polygon into a point (or sometimes a set of points) and computes the driving time between the polygon's point(s) and the point of interest. These computations can vary substantially based on the point(s) selected. For example, consider a point being selected where there are no nearby road segments: Should the point be linked to whatever road segment is the closest? Should the point only be associated with a road(s) within some maximum distance? What if the nearest road segment is not logical (i.e. one-way, dead-end, etc.) or what if multiple nearest road options exists?

An alternative method for assessing proximity is to identify all areas within some service time or distance of a facility. In the context of a GIS, buffering of a point facility by some distance to generate a polygonal service area is one example. Provided that movement over the landscape is limited to a network, similar transformations from point to polygon are available. For instance, in ArcGIS, a point location in a network can be transformed into a polygon, termed a service area. Service areas can be generated for a single time/distance threshold or a series of thresholds, each of which indicates the travel time back to the facility. A simplistic representation of service areas with a single census unit superimposed over them is shown in Figure 1.

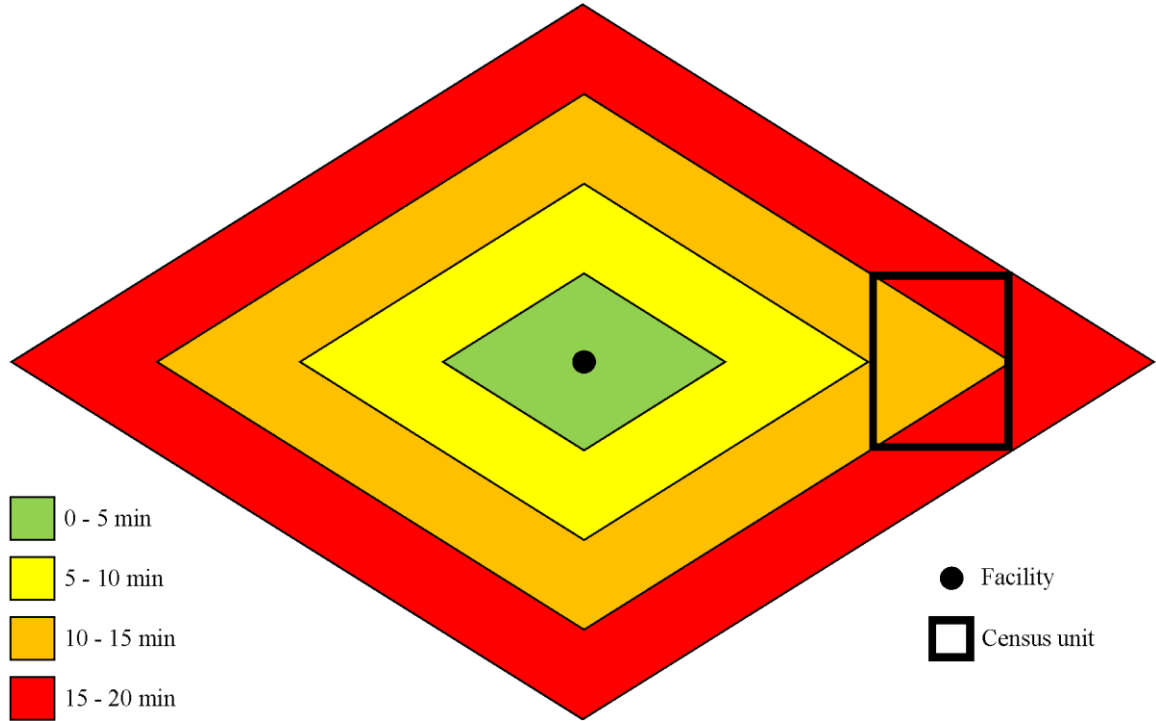


Figure 1: Pictorialized service areas and a census unit

For the census unit shown in Figure 1, the average travel time can be computed with the relative area of the census unit within each of the service areas according to the equation below (where TT_a is the average travel time, a_n is the percentage of the census unit's area covered by the service area, and mp_n is the midpoint of the service area):

$$TT_a = a_1mp_1 + \dots + a_nmp_n$$

For the census unit shown in Figure 1, the average weighted travel time is:

$$TT_a = \left(0.5 \cdot \frac{10+15}{2}\right) + \left(0.5 \cdot \frac{15+20}{2}\right) = 15 \text{ minutes}$$

If the population within the census unit was uniformly distributed, this travel time estimate would be appropriate. If the population within the census unit is not uniformly distributed, this travel time estimate would be over- or under- estimated. If the population is known for a smaller census unit, for example in Figure 2, a population-adjusted average travel time can be computed with the relative area of each census sub-

unit within each of the service areas and the relative population of each census sub-unit according to the equation below (where TT_{paa} is the population-adjusted average travel time for the census unit, TT_{a_n} is the average travel time for the census sub-unit, and p_n is the percentage of the census unit's population within the census sub-unit):

$$TT_{paa} = TT_{a_1}p_1 + \dots + TT_{a_n}p_n$$

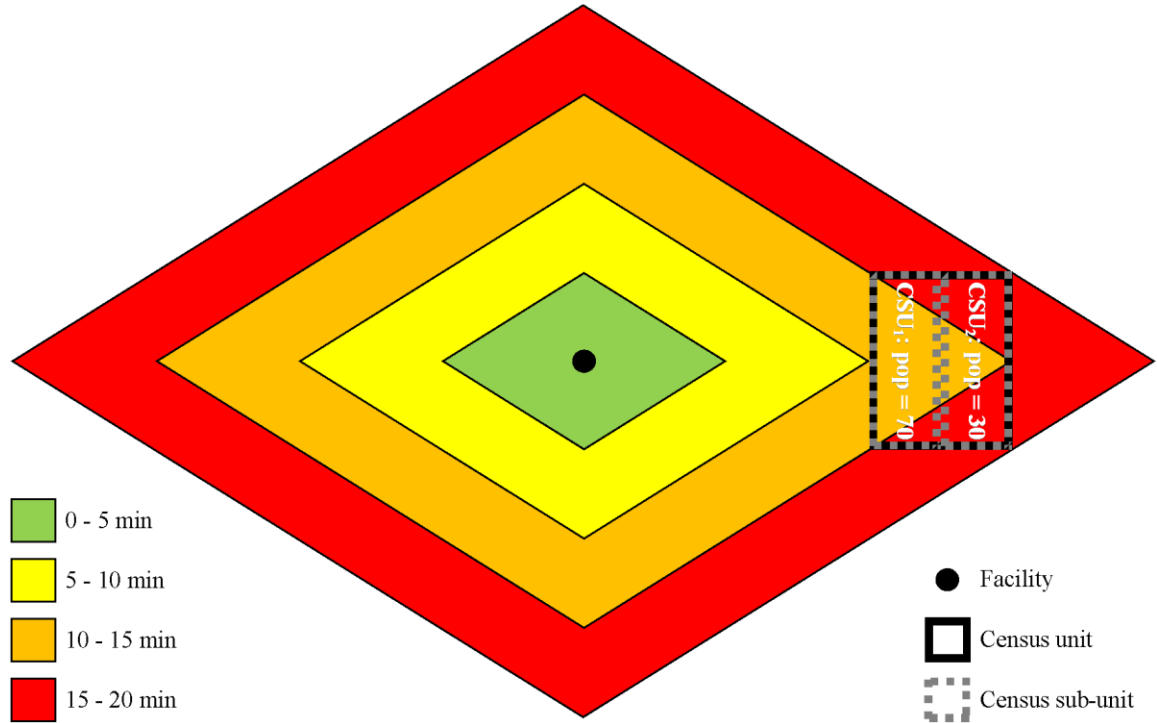


Figure 2: Pictorialized service areas, a census unit, and sub census units

Calculation of the average weighted travel for the census sub-units requires computing the intersection of the census sub-units and the service areas. For the two census sub-units shown in Figure 2, this intersection results in six polygons (A, B, C, D, E, and F) as shown in Figure 3. Census sub-unit 1 (CSU_1) is composed of polygons A, B, and C. Census sub-unit 2 (CSU_2) is composed of polygons D, E, and F. The average travel time for CSU_1 and CSU_2 are computed using the polygons from the intersection:

$$CSU_1: TT_a = \left(0.125 \cdot \frac{15+20}{2}\right)_A + \left(0.75 \cdot \frac{10+15}{2}\right)_B + \left(0.125 \cdot \frac{15+20}{2}\right)_C = 13.75 \text{ minutes}$$

$$\text{CSU}_2: TT_a = \left(0.375 \cdot \frac{15+20}{2}\right)_D + \left(0.25 \cdot \frac{10+15}{2}\right)_E + \left(0.375 \cdot \frac{15+20}{2}\right)_F = 16.25 \text{ minutes}$$

With the average weighted travel computed for each census sub-unit, the population-adjusted average travel time is:

$$TT_{paa} = 13.75 \cdot \frac{70}{100} + 16.25 \cdot \frac{30}{100} = 14.5 \text{ minutes}$$

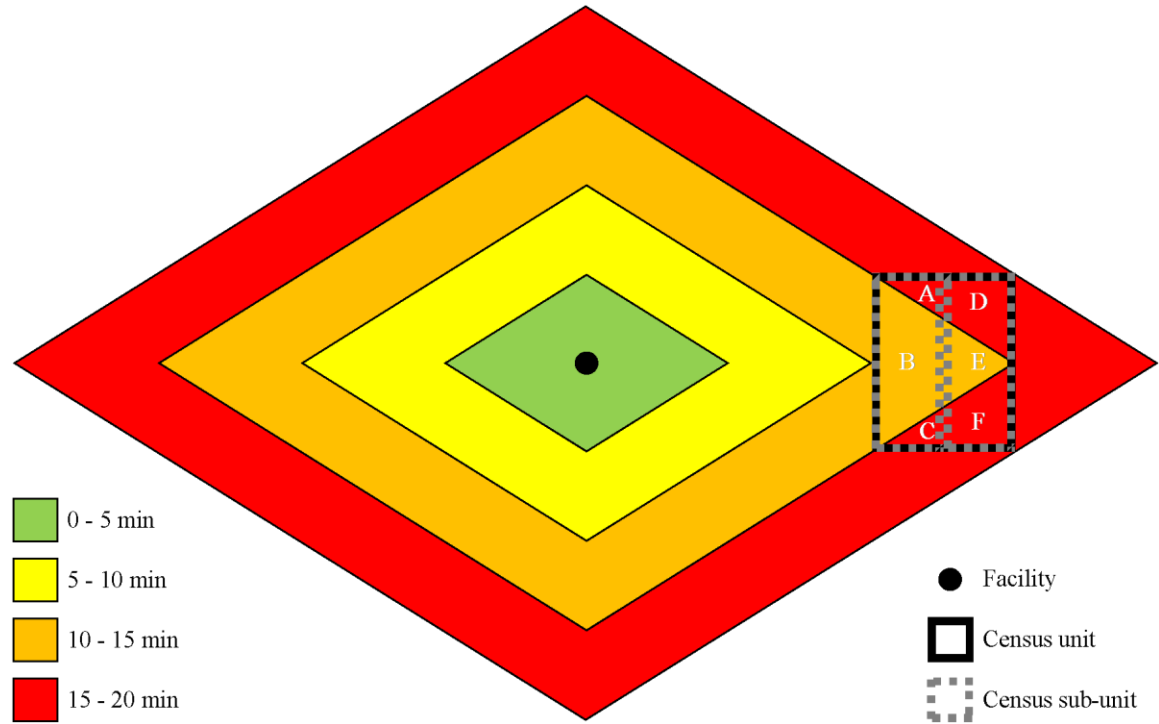


Figure 3: Pictorialized intersection of service areas and census sub-units

While the difference between the computed TT_a and TT_{paa} may seem negligible (particularly in this example), as the census units increase in area, the population of census sub units become less uniformly distributed, and/or the service area interval increases, the difference between TT_a and TT_{paa} increases.

3.2: Measuring access

With the population-adjusted average travel time from a census unit to a facility computed, it must be determined if the travel time is within a commutable distance for

the average person within the census unit. One could define a travel time threshold, but how would such a threshold be decided upon?

Many studies “set somewhat arbitrary cutoff points as threshold of access beyond which one can assume that individual are not meeting their needs” (Rosero-Bixby 2004, 1276). These thresholds assume equal willingness and ability to commute between every commuter. However, this assumption is not realistic given that commuting times have been shown to vary considerably, regionally and demographically (McKenzie and Rapino 2011).

Acknowledging the variability of commuting times, the threshold separating accessible from inaccessible could be set per census unit using commuting data collected within the census unit. The American Community Survey asks respondents the number of minutes they spend commuting to work. Although the respondent writes-in their answer as an integer number, the U.S. Census Bureau reports this data as the number of people whose commute time falls within specified time intervals (for example, less than 5 minutes, 5 – 9 minutes, etc.).

This reporting format requires additional processing to be used as a threshold distinguishing accessible facilities from inaccessible ones. First, each time interval must be converted into a single number by calculating its midpoint. Then, using the relative number of people within each commuting time interval, the weighted average of commute time could be computed. However calculating the median commute time would be a more appropriate measure because, for most census units, the distribution of commute times was highly skewed to the right and the data is recorded in an ordinal scale rather than continuous (Manikandan 2011). The large skew to the high (caused by those

with abnormally long commutes) would cause the weighted average to be artificially inflated. Compared to the weighted average, the median commute time is more representative of the census unit populations' ability and willingness to commute. Therefore, the median commute time per census unit can be used a threshold to differentiate accessible facilities from inaccessible facilities.

The commuting times reported by the US Census Bureau include all modes of transportation (driving, public transportation, walking, biking, etc.). If the service areas generated to calculate proximity include only a single mode of transportation (for example, driving), using the median commute time as a threshold may not be appropriate. However, if the census unit's predominant mode of transportation is the same mode of transportation for which the service areas were generated, this potential issue is diminished. It is possible to determine if this is the case since the US Census Bureau reports the number of commuters per census unit who use each transportation mode.

Chapter 4: Application of Methodology

At the most basic level, application of the methods described in Chapter 3 for college access requires the location of IHEs, the location of the people who could potentially be served by these IHEs, and a road network once which people would to commute on to the IHEs. In the next section, the collection and processing of these datasets is detailed. In this study, the objective was to assess whether access to IHEs affects higher education obtainment. To analyze this potential relationship, socially and economically similar MSAs with differing levels of higher education attainment were identified. Following this, the methodology described in the previous chapter is used to measure access to IHEs in the identified study areas.

4.1: Data Sources

4.1.1: College Scorecard

In the past, gathering information about IHEs was not straightforward, particularly for the average college-destined high schooler. To simplify the information-gathering for these prospective students, in September 2015 the Obama administration announced College Scorecard. This website was intended to help students find a suitable college by providing them with information about predominantly awarded degree, institutional control, program availability, graduation rates, student-body composition, average financial aid, student loan default rates, etc. (Office of the Press Secretary 2015).

The data behind this website is publically available for download in comma-separated values (CSV) format. In total, the College Scorecard covers over 7,000 IHEs (all those receiving any federal funding) with nearly 2,000 attributes for each. These attributes record the same kind of information (listed above) as the College Scorecard website but with greater specificity. The location of each college is known via latitude and longitude attributes allowing each IHE to be represented in a GIS (Figure 4).

Many of these attributes are from the Integrated Postsecondary Education Data System (IPEDS), but additional attributes originate from the National Student Loan Data System (NSLDS) and from the Department of Treasury to provide data about student financial outcomes after attending an IHEs. The College Scorecard has many missing values (due to issues with reporting and/or privacy concerns), but the most recent dataset, the 2014 – 2015 academic school year, is mostly complete and was used in this study.

Some IHEs were excluded: IHEs outside of the conterminous United States, IHEs without their predominant degree awarded or institutional control classified, IHEs

classified as predominantly awarding graduate degrees, IHEs operating online-only, and IHEs not currently open. In the end, a total of 6,032 IHEs were retained for analysis.

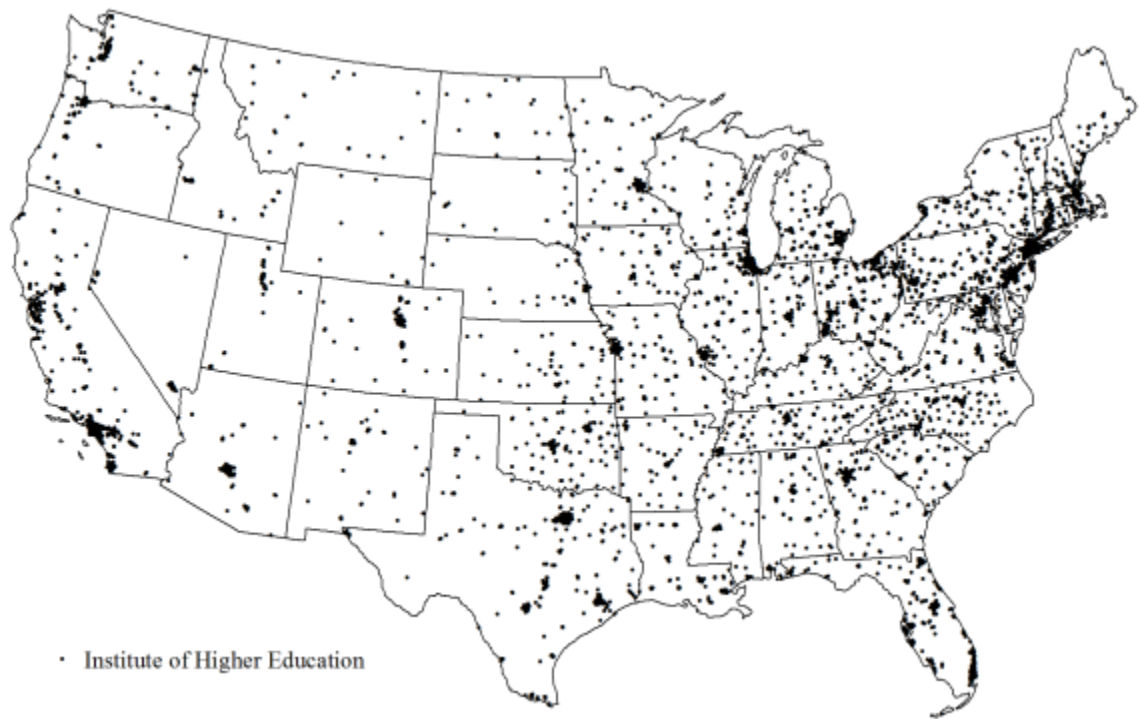


Figure 4: IHEs within the 2014 – 2015 College Scorecard dataset

4.1.2: American Community Survey

The United States Census Bureau has conducted a decennial census since 1790, the last one being in 2010. To best ensure temporal consistency the IHE data from the 2014 – 2015 academic school year, a more recent census data product was used: the 2011 – 2015 American Community Survey (ACS). Unlike the decennial census where the entire population is surveyed, the ACS samples housing unit addresses monthly. Over the course of a year, about 2 million addresses will have been sampled. ACS data products are released annually, with estimates in rolling 1-, 3-, or 5-year samples. Although the margin of errors are higher than the decennial census (Citro and Kalton 2007, 1–2), not using the 2011 – 2015 ACS estimates would mean using data collected during the 2010 decennial census (five years prior to the IHE data). If the population and

demographics were stable during this time period, using the decennial census may have been appropriate, but lacking confirmation of this stability, the uncertainty introduced by using the 2011 – 2015 ACS was deemed acceptable.

Although the data used in this study was originally collected by the U.S. Census Bureau, the National Historical Geographic Information System (NHGIS) provides a more accessible interface for downloading and using the data. For example, rather than downloading an entire table from the ACS, using the NHGIS, the user can choose to download only a subset of variables from a table or create a dataset which combines variables from multiple tables. To simplify the process of joining the tabular CSV data into the attribute table of the shapefile geometry, the NHGIS has created a text-based GISJOIN field to help avoid potential issues with mismatching variable types (integer vs string) caused by leading zeros in the Federal Information Processing Standard (FIPS) code of each census unit.

4.1.3: Transportation Network

As described in Chapter 3, computing travel time requires a network dataset composed of the arcs and the nodes of the physical road network. Such datasets can be created from freely available datasets such as OpenStreetMaps (OSM), however the quality of these road networks is not nearly as high as commercially made products. Network datasets created from OSM must be small in spatial extent due to OSM restrictions of 50,000 nodes per download (White 2013). For these reasons, this study utilized Esri's StreetMap Premium for ArcGIS. This commercially produced and maintained network dataset covers all of North America and is optimized for network analysis (by establishing network hierarchy).

4.2: Study areas selection

College access (or conversely, lack of college access) can vary over different spatial scales. Previous research has found that college access differs state-to-state (Kipp, Wohlford, and Price 2002) and within a single metro area where college access differed in the urban core compared to the suburbs (Dache-Gerbino 2016). This study analyzed college access at two relatively local spatial scales: the metropolitan statistical area (MSA) and the block group. Block groups are the smallest enumeration unit of ACS data enabling access to be measured at a very local level.

Varying levels of access at the block group spatial scale might be reflected at the larger MSA spatial scale. The methodology of this study was designed to test this hypothesis: Do MSAs with greater access to IHEs have more college educated citizens? In the most basic sense, access to IHEs for MSAs with low rates of college educated citizens needs to be compared to that of MSAs with high rates of college educated citizens. However, a more nuanced approach was required because the MSAs with the lowest rates of college educated citizens (Dalton, GA and Houma-Thibodaux, LA) are not comparable socio-economically to the MSAs with the highest rates of college educated citizens (Boulder, CO and Ithaca, NY). For a more reasonable comparison, socio-economically comparable MSAs with different rates of college educated citizens were identified.

4.2.1: Grouping of metropolitan statistical areas

To find socially and economically similar MSAs of relatively similar size, ten socio-economic variables (shown in Table 1) were appended to the MSA feature class. This variables were selected based on precedent set in past literature (Dache-Gerbino 2016). With these additional attributes, ArcGIS's Grouping Analysis could be used to

group the MSAs. This tool can consider spatial constraints, but since it is possible that socially and economically similar MSAs may not be spatially located near each other, the Grouping Analysis tool was configured not to consider any spatial constraints and use the ten socio-economic variables as the sole basis for the grouping. When spatial constraints are ignored, the Grouping Analysis uses the *k*-means clustering technique.

Uncertain of the number of clusters that would be appropriate to identify socio-economically similar MSAs, the Grouping Analysis tool was configured to algorithmically find the optimal number of clusters. The algorithm tests all the possible number of clusters between two to fifteen. For each number of clusters, the algorithm clusters the data ten times. Since the *k*-means clustering algorithm is dependent on the initial seeds (chosen randomly), the results varied slightly between these ten runs. To measure the efficiency of using this number of clusters, for each run of the *k*-means clustering algorithm, a pseudo F-statistic was computed. Higher pseudo F-statistics indicate more similarity within clusters and more dissimilarity between clusters (Esri 2017a). Using these socio-economic variables, the highest pseudo F-statistic and narrowest range of pseudo F-statistics was obtained when using three clusters (Figure 5). The Grouping Analysis was then performed accordingly (Figure 7).

The Grouping Analysis tool computes an R^2 value for each attribute indicating how effective that attribute is for discriminating between clusters (Figure 6). Higher R^2 values indicate more discrimination between clusters (Esri 2017a).

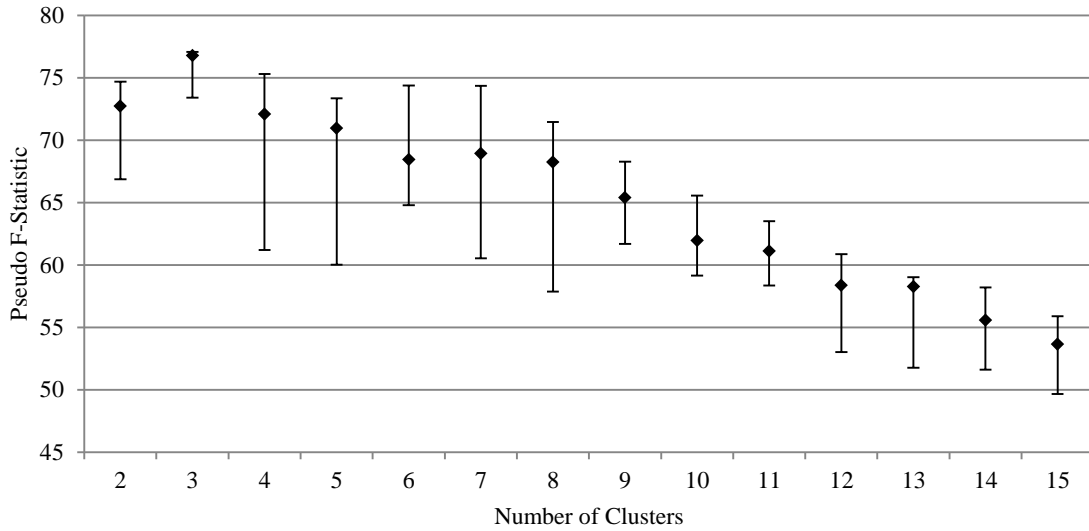


Figure 5: Pseudo F-statistic by number of clusters

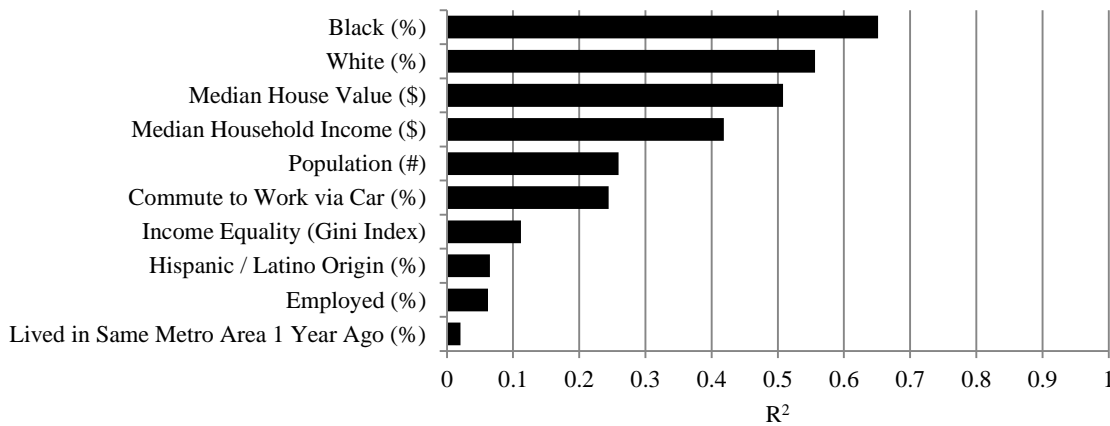


Figure 6: Effectiveness of ACS variables to discrimination between MSA clusters

The Grouping Analysis tool does not report what makes the clusters internally similar and likewise does not report what makes the clusters externally dissimilar. However, by analyzing the ACS variables of these distributions (Table 1), it becomes apparent what distinguishes these clusters: Cluster A is comprised of relatively wealthy metro areas with large racially diverse populations. Cluster B is comprised of relatively small predominantly middle-class White metro areas. Cluster C is comprised of

relatively poorer metros areas with large Black populations. Despite the variation within each cluster (Table 1), these distinctions are clear from the mean of these ACS variables.

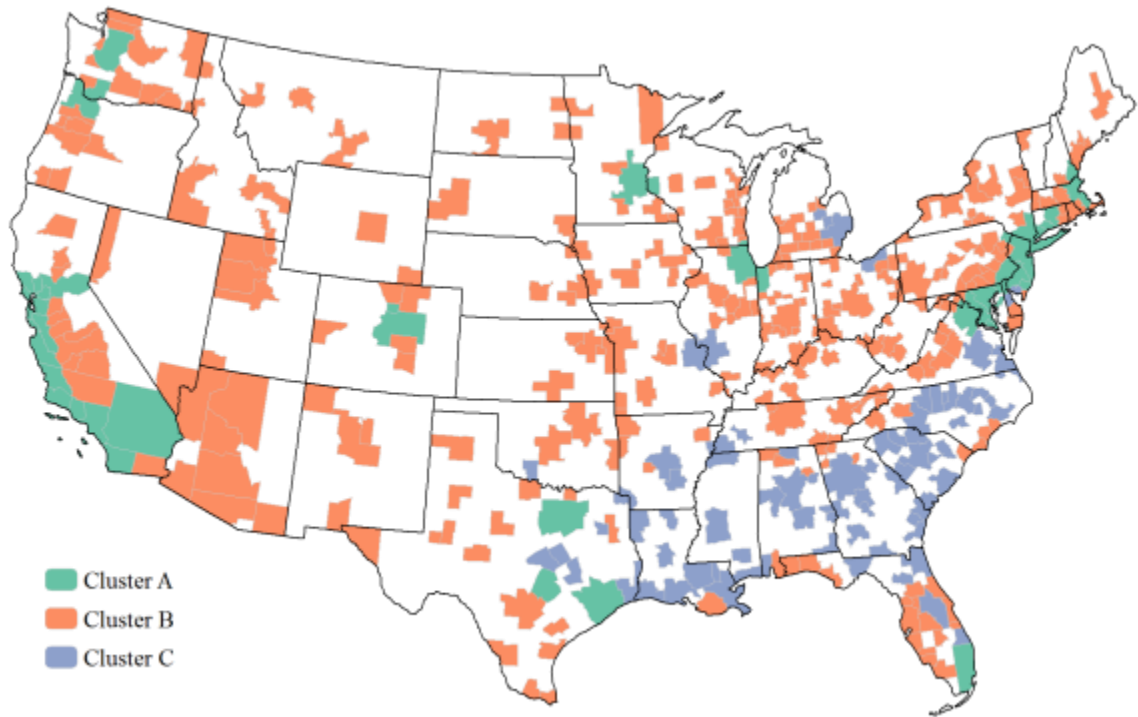


Figure 7: Socio-economic clusters of metropolitan statistical areas

ACS Variable	Cluster A		Cluster B		Cluster C	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Population (#)	3,167,670	4,063,947	379,791	497,465	649,102	928,650
White (%)	68.29	13.50	85.25	6.91	64.82	8.74
Black (%)	9.40	7.56	5.84	4.59	28.22	9.51
Hispanic / Latino Origin (%)	23.12	13.48	12.84	16.84	7.58	5.94
Employed (%)	66.52	3.11	61.98	6.10	61.56	3.53
Median Household Income (\$)	68,253	10,171	49,767	6,794	46,170	5,772
Income Equality (Gini Index)	0.4642	0.0273	0.4495	0.0217	0.4679	0.0218
Median House Value (\$)	350,884	133,843	157,160	46,023	140,064	32,483
Lived in Same Metro Area 1 Year Ago (%)	94.90	1.74	93.76	2.44	93.61	2.82
Commute to Work via Car (%)	87.76	6.90	93.85	3.19	95.34	1.92

Table 1: Socio-economic variation of MSA clusters

4.2.2: Identifying similar metropolitan statistical areas

Within each of the identified clusters, the metro area with the lowest rate of college attainment (bachelor’s degree or higher) was determined: Within Cluster A, the least college educated metro area was Riverside-San Bernardino-Ontario, CA (27.87%). Within Cluster B, the least college educated metro area was Dalton, GA (18.02%).

Within Cluster C, the least college educated metro area was Vineland-Bridgeton, NJ (20.74%).

For each of these metro areas, the most socio-economically similar metro area within the same cluster was identified using ArcGIS’s Similarity Search tool. This tool uses the attributes of a single feature and compares those attributes to all other features in the dataset. A measure of similarity is computed and these similarity scores are ranked from most similar to least similar (Esri 2017b). The variables used in the Similarity Search are shown in Table 2 with the exception that the college educated rate was excluded from the Similarity Search. Rates of college education were excluded because the purpose of finding these pairs of similar metro areas is to determine whether or not the degree of college access is correlated with the college educated rate. The relatively large variation in the college educated rate between these pairs of otherwise socio-economically similar MSAs is suitable for such exploration.

ACS Variable	Cluster A		Cluster B		Cluster C	
	Dalton, GA	Joplin, MO	Vineland-Bridgeton, NJ	Tuscaloosa, AL	Riverside-San Bernardino-Ontario, CA	Sacramento-Roseville-Arcade, CA
Population (#)	142,857	175,961	157,035	235,570	4,392,801	2,221,525
White (%)	90.19	91.12	63.05	63.61	63.44	66.38
Black (%)	3.14	1.82	20.47	33.37	7.35	7.13
Hispanic / Latino Origin (%)	27.78	6.57	28.64	2.95	48.94	20.85
Employed (%)	63.23	63.7	57.47	56.98	60.49	62.18
Median Household Income (\$)	39,459	42,255	49,984	43,697	55,092	60,003
Income Equality (Gini Index)	0.4554	0.4531	0.4404	0.4731	0.4446	0.4585
Median House Value (\$)	111,700	108,000	162,400	146,000	245,900	286,500
Lived in Same Metro Area 1 Year Ago (%)	96.39	95.46	95.18	92.93	95.22	95.39
Commute to Work via Car (%)	96.74	96.79	94.11	97.29	95.23	91.74
College Educated (%)	18.02	45.31	20.74	47.31	27.87	41.02

Table 2: Socio-economic similarity and education attainment dissimilarity of study areas

The slight variation between each pair of MSAs (Table 2) is acceptable because the Similarity Search tool proves these pairs are more similar (with respect to this subset

of ACS variables) than any other potential pair. The locations of these MSA pairs are shown in Figure 8.



Figure 8: Locations of study areas

4.3: Measuring proximity

As described earlier in the Methodology section, the first step to measure proximity was generating service areas. ArcGIS has tools built-in to create service areas, but the tools were not designed to be used with large datasets such as the one used in this dataset. As such, routines were programmed in Python (using arcpy to interface with ArcGIS) to extend the capabilities of the built-in service areas generation tools to accommodate large datasets. All of the service areas generated used a 1-minute time interval up to a maximum of 135 minutes. These service areas represent the amount of driving time from a location to the IHE. Over 91% of commuters in the MSAs studied commute to work via car, therefore using drive time as a measure of proximity was deemed appropriate. To decrease computation time, ArcGIS was configured to generate

generalized service areas instead of detailed ones. In some cases, the service areas generated by ArcGIS were not correctly formatted. Rather than the service area polygons being a ring of the area accessible within that given time interval, some service area polygons were a disk of the area accessible within that given time interval or less.

Additional Python routines were coded to identify and corrected these erroneous disks.

An example of the output from these Python routines is shown below in Figure 9.

These are the service areas for Missouri Southern State University (MSSU) in Joplin, MO (Figure 9). MSSU is located near the center of these service areas, but the shape of the service areas are not radially symmetrical (like a geometric buffer would be) because these drive time distance are dependent on the underlying road network. The service areas are elongated where there are major highways where travel is faster. This is especially evident near the edges of the service areas where the service areas end along major highways.

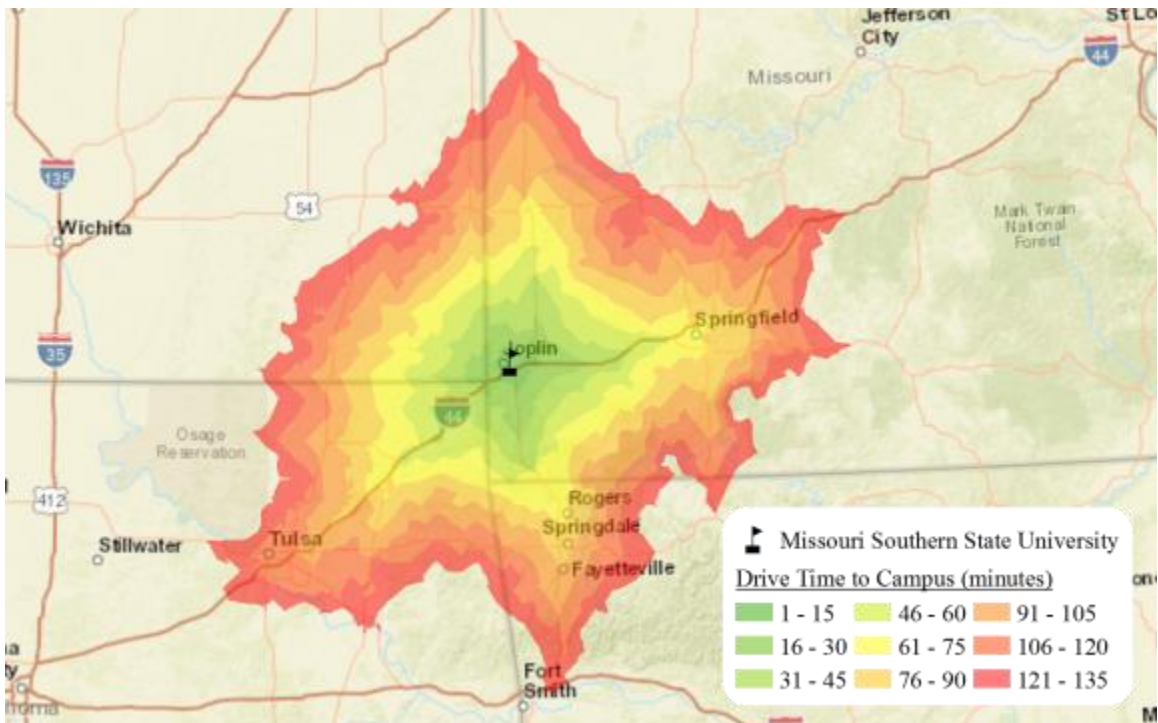


Figure 9: Service areas for Missouri Southern State University, Joplin, MO

As described Chapter 3, using the service areas, the population adjusted average driving time to a census block group can be computed. The census block groups of the Joplin, MO MSA are shown in Figure 10 superimposed on MSSU’s service areas.

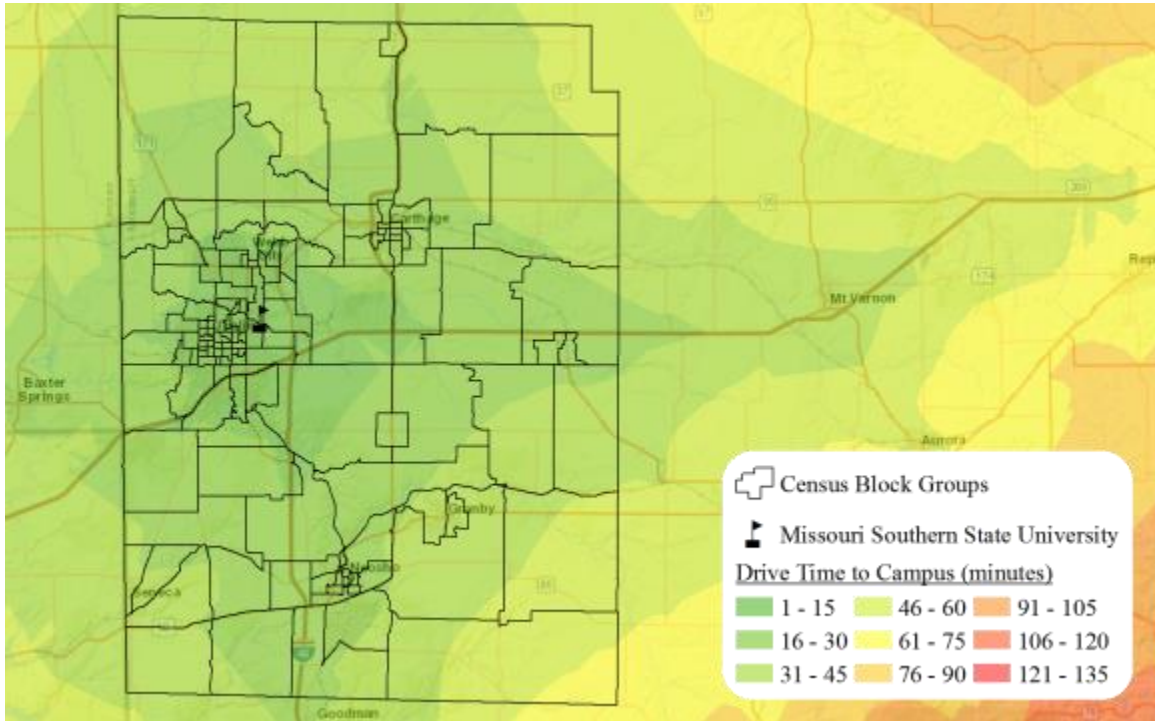


Figure 10: Joplin, MO MSA census block groups & service areas for MSSU

Each census block group is composed of one or more census blocks. Since the population of each census block is known, the population-adjusted average travel time can be computed as described in Chapter 3. This additional step alleviates the underlying assumption that population within a census block group is uniformly distributed. This assumption is still assumed for the census blocks. Figure 11 shows an example where the census block group’s population is not uniformly distributed. This census block group, located on the western fringe of Joplin, MO, has its population clustered on the east side closest to the urban core. The majority of the population lives less than 17 minutes from MSSU, but the at least half of the census block group by area is situated more than 17 minutes from MSSU. If the average travel time was based on the areal coverage of the

service areas, the average travel time would be overestimated. To avoid this, the travel time was computed as the population-adjusted average travel time described in Chapter 3.

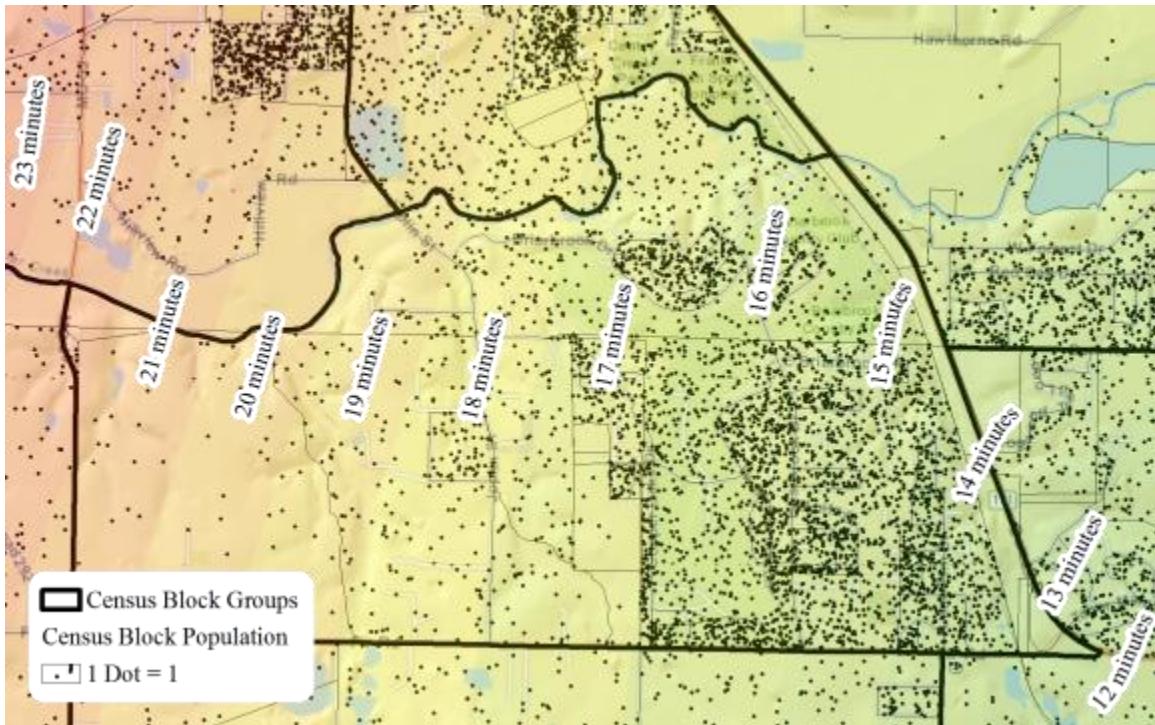


Figure 11: Census block population of a census block group in western Joplin, MO

Neither ArcGIS nor other commercial GIS software has this algorithm for population-adjusted driving time analysis implemented. Each of these methodological steps was implemented in Python using arcpy to integrate ArcGIS and its functionalities to Python and its functionalities. In some cases, neither Python nor ArcGIS functionalities were sufficient necessitating the implementation of other software-packages. For example, computing weighted averages in ArcGIS or in Python was not easily implemented or efficient. However such a task is readily accomplished using Structured Query Language (SQL). Unfortunately, ArcGIS's proprietary file geodatabase format does not support the SQL functions necessary for computing weighted averages. Python has native support for the SQLite database engine so the code was written to extract the information from the ArcGIS file geodatabase, insert it into a

temporary SQLite database, compute the weighted averages, and return the output back into the ArcGIS file geodatabase. These steps, despite sounding unnecessarily complicated, proved to be over 50 times faster than using tools built-in to ArcGIS.

Multiprocessing was also used to increase the efficiency of the developed methodology. Most modern computers have more than one core in their central processing unit (CPU) but software must be written specifically to use multiple cores. The computational burden was relatively simple to distribute across multiple cores because computing each IHE is independent from each other. Therefore a group of IHEs were assigned to each of the computer's core and were run simultaneously to speed up computation.

4.4: Measuring access

Commuting time varies between census block groups. For example rural communities on average are willing to commute further than urban communities (Turley 2009). Therefore, if the travel time from the nearby IHE exceeds the median travel time reported for the census block group, the IHE was concerned inaccessible. If the travel time from an IHE was less than or equal to the median travel time for the census block group, the IHE was concerned accessible.

Once again, commercial GIS software lacked the ability to complete the methodology described above so it was implemented in Python using ArcGIS's arcpy module. Once the determination had been made if an IHE was accessible to a census block group, the number of accessible IHEs per census block group was calculated. In addition to an overall count of accessible IHEs, counts were computed for two IHE characteristics: predominant degree (certificate, associate's, or bachelor's) and

institutional control (public, private not-for-private, or public for-profit). These counts of accessible IHE serve as measures of access.

4.5: Measuring socio-economic status

In education literature, socio-economic status (SES) has three primary components: family income, parental educational attainment, and parental occupational status. Combining these components into a composite variable is possible however there is no established methodology for doing so. Determining SES for a census unit is possible if the data for each component is available. The ACS has variables related to these components but again, there is no standardized methodology for combining them into a single quantity (Cowan et al. 2012). Directly comparing the median household income between census units is very easy, but how might the occupational distribution between census units be compared?

Many univariate proxies for SES have been used extensively in other research fields. These measures, because of their single variable nature, are very simple to compare but may not encompass all the nuances and complexities of SES. Median value of owner-occupied housing is a practically usable proxy for SES because of the “Great Sort” which has moved communities of people based on political ideology and economic means (Oakes 2017).

Based upon these recommendations, in this study, SES has been measured using two SES proxies: median household income and the median value of owner-occupied housing. Within each MSA, each SES proxy was divided into quartiles. The census block groups within the bottom quartiles are considered low SES and the census block groups within the top quartiles are considered high SES. In most, but not all, census block groups both SES proxies were in the same quartile (i.e. both in the top quartile).

Chapter 5: Results

5.1: Computational performance

All of the computation for this study was completed on a personal computer (PC) with the following software setup: Windows 10 Home 64-bit, ArcGIS 10.4 (advanced license), Python 2.7.10 32-bit, and SQLite 3.14.1. The hardware of this PC included an AMD FX-8150 CPU (8 cores at 3.6 GHz each), 8 gigabytes (GB) of DDR3 memory, and a Samsung 850 EVO 250GB solid-state drive (SSD).

The scripts programmed to utilize multiprocessing were limited to six processes so the operating system had two CPU cores to maintain system stability. Out of memory errors occasionally occurred, despite the PC never actually running out of memory. Since the version of Python used in this study was 32-bit, it was unable to allocate a sufficient amount of memory. The 64-bit version of Python does not have this limitation.

The task of generating service areas was subdivided by census region (where the number of IHES in the census region was relatively small) or by state (where the number of IHEs in the census region was relatively large). This resulted in 27 different sets of IHEs to generate service areas for. On average, each IHE took about 13.4 seconds to compute its 135 service areas (two hour and fifteen maximum in one minute time intervals). Road density caused variation in computation time, but overall the computation time appears linear (Figure 12).

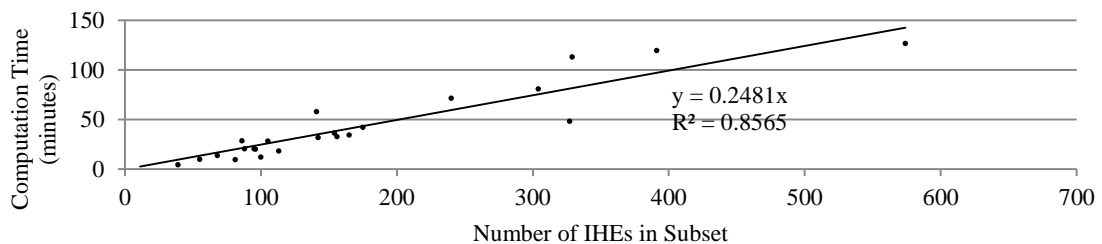


Figure 12: Computation time for service area generation per IHE subset

After generating the service areas for each IHE, it was necessary to repair them as described in Section 4.3. On average, this process took 18.5 seconds per IHE. Since some sets of service areas had no incorrectly generated service areas but other sets of service areas had many (10+ in some cases), the computation ranged from under 5 seconds to just over 60 seconds. This task was parallelized to speed up computation.

Tabulating the intersection of the IHE’s service areas to the census blocks and calculating the population-adjusted average travel time for the census block group took an average of 42.3 seconds per IHE. For many IHEs, these calculations were much quicker than the average, but for others, the computation time was substantially longer (Figure 13).

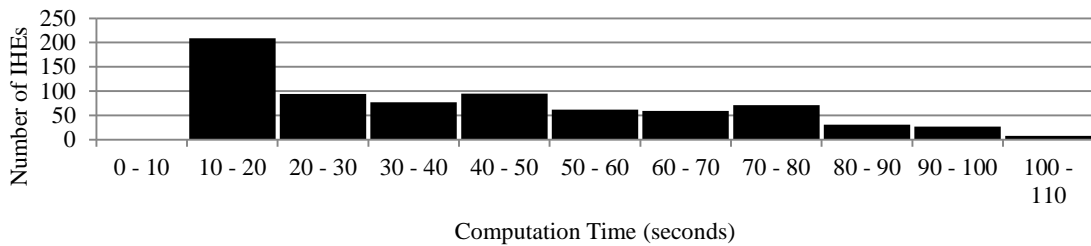


Figure 13: Computation time for calculating the population-adjusted average travel time from a census block group to an IHE

5.2: Types of results

5.2.1: Maps of IHE access

As described in Chapter 4.4, each IHE was classified as accessible or inaccessible for the census block groups within its service areas. The count of accessible IHEs within each census block group was shown in a series of choropleth maps. Within the MSAs where the overall count of IHEs was large, additional choropleth maps were created illustrating the count of accessible IHEs of a particular type (such as predominantly bachelor’s degree granting). Maps of this type are shown in the sections below. Maps

were created using ArcGIS® software by Esri. ArcGIS® and ArcMap™ are the intellectual property of Esri and are used herein under license. Copyright © Esri.

5.2.2: Tables of IHE access by demography and SES

Since the census block group geometry was enriched with ACS data, it is possible to relate the count of accessible IHEs to demographic and economic factors. The population within these census block groups was summarized for each MSA. Table 3, for example, shows the number of accessible IHEs for the population within the Tuscaloosa, AL MSA.

Count of Accessible IHEs	Count of Census Block Groups	Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
0	25	25,668	11,427	13,813	55	25,282	386	10,655	1,620	15,247	268
1	13	19,583	14,833	4,391	4	19,183	400	1,007	6,921	9,788	0
2	6	11,193	9,266	1,282	356	10,815	378	1,594	2,075	0	2,075
3	39	68,816	53,238	12,887	1,423	65,726	3,090	12,772	31,534	3,228	41,096
4	62	105,588	56,820	45,851	1,152	102,904	2,684	21,241	32,424	11,400	25,792
5	2	4,722	4,268	381	18	4,719	3	0	4,722	0	4,126

Table 3: Example from Tuscaloosa, AL of the count of any accessible IHE for different demographic and economic factors

Tables of this form were also generated for each IHE subset (i.e. public bachelor’s degree granting, private for-profit certificate degree granting, etc.) used in this study. Therefore, for each MSA, a total of seven tables of this form were generated. These tables were summarized into two separate tables by summarizing each column for each IHE subset as a weighted average or a median. The median is a more appropriate measure of central tendency since this data is measured on an ordinal scale and not normally distributed (Manikandan 2011), but the weighted average reflects outlier census

block groups with a larger number of accessible IHEs more clearly. Tables with the median and weighted average number of accessible IHEs by demographic and economic factors are shown in the sections below.

5.3: Dalton, GA MSA

The map of the median commute time to work for the census block groups of this MSA (Figure 14), indicates longer commutes are taken by those living outside of Dalton, GA. The longest commutes are found in the northernmost census block groups of the MSA. Using these median commute times to work as the maximum travel time for which an IHE is deemed accessible, the number of accessible IHEs per census block group was tabulated.

The number of accessible IHEs of any type is shown in Figure 15. Only two IHEs exist within the MSA, whereas four additional IHEs outside of the MSA are accessible to some census block groups. Figure 16 through Figure 21 illustrate the number of accessible IHEs by IHE type. Only two census block groups have any accessible predominantly bachelor's-degree granting IHEs (Figure 18).

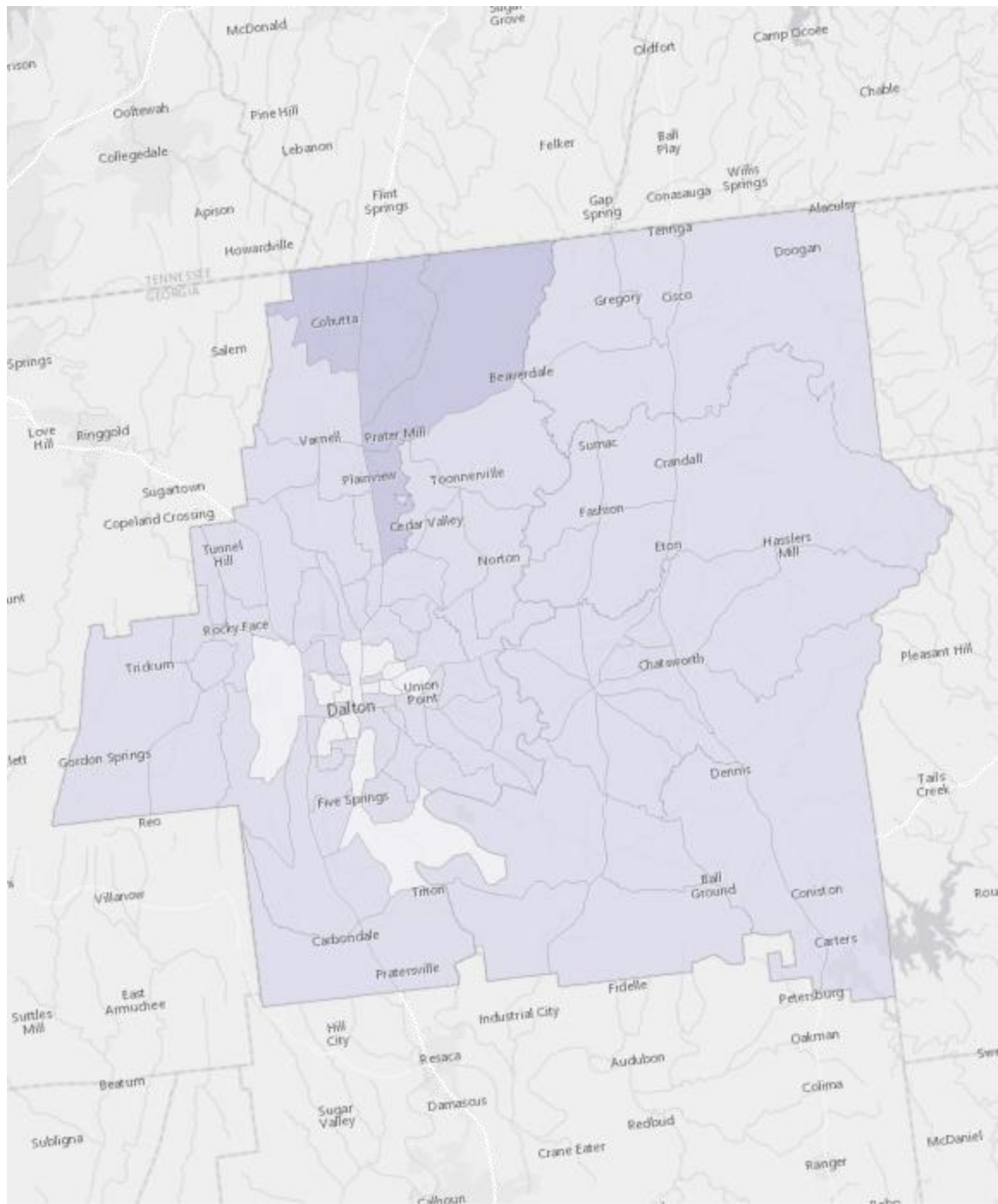
Table 4 and Table 5 summarize the number of accessible IHEs (by type) for demographic and SES factors. Table 4 summarizes by calculating the median number of accessible IHEs. Table 5 summarizes by calculating the average number of accessible IHEs. Overall, within this MSA, the number of accessible IHEs is quite small and there is no discernable difference in the amount of access between different demographic or SES factors.

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		2	2	2	2	2	2	2	2	2	2
Predominant Degree	Certificate	1	1	1	1	1	1	1	1	1	1
	Associate's	1	1	1	1	1	1	1	1	1	1
	Bachelor's	0	0	0	0	0	0	0	0	0	0
Institution Control	Public	1	1	1	1	1	1	1	1	1	1
	Private Not-For-Profit	0	0	0	0	0	0	0	0	0	0
	Private For-Profit	1	1	1	1	1	1	1	1	1	1

Table 4: Median number of accessible IHEs (by type) for demographic and SES factors of the Dalton, GA MSA

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		1.59	1.56	1.87	1.86	1.50	1.82	1.36	1.80	1.25	1.73
Predominant Degree	Certificate	0.86	0.85	0.95	0.94	0.82	0.95	0.74	0.95	0.65	0.93
	Associate's	0.71	0.69	0.91	0.91	0.65	0.85	0.62	0.85	0.60	0.80
	Bachelor's	0.02	0.02	0.00	0.00	0.03	0.01	0.00	0.00	0.00	0.00
Institution Control	Public	0.71	0.69	0.91	0.91	0.65	0.85	0.62	0.85	0.60	0.80
	Private Not-For-Profit	0.02	0.02	0.00	0.00	0.03	0.01	0.00	0.00	0.00	0.00
	Private For-Profit	0.86	0.85	0.95	0.94	0.82	0.95	0.74	0.95	0.65	0.93

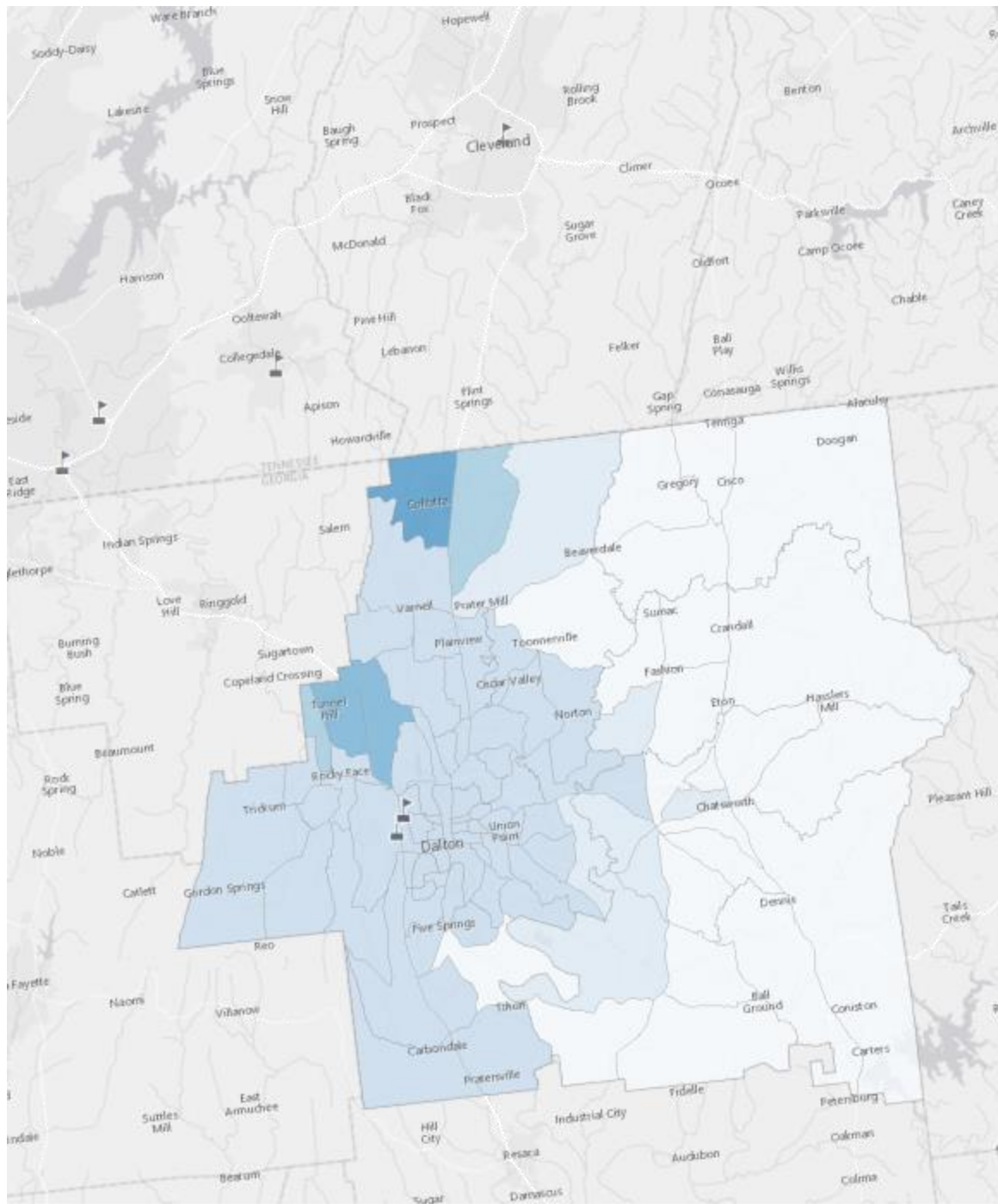
Table 5: Average number of accessible IHEs (by type) for demographic and SES factors of the Dalton, GA MSA



Median Commute Time to Work < 15 15 - 30 30 - 45 45 - 60 60 - 90 > 90



Figure 14: Median commute time to work for the census block groups of the Dalton, GA MSA



Number of Accessible IHEs (all types) 0 1 2 3 4 5 6 7 8+



Figure 15: Number of accessible IHEs (all types) for the census block groups of the Dalton, GA MSA

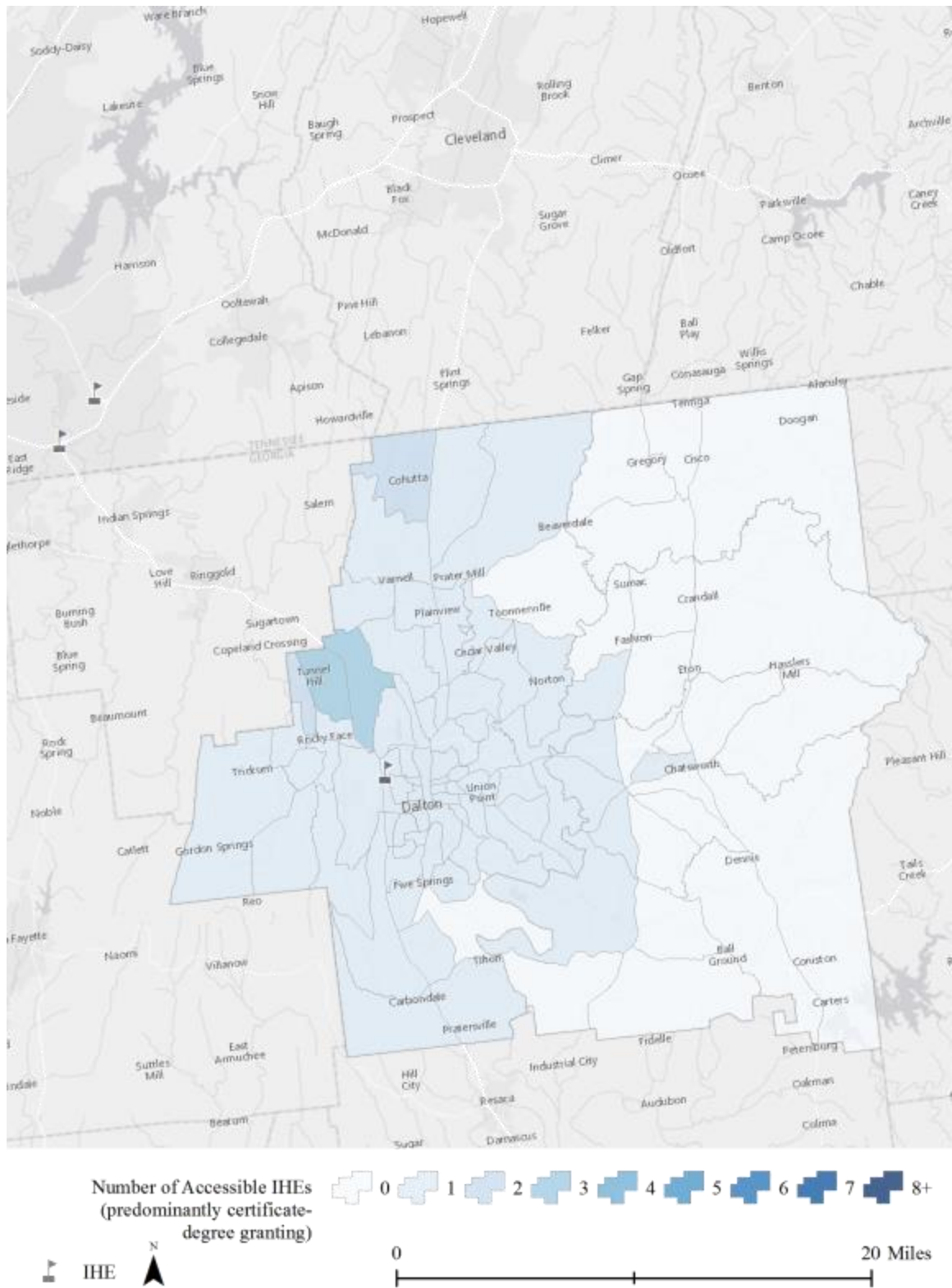
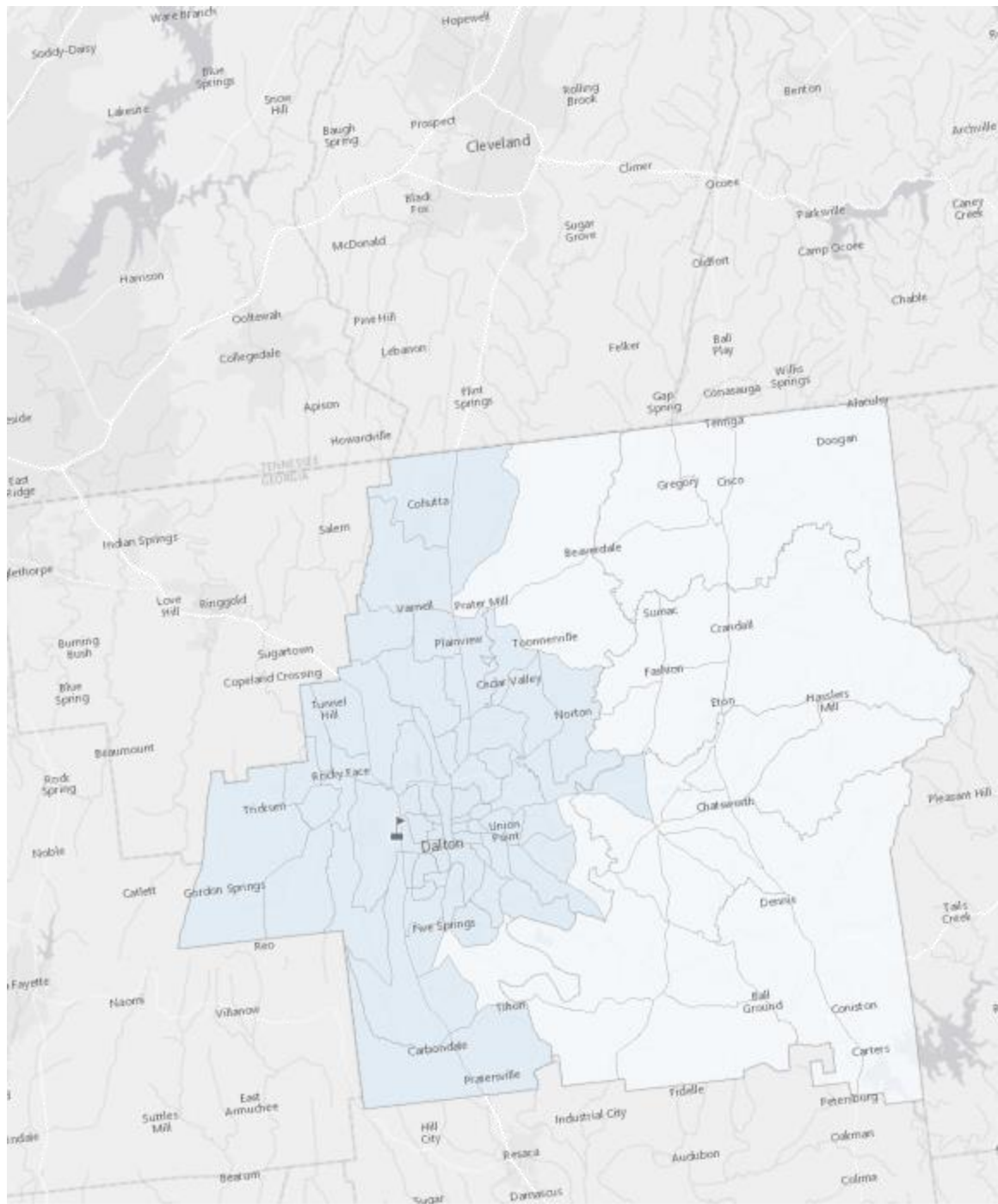


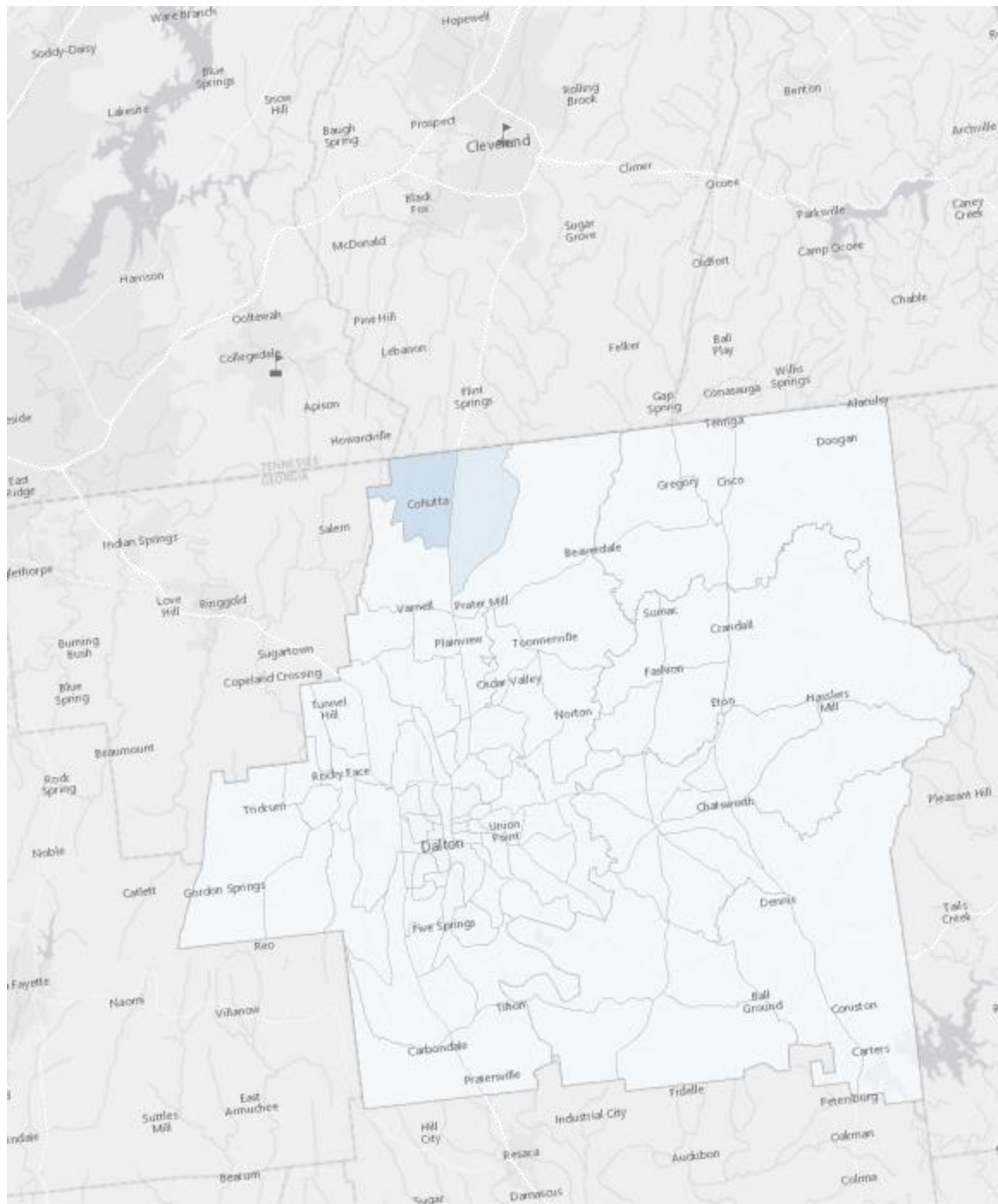
Figure 16: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Dalton, GA MSA



Number of Accessible IHEs
(predominantly associate's-
degree granting)



Figure 17: Number of accessible IHEs (predominantly associate's-degree granting) for the census block groups of the Dalton, GA MSA



Number of Accessible IHEs
(predominantly bachelor's
degree granting)

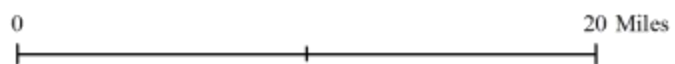
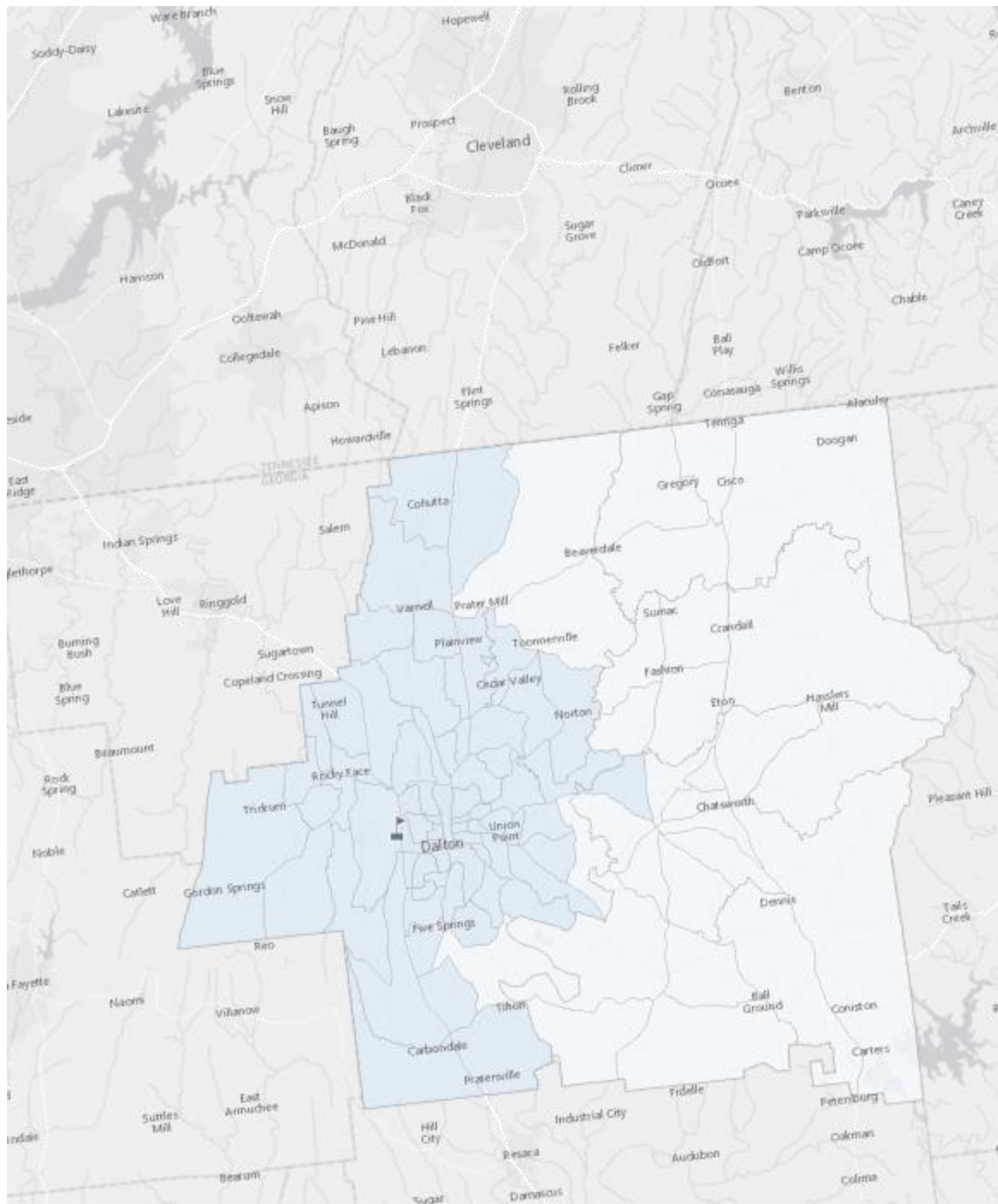


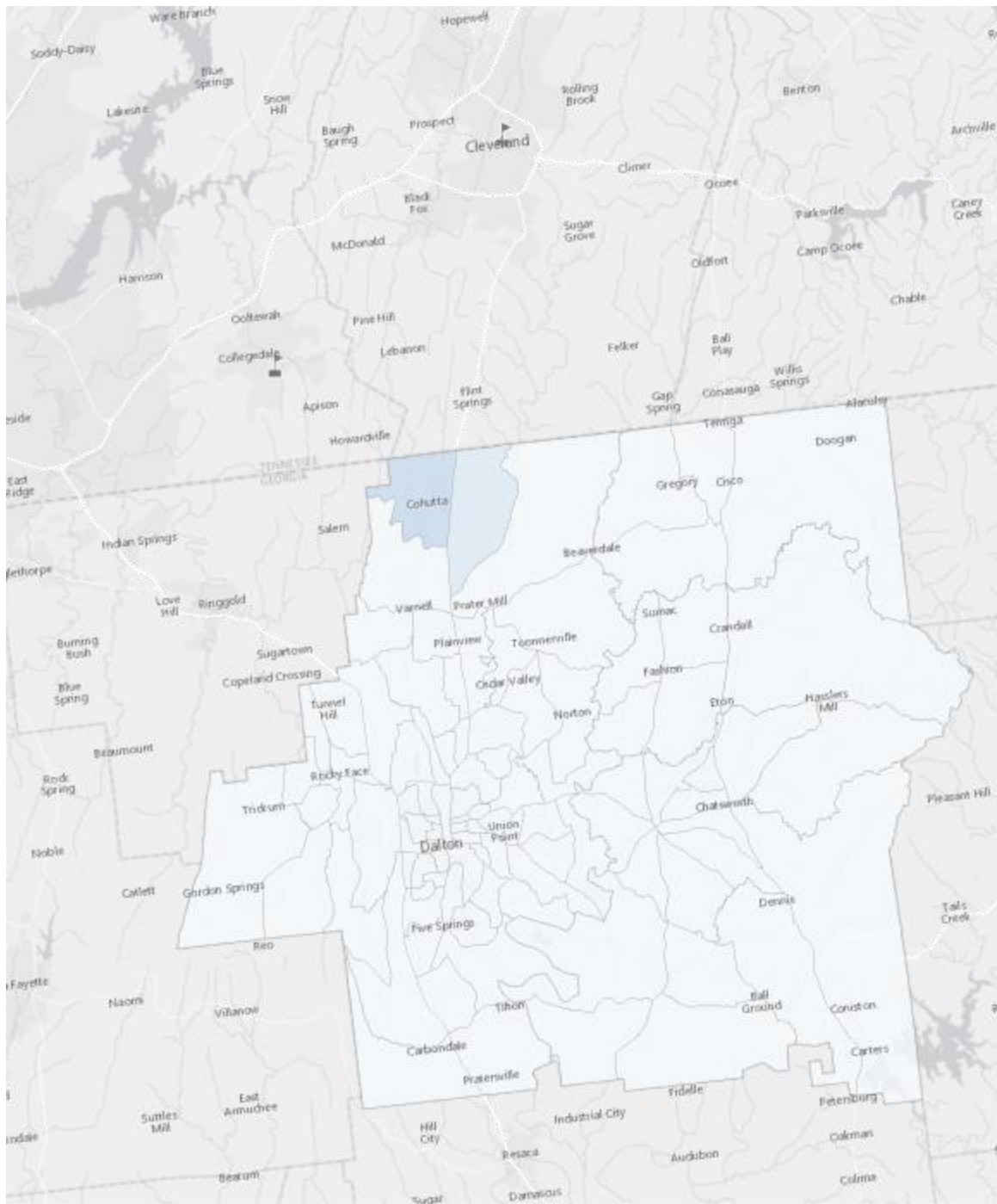
Figure 18: Number of accessible IHEs (predominantly bachelor's degree granting) for the census block groups of the Dalton, GA MSA



Number of Accessible IHEs (public) 0 1 2 3 4 5 6 7 8+



Figure 19: Number of accessible IHEs (public) for the census block groups of the Dalton, GA MSA



Number of Accessible IHEs (private non-profit) 0 1 2 3 4 5 6 7 8+



0 20 Miles

Figure 20: Number of accessible IHEs (private non-profit) for the census block groups of the Dalton, GA MSA

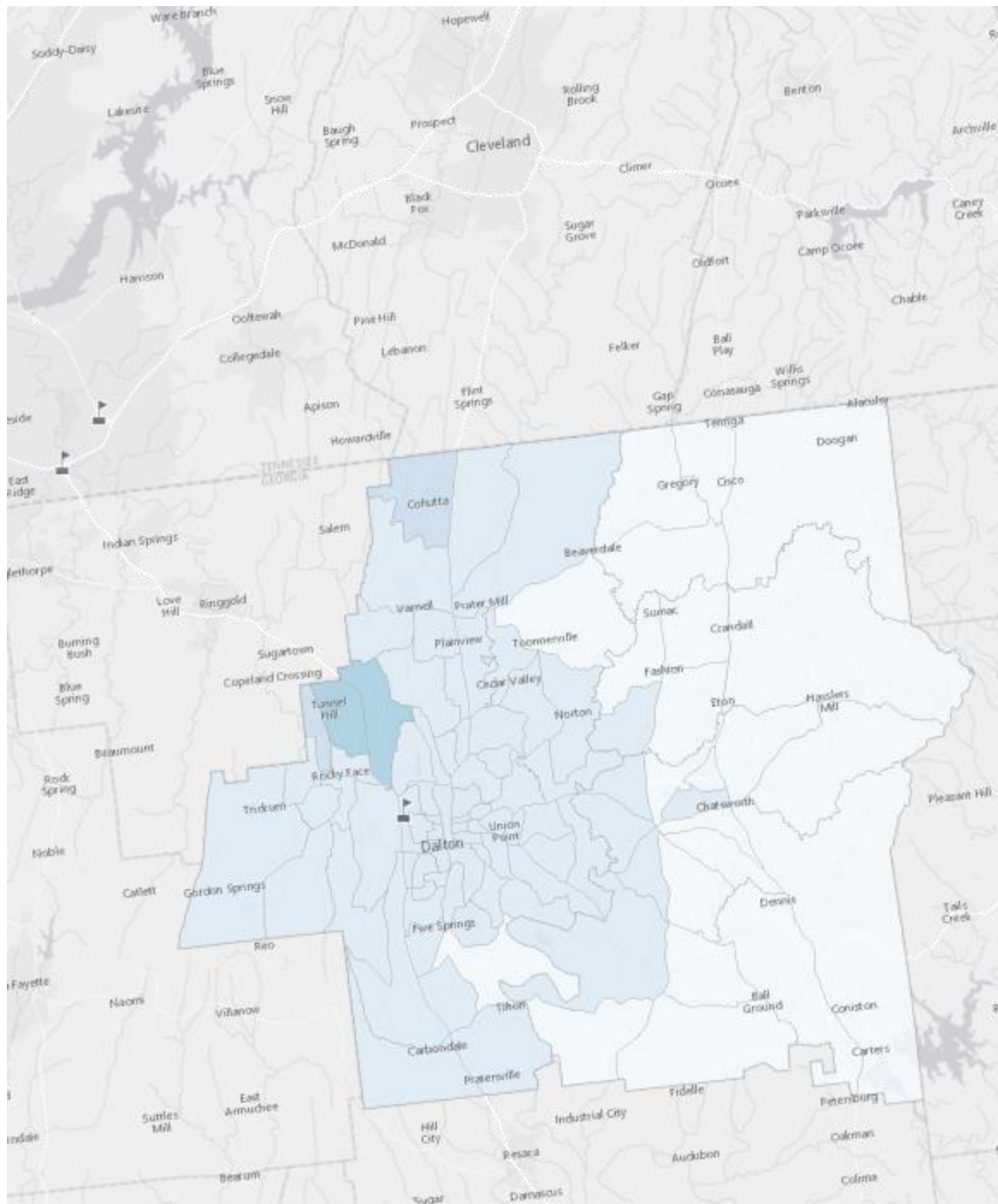


Figure 21: Number of accessible IHEs (private for-profit) for the census block groups of the Dalton, GA MSA

5.4: Joplin, MO MSA

The map of the median commute time to work for the census block groups of this MSA (Figure 22), indicates longer commutes are taken by those living outside of Joplin or Carthage. The longest commutes are found in the periphery census block groups of the MSA. Using these median commute times to work as the maximum travel time for which an IHE is deemed accessible, the number of accessible IHEs per census block group was tabulated.

The number of accessible IHEs of any type is shown in Figure 23. There are seven accessible IHEs within the MSA with an additional IHE outside of the MSA accessible to some census block groups within the MSA. Figure 24 through Figure 29 illustrate the number of accessible IHEs by IHE type. Compared to predominantly certificate- or bachelor-degree granting IHEs, predominantly associate's-degree granting institutions are the least accessible.

Table 6 and Table 7 summarize the number of accessible IHEs (by type) for demographic and SES factors. Table 6 summarizes by calculating the median number of accessible IHEs. Table 7 summarizes by calculating the average number of accessible IHEs. The Hispanic / Latino population has fewer accessible IHEs than any other race / ethnicity. The low SES census block groups have fewer accessible IHEs than the high SES census block groups. The difference is greater when using the median home value SES proxy than the median household income SES proxy, but the trend is consistent between both.

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		5	5	6	5	5	2	5	6	3	6
Predominant Degree	Certificate	2	2	3	2	2	1	2	3	2	3
	Associate's	1	1	1	1	1	1	1	1	1	1
	Bachelor's	2	2	2	2	2	0	2	2	1	2
Institution Control	Public	2	2	2	2	2	1	2	2	2	2
	Private Not-For-Profit	1	1	1	1	1	0	1	1	0	1
	Private For-Profit	2	2	3	2	2	1	2	3	1	3

Table 6: Median number of accessible IHEs (by type) for demographic and SES factors of the Joplin, MO MSA

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		4.11	4.11	4.97	4.08	4.18	3.08	4.10	4.56	3.72	4.58
Predominant Degree	Certificate	2.18	2.18	2.54	2.03	2.20	1.82	2.10	2.35	1.96	2.42
	Associate's	0.76	0.75	0.90	0.91	0.77	0.52	0.78	0.80	0.77	0.84
	Bachelor's	1.17	1.17	1.53	1.15	1.20	0.74	1.22	1.41	1.00	1.32
Institution Control	Public	1.64	1.64	1.84	1.38	1.65	1.46	1.66	1.73	1.62	1.73
	Private Not-For-Profit	0.58	0.58	0.75	0.58	0.59	0.36	0.60	0.72	0.48	0.65
	Private For-Profit	1.90	1.89	2.38	2.12	1.94	1.26	1.84	2.12	1.63	2.20

Table 7: Average number of accessible IHEs (by type) for demographic and SES factors of the Joplin, MO MSA

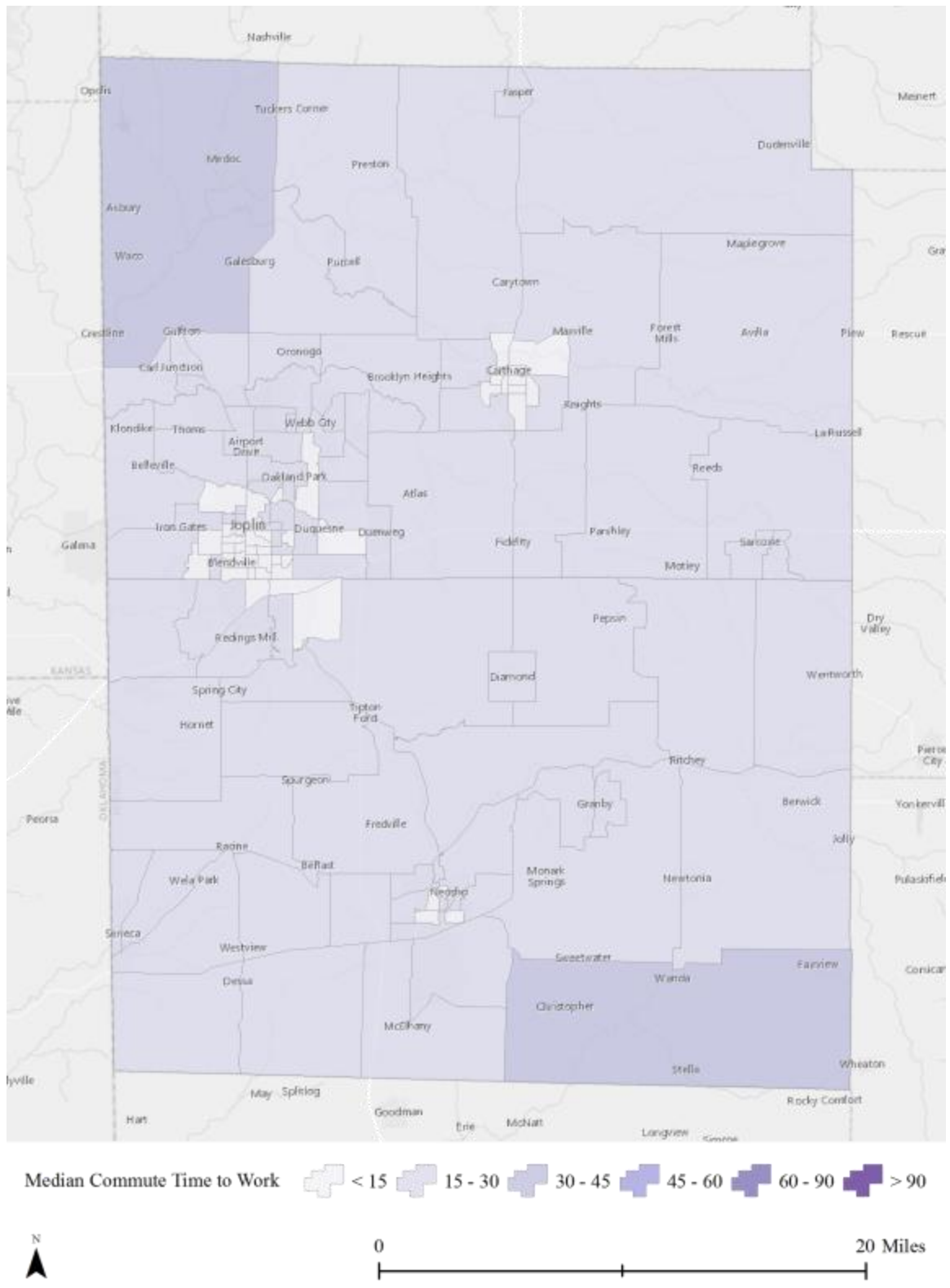


Figure 22: Median commute time to work for the census block groups of the Joplin, MO MSA

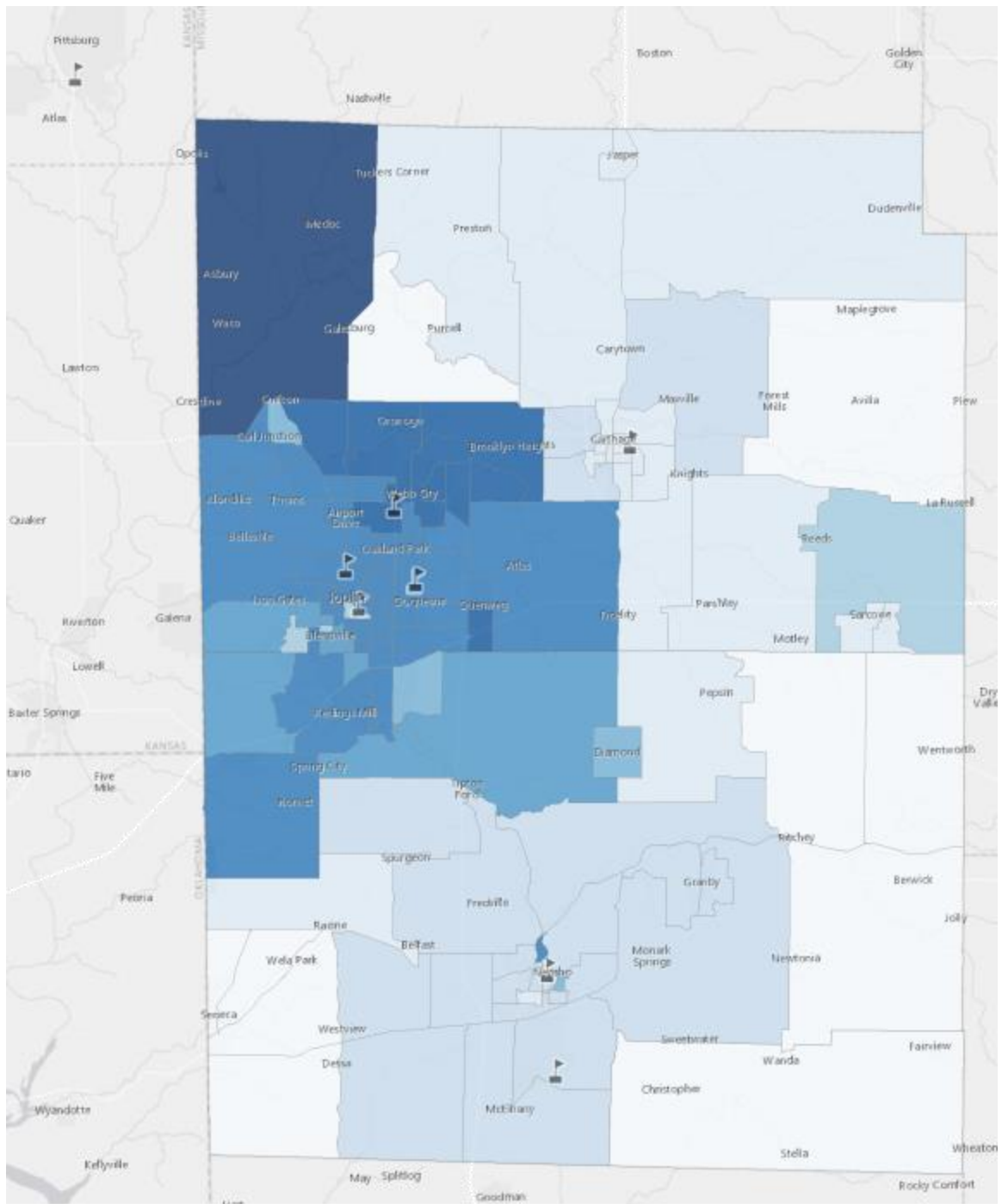


Figure 23: Number of accessible IHEs (all types) for the census block groups of the Joplin, MO MSA

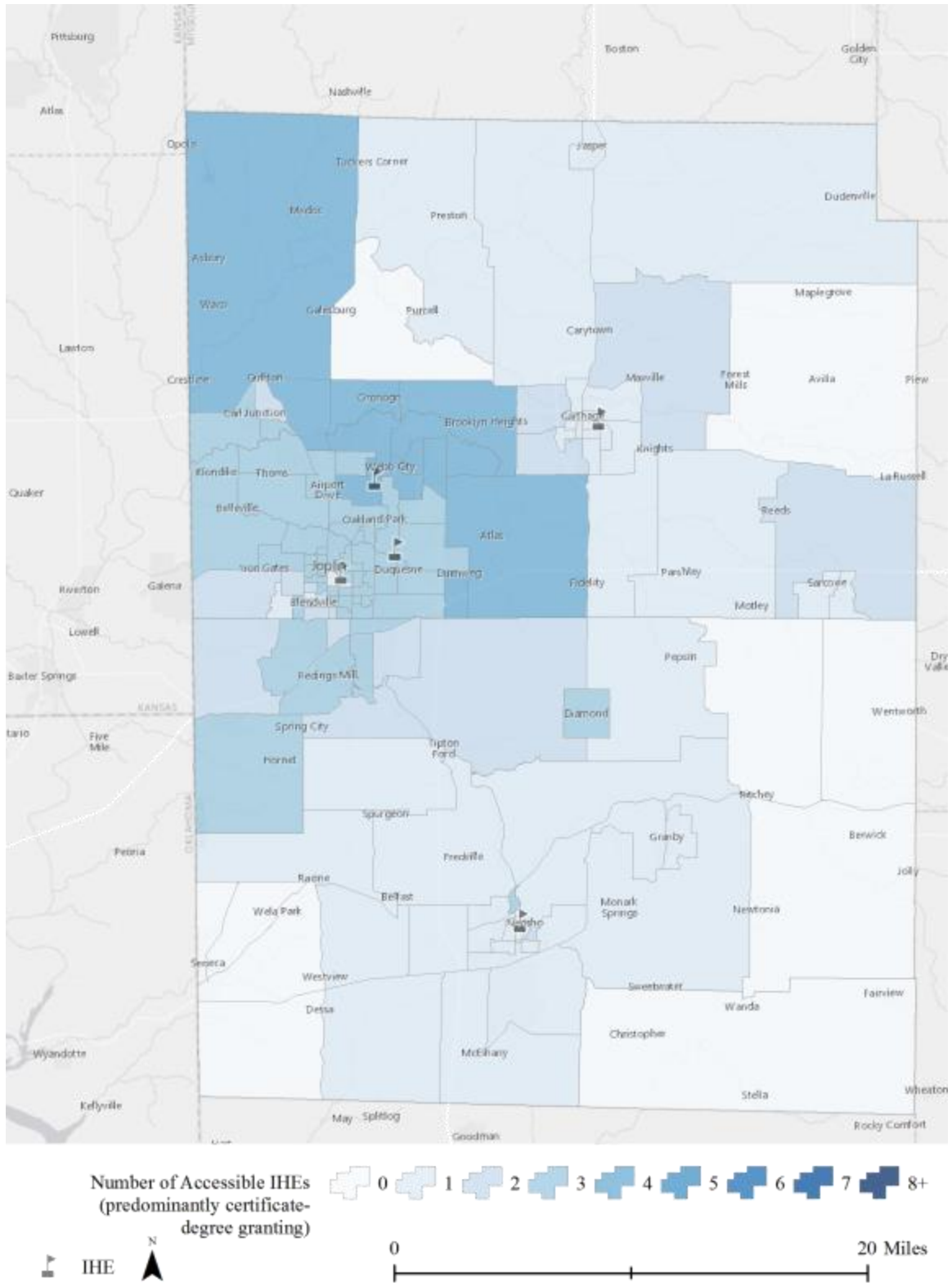


Figure 24: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Joplin, MO MSA

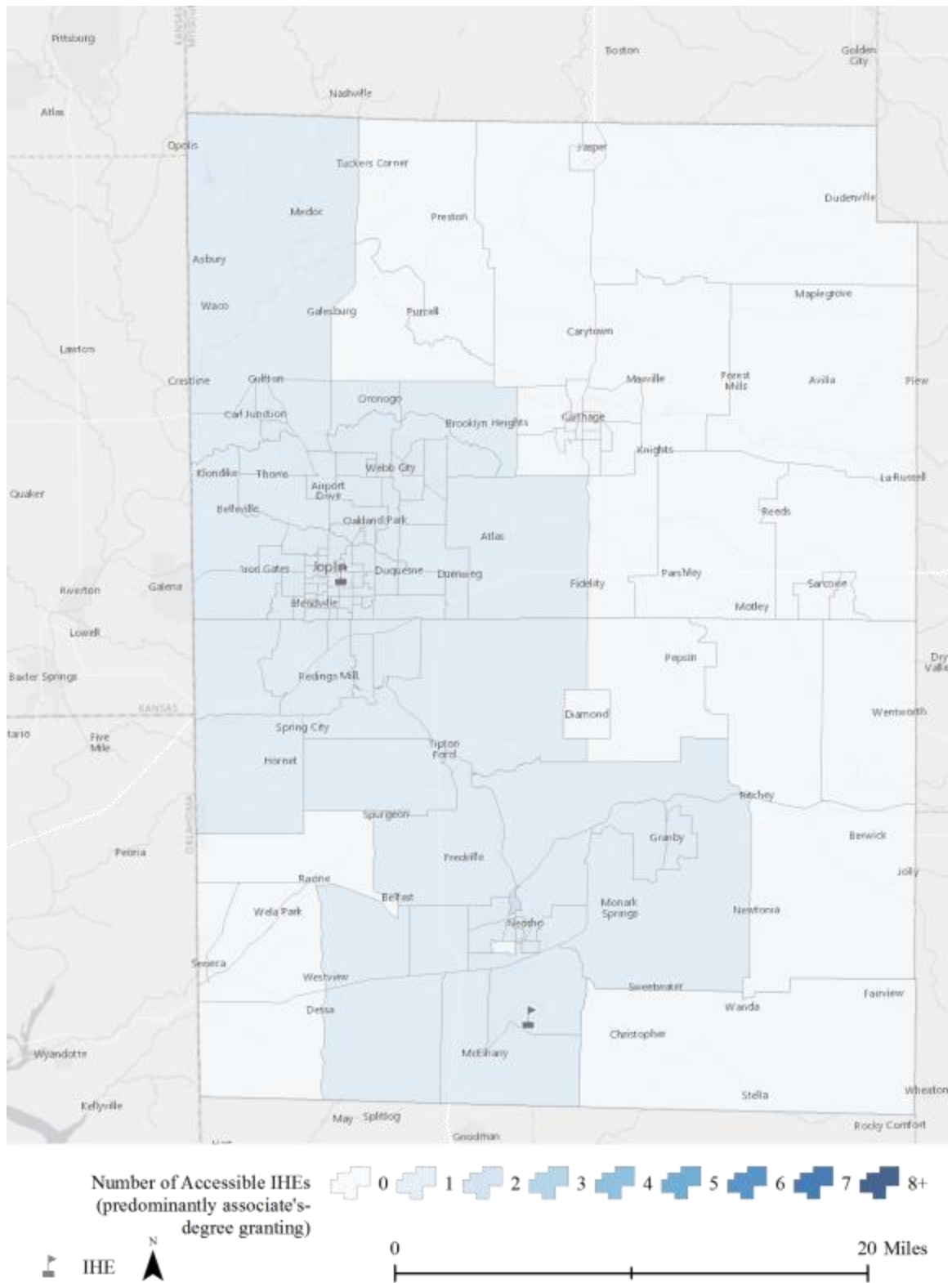


Figure 25: Number of accessible IHEs (predominantly associate's-degree granting) for the census block groups of the Joplin, MO MSA

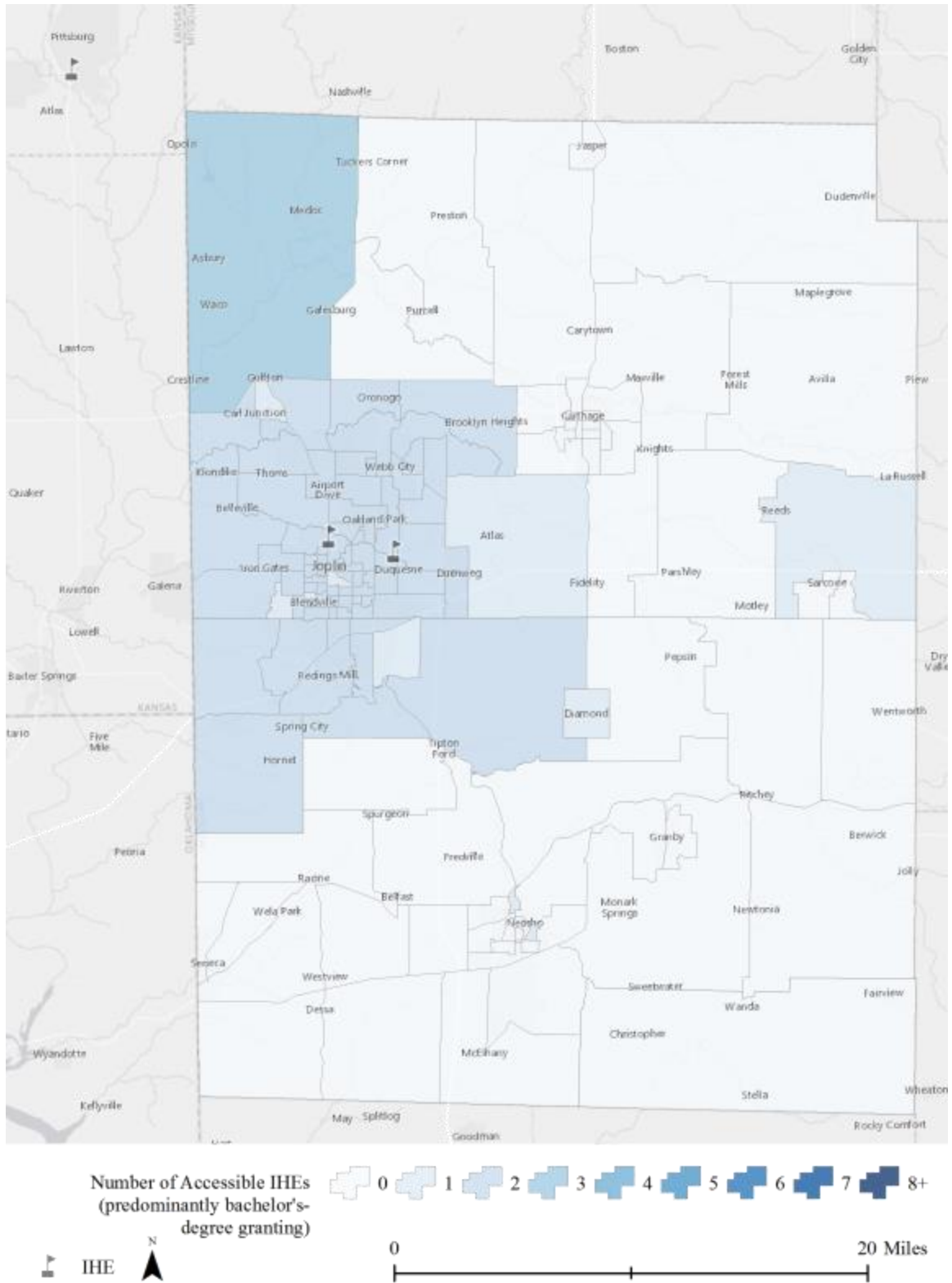
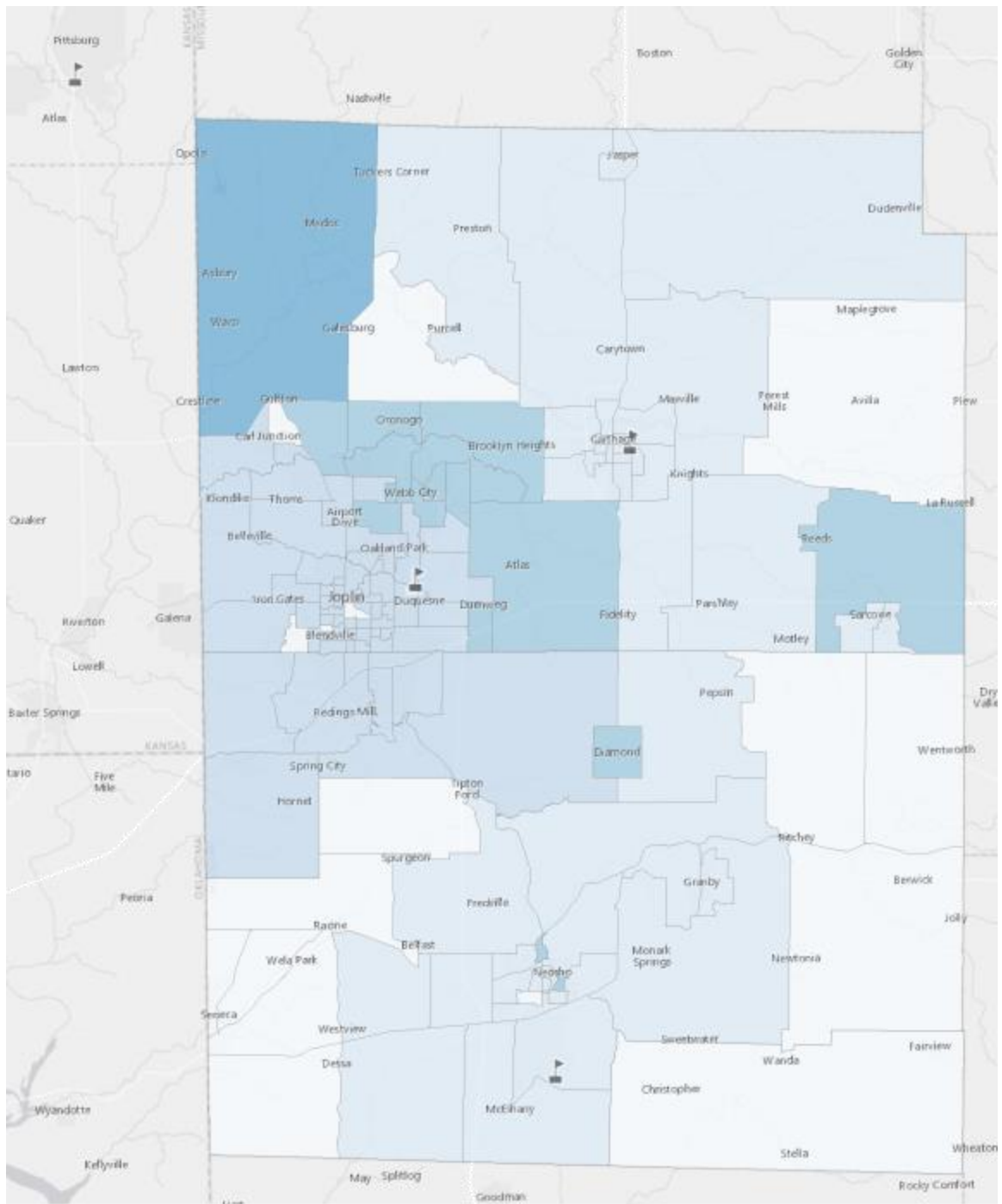


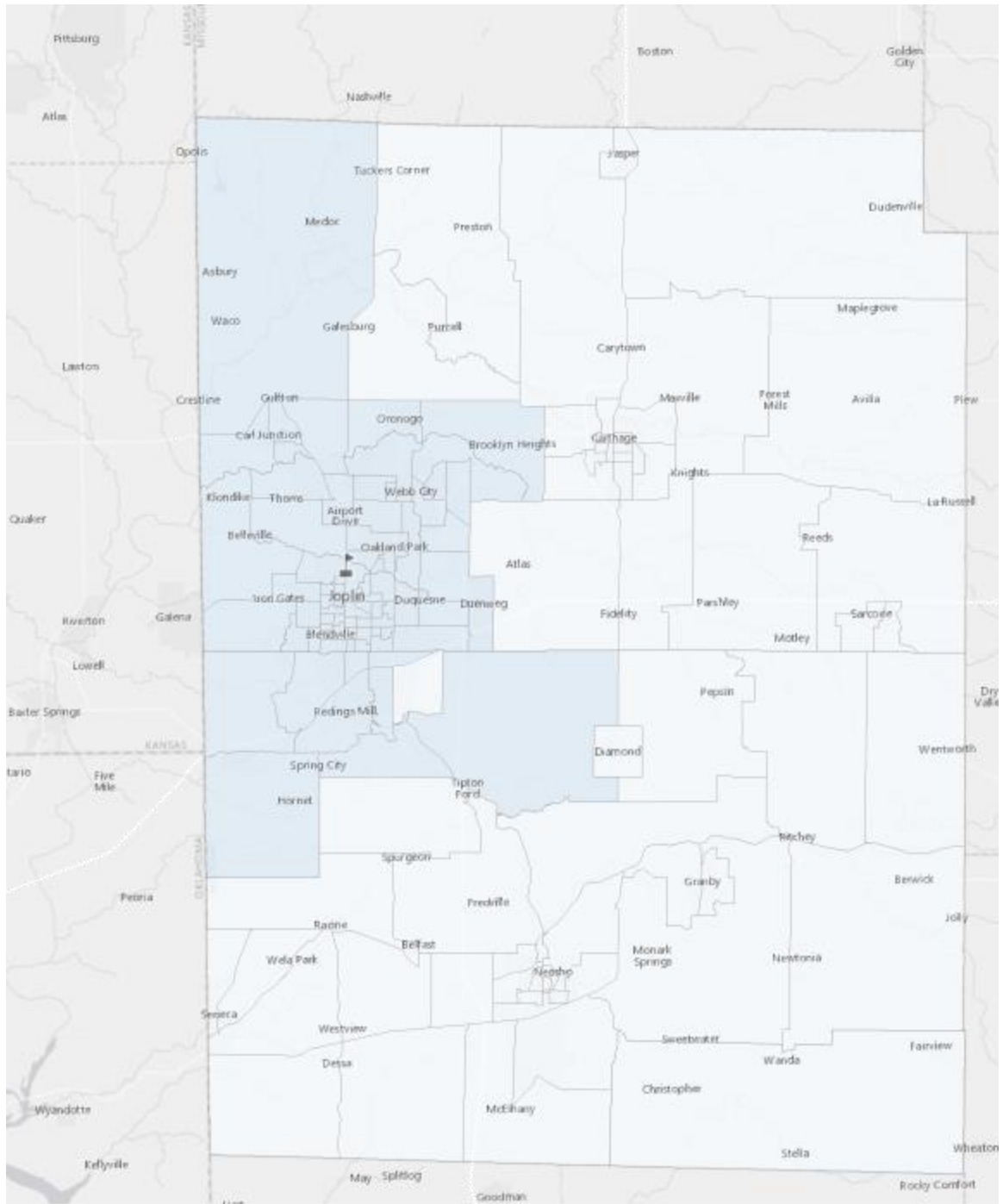
Figure 26: Number of accessible IHEs (predominantly bachelor's-degree granting) for the census block groups of the Joplin, MO MSA



Number of Accessible IHEs (public) 0 1 2 3 4 5 6 7 8+



Figure 27: Number of accessible IHEs (public) for the census block groups of the Joplin, MO MSA



Number of Accessible IHEs (private non-profit)

0	1	2	3	4	5	6	7	8+
---	---	---	---	---	---	---	---	----



Figure 28: Number of accessible IHEs (private non-profit) for the census block groups of the Joplin, MO MSA

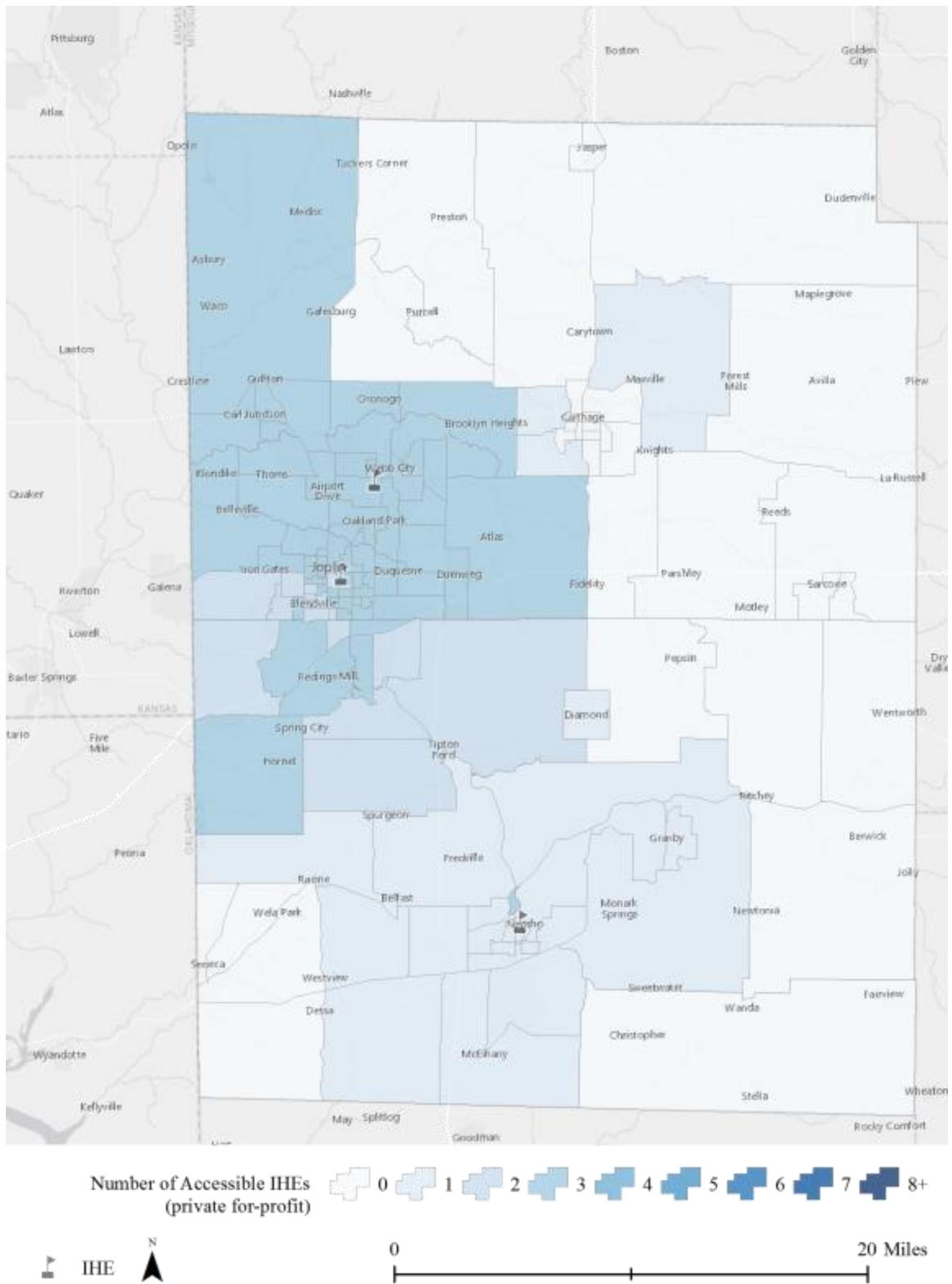


Figure 29: Number of accessible IHEs (private for-profit) for the census block groups of the Joplin, MO MSA

5.5: Vineland – Bridgeton, NJ MSA

The map of the median commute time to work for the census block groups of this MSA (Figure 30) indicates commuters in Bridgeton, NJ are, on average, willing and able to commute further than commuters in Vineland, NJ. The longest commutes are near in the southernmost coastal region of the MSA. Using these median commute times to work as the maximum travel time for which an IHE is deemed accessible, the number of accessible IHEs per census block group was tabulated.

The number of accessible IHEs of any type is shown in Figure 31. One IHE exists within the MSA, whereas thirteen additional IHEs outside of the MSA are accessible to some census block groups. Figure 32 through Figure 37 illustrate the number of accessible IHEs by IHE type. One predominantly associate's degree-granting IHE is available to the majority of census block groups within the MSA (Figure 33). Very few census block groups within the MSA have any access to a predominantly bachelor's degree-granting IHE (Figure 34).

Error! Reference source not found. and **Error! Reference source not found.** summarize the number of accessible IHEs (by type) for demographic and SES factors. **Error! Reference source not found.** summarizes by calculating the median number of accessible IHEs. **Error! Reference source not found.** summarizes by calculating the average number of accessible IHEs. Overall, within this MSA, the number of accessible IHEs is quite small and there is no discernable difference in the amount of access between different demographic or SES factors.

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		1	1	1	1	1	1	1	1	1	1
Predominant Degree	Certificate	0	0	0	0	0	0	0	0	0	0
	Associate's	1	1	1	1	1	1	1	1	1	1
	Bachelor's	0	0	0	0	0	0	0	0	0	0
Institution Control	Public	1	1	1	1	1	1	1	1	1	1
	Private Not-For-Profit	0	0	0	0	0	0	0	0	0	0
	Private For-Profit	0	0	0	0	0	0	0	0	0	0

Table 8: Median number of accessible IHEs (by type) for demographic and SES factors of the Vineland – Bridgeton, NJ MSA

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		0.95	0.95	0.85	0.93	0.89	1.10	1.31	0.76	1.23	0.91
Predominant Degree	Certificate	0.06	0.07	0.02	0.08	0.06	0.06	0.20	0.04	0.23	0.00
	Associate's	0.84	0.84	0.81	0.81	0.80	0.95	1.01	0.72	0.95	0.91
	Bachelor's	0.04	0.03	0.02	0.04	0.02	0.09	0.09	0.00	0.06	0.00
Institution Control	Public	0.87	0.86	0.81	0.83	0.82	0.99	1.10	0.72	1.01	0.86
	Private Not-For-Profit	0.02	0.02	0.02	0.02	0.01	0.04	0.00	0.00	0.00	0.05
	Private For-Profit	0.06	0.07	0.02	0.08	0.06	0.06	0.20	0.04	0.23	0.00

Table 9: Average number of accessible IHEs (by type) for demographic and SES factors of the Vineland – Bridgeton, NJ MSA

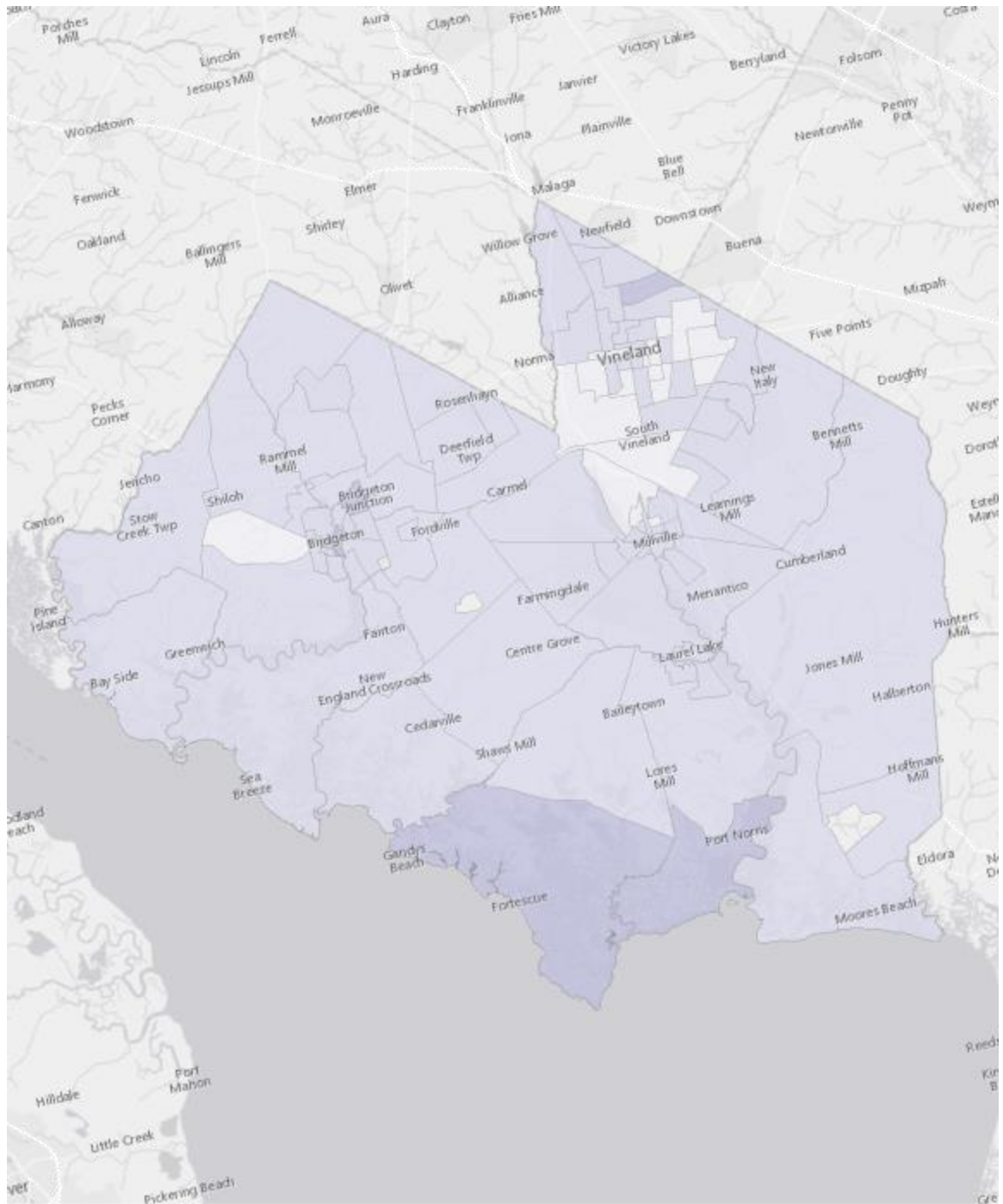


Figure 30: Median commute time to work for the census block groups of the Vineland – Bridgeton, NJ MSA

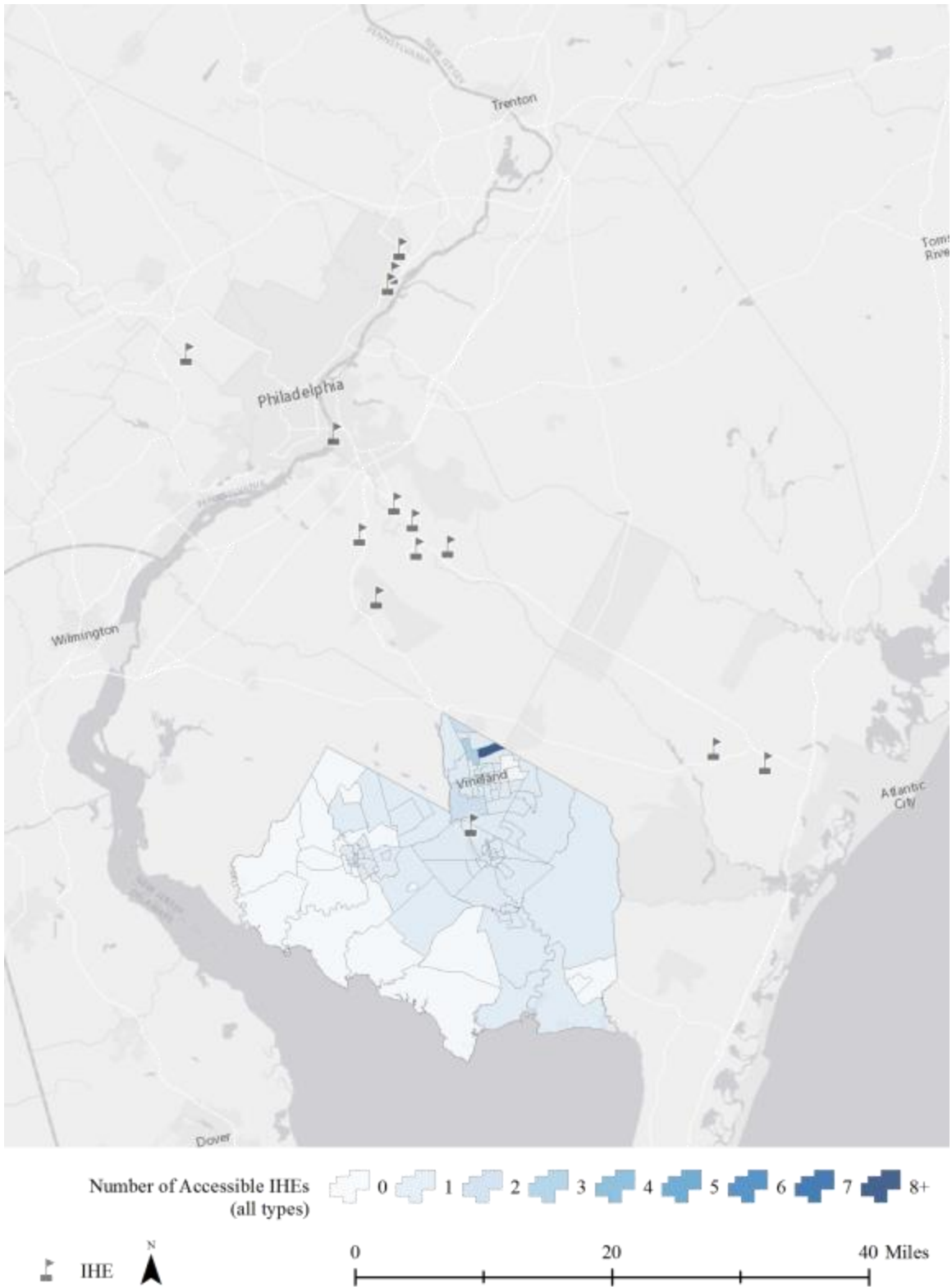


Figure 31: Number of accessible IHEs (all types) for the census block groups of the Vineland – Bridgeton, NJ MSA

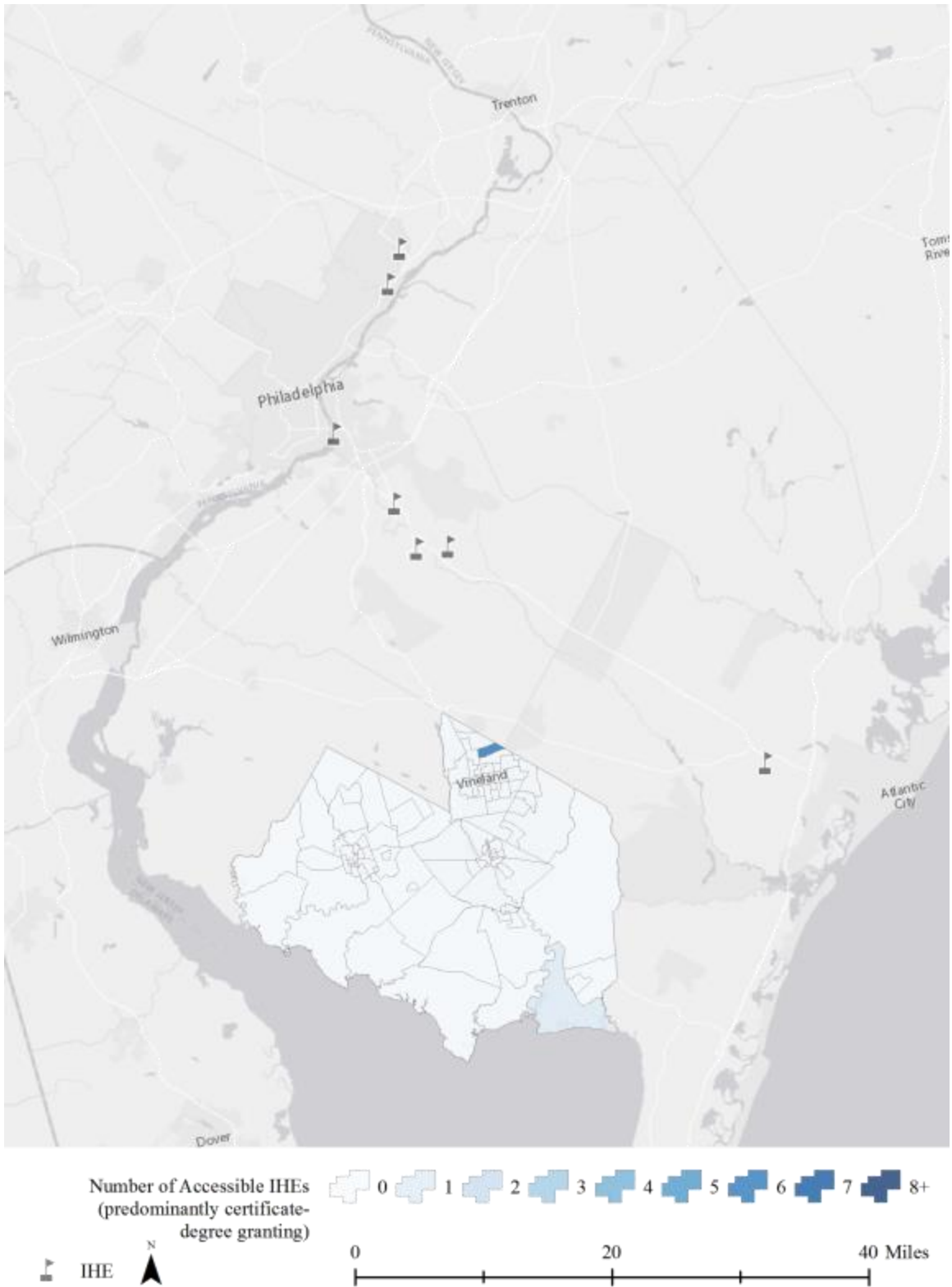


Figure 32: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Vineland – Bridgeton, NJ MSA

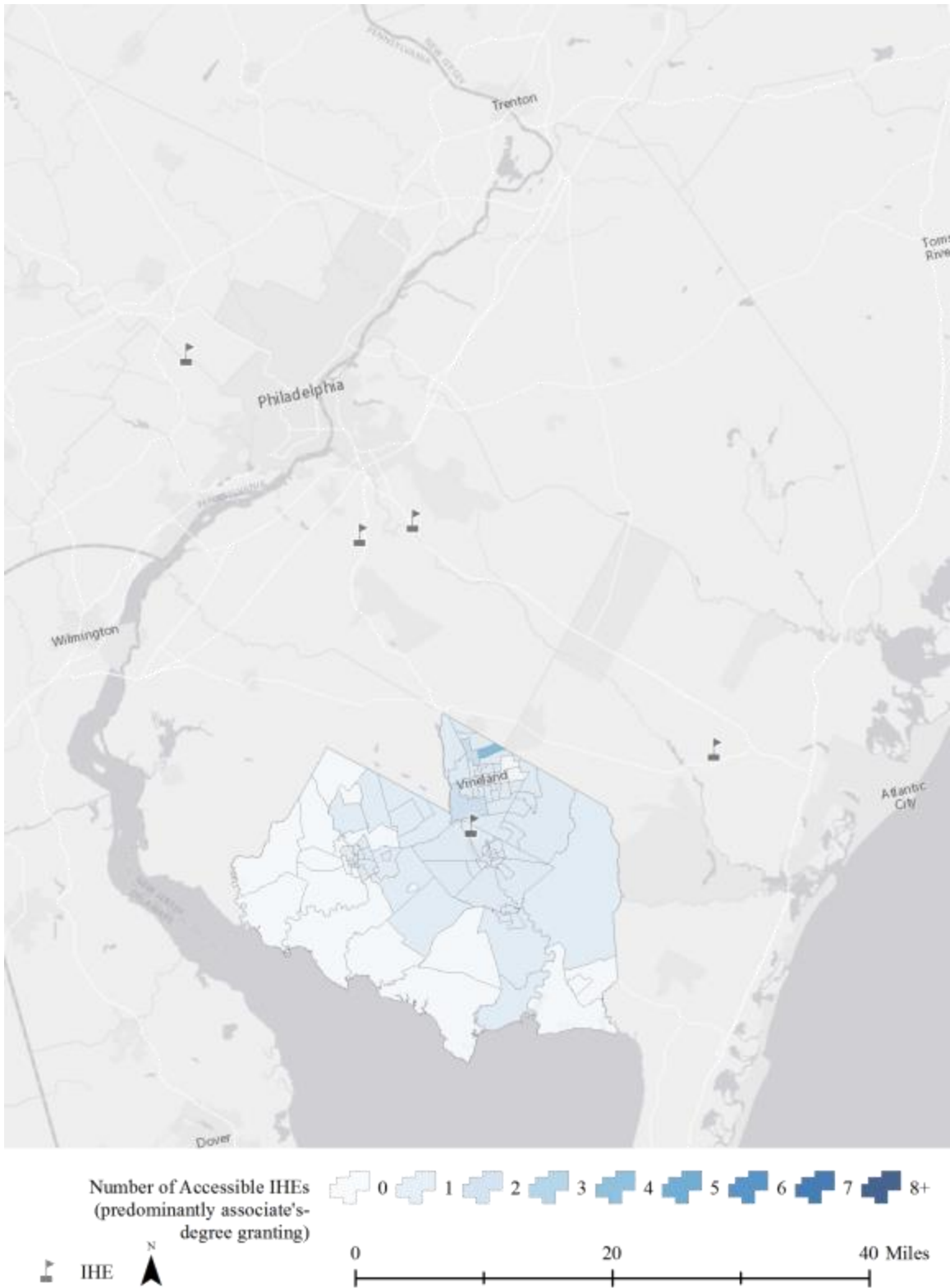


Figure 33: Number of accessible IHEs (predominantly associate's-degree granting) for the census block groups of the Vinland – Bridgeton, NJ MSA

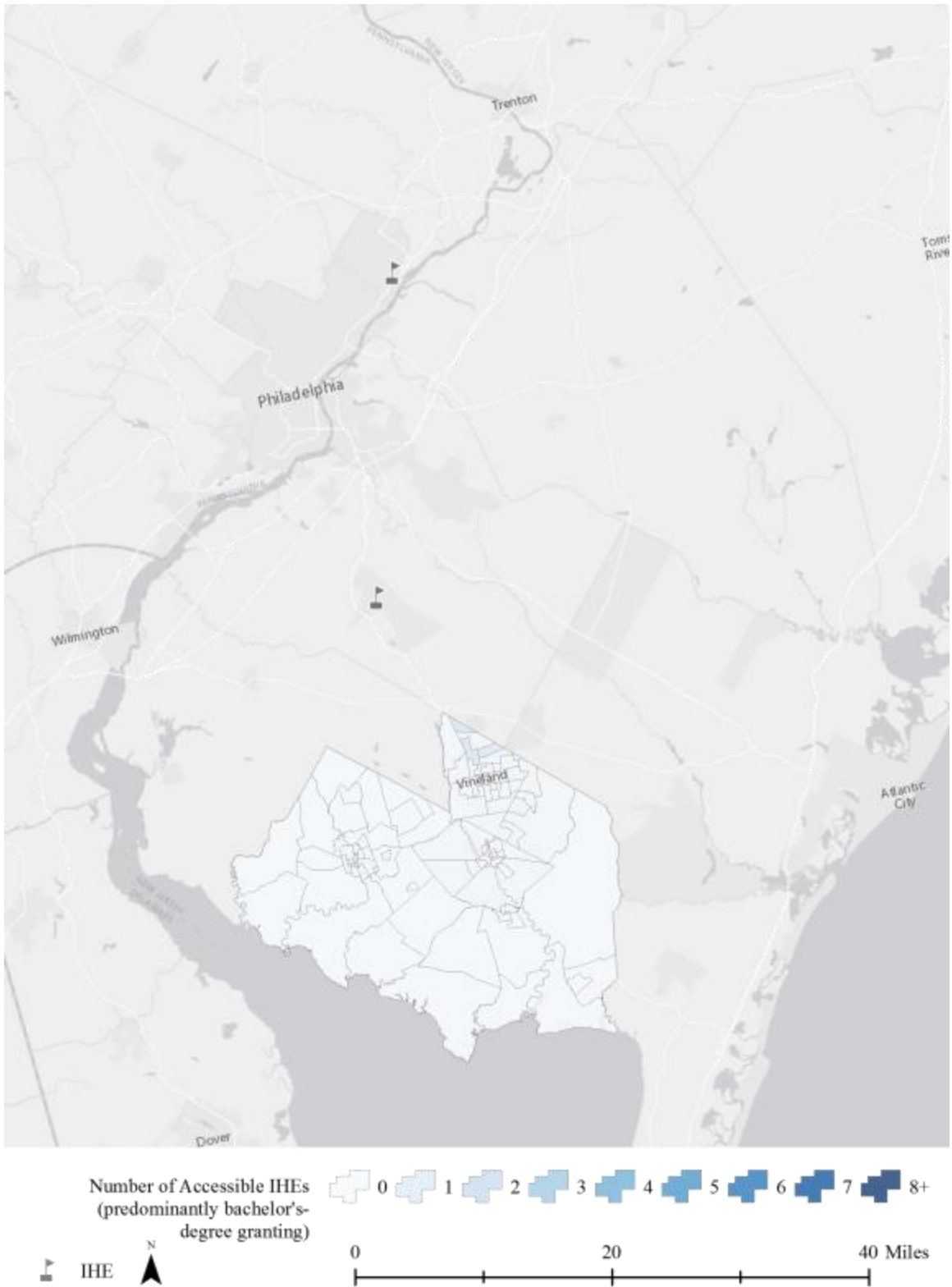


Figure 34: Number of accessible IHEs (predominantly bachelor's-degree granting) for the census block groups of the Vinland – Bridgeton, NJ MSA

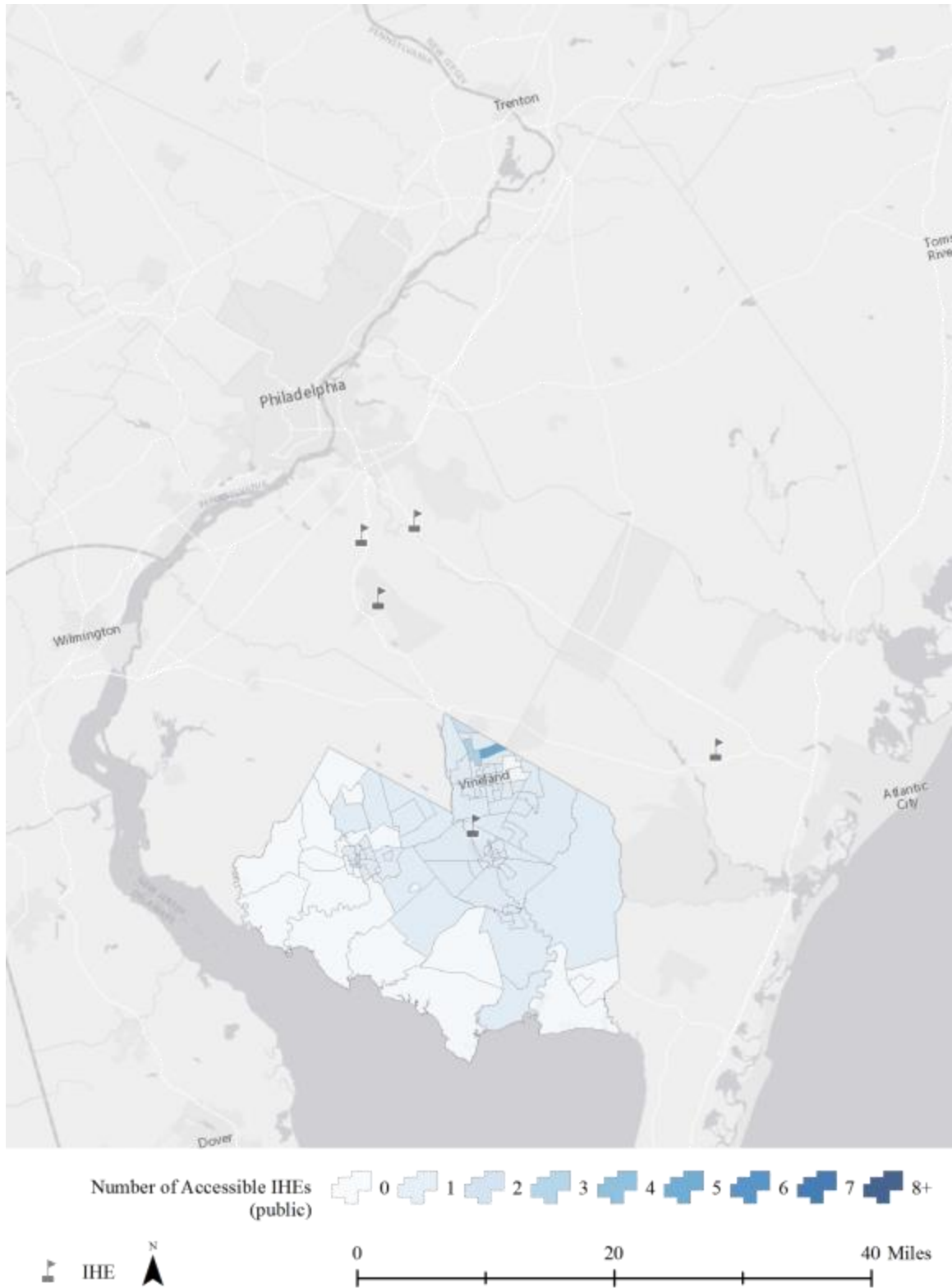


Figure 35: Number of accessible IHEs (public) for the census block groups of the Vineland – Bridgeton, NJ MSA

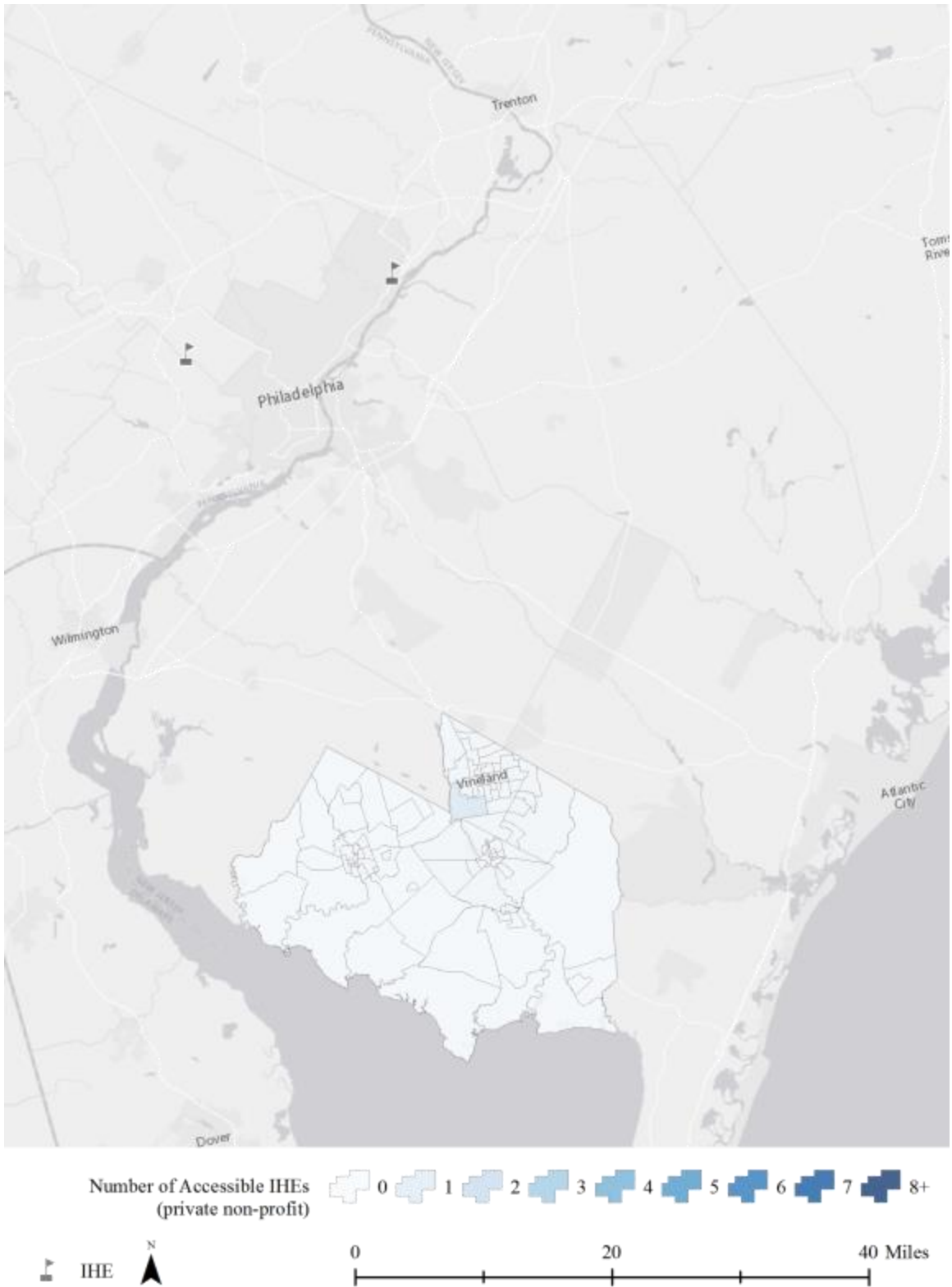


Figure 36: Number of accessible IHEs (private non-profit) for the census block groups of the Vineland – Bridgeton, NJ MSA

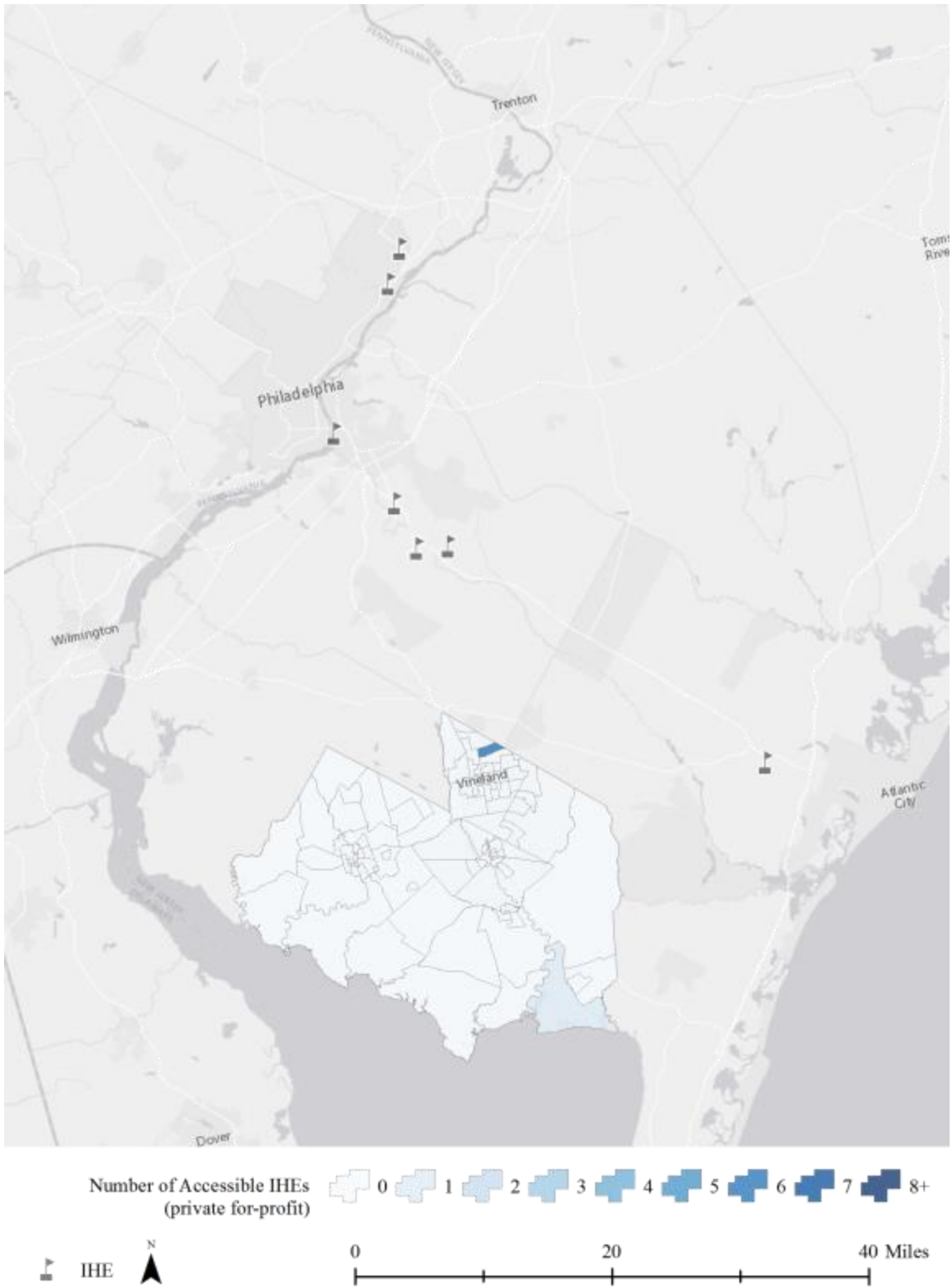


Figure 37: Number of accessible IHEs (private for-profit) for the census block groups of the Vineland – Bridgeton, NJ MSA

5.6: Tuscaloosa, AL MSA

The map of the median commute time to work for the census block groups of this MSA (Figure 38), indicates the shortest commutes are taken by those living in Tuscaloosa. Outside of Tuscaloosa, the commute times are typically longer, but with no generalizable trend. Using these median commute times to work as the maximum travel time for which an IHE is deemed accessible, the number of accessible IHEs per census block group was tabulated.

The number of accessible IHEs of any type is shown in Figure 39. Four IHEs exists within the MSA, whereas ten additional IHEs outside of the MSA are accessible to some census block groups. Figure 40 through Figure 45 illustrate the number of accessible IHEs by IHE type. There is more access to predominantly bachelor's degree-granting (Figure 42) institutions than associate's-degree granting institutions (Figure 41).

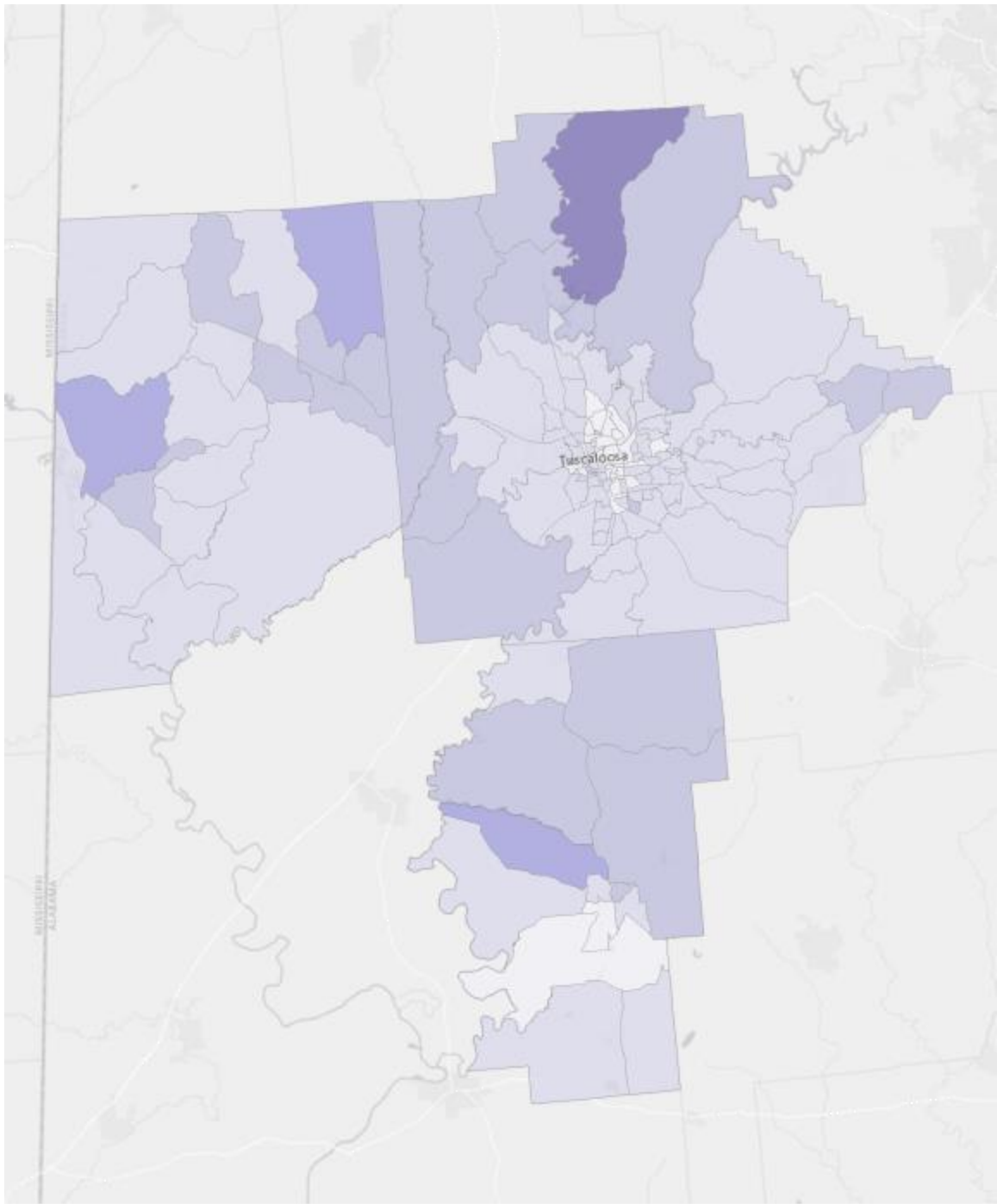
Table 10 and Table 11 summarize the number of accessible IHEs (by type) for demographic and SES factors. Table 10 summarizes by calculating the median number of accessible IHEs. Table 11 summarizes by calculating the average number of accessible IHEs. The amount of access to IHEs is generally consistent between demographic factors with the notable exception of the lack of access to predominantly associate's degree-granting IHEs for the Asian and Hispanic / Latino population. Between SES factors, the amount of access to IHEs is unequal. The census block groups that compose the bottom quartile of median house value, have considerably fewer accessible IHEs than the census block groups that compose the top quartile of median house value. The same trend is observable between the bottom and top quartile of median household income but the difference is less.

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		3	3	4	3	3	3	3	3	1	3
Predominant Degree	Certificate	1	1	1	1	1	1	1	1	0	1
	Associate's	1	0	1	0	1	0	0	1	0	0
	Bachelor's	2	2	2	2	2	2	2	2	0	2
Institution Control	Public	1	1	2	1	1	1	1	1	1	1
	Private Not-For-Profit	1	1	1	1	1	1	1	1	0	1
	Private For-Profit	1	1	1	1	1	1	1	1	0	1

Table 10: Median number of accessible IHEs (by type) for demographic and SES factors of the Tuscaloosa, AL MSA

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		2.95	2.95	2.94	3.22	2.94	3.05	2.70	3.27	1.64	3.42
Predominant Degree	Certificate	0.79	0.83	0.73	0.97	0.79	0.86	0.70	0.93	0.44	1.02
	Associate's	0.53	0.47	0.66	0.39	0.54	0.44	0.49	0.53	0.38	0.44
	Bachelor's	1.62	1.65	1.55	1.86	1.62	1.75	1.51	1.81	0.81	1.96
Institution Control	Public	1.30	1.24	1.43	1.36	1.30	1.30	1.24	1.33	0.83	1.38
	Private Not-For-Profit	0.88	0.93	0.78	0.89	0.88	0.89	0.75	1.10	0.37	1.08
	Private For-Profit	0.76	0.78	0.72	0.96	0.76	0.86	0.70	0.84	0.44	0.96

Table 11: Average number of accessible IHEs (by type) for demographic and SES factors of the Tuscaloosa, AL MSA



Median Commute Time to Work < 15 15 - 30 30 - 45 45 - 60 60 - 90 > 90

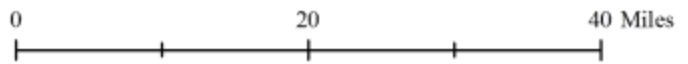


Figure 38: Median commute time to work for the census block groups of the Tuscaloosa, AL MSA

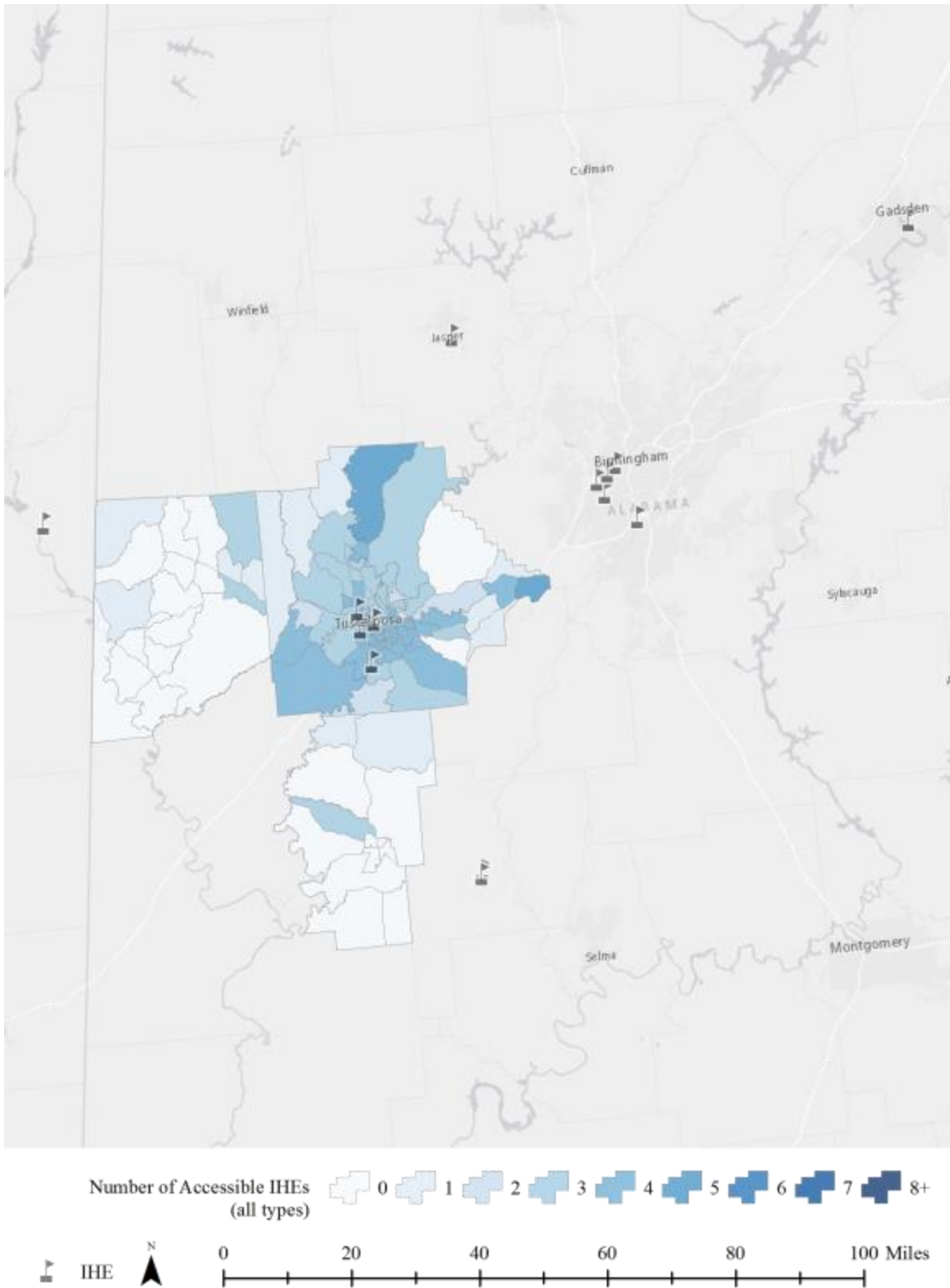


Figure 39: Number of accessible IHEs (all types) for the census block groups of the Tuscaloosa, AL MSA

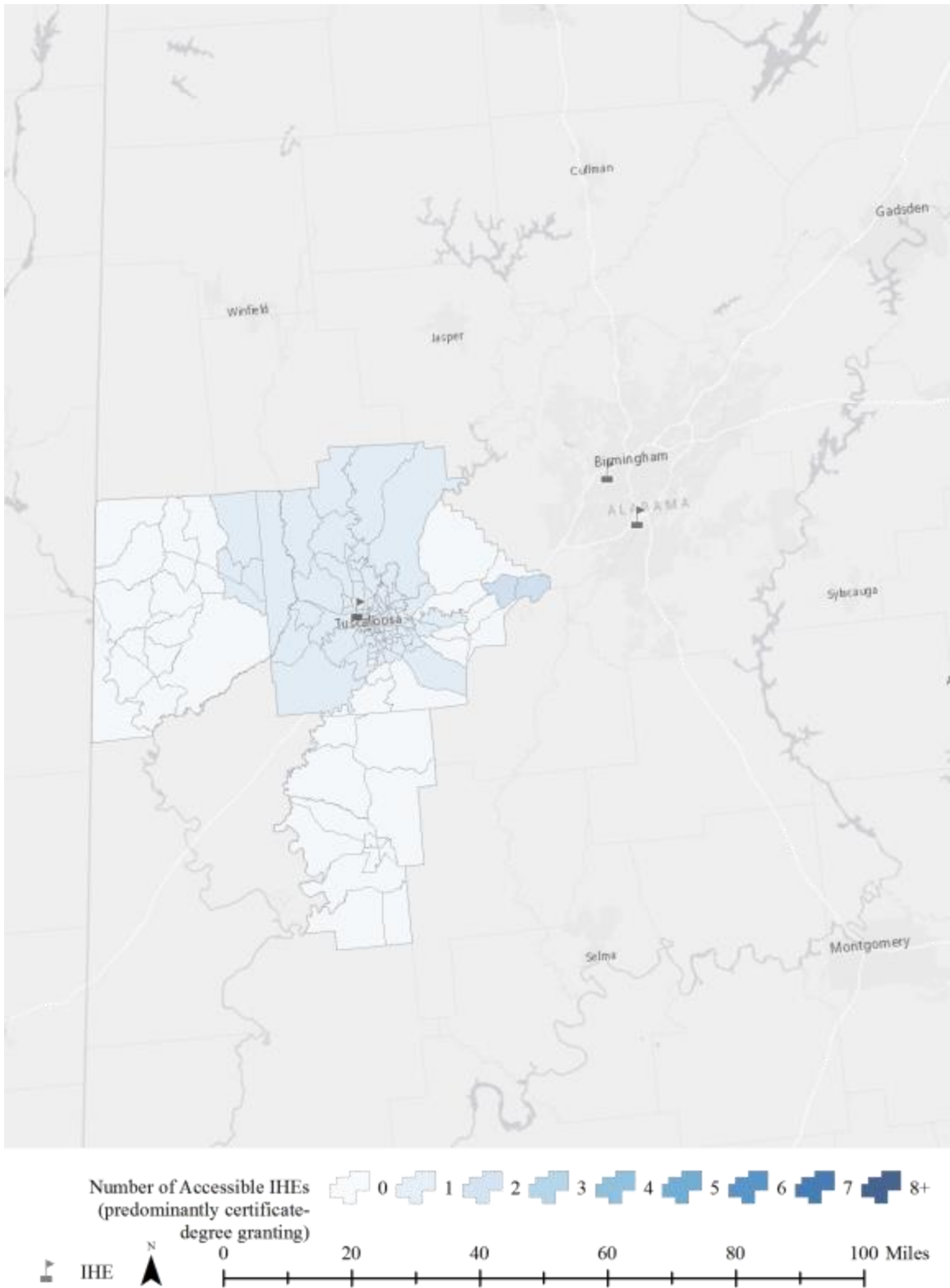


Figure 40: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Tuscaloosa, AL MSA

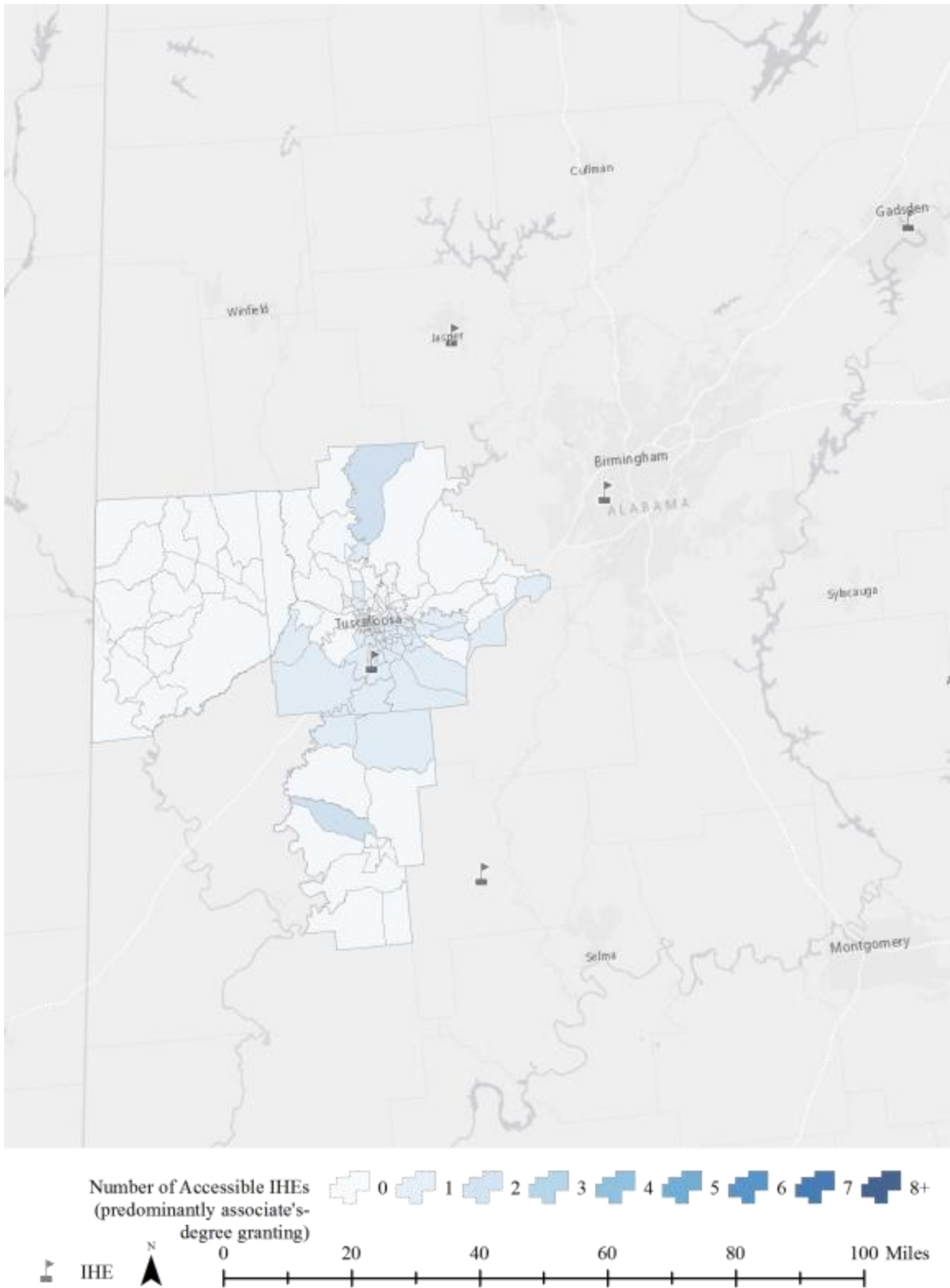


Figure 41: Number of accessible IHEs (predominantly associate's-degree granting) for the census block groups of the Tuscaloosa, AL MSA

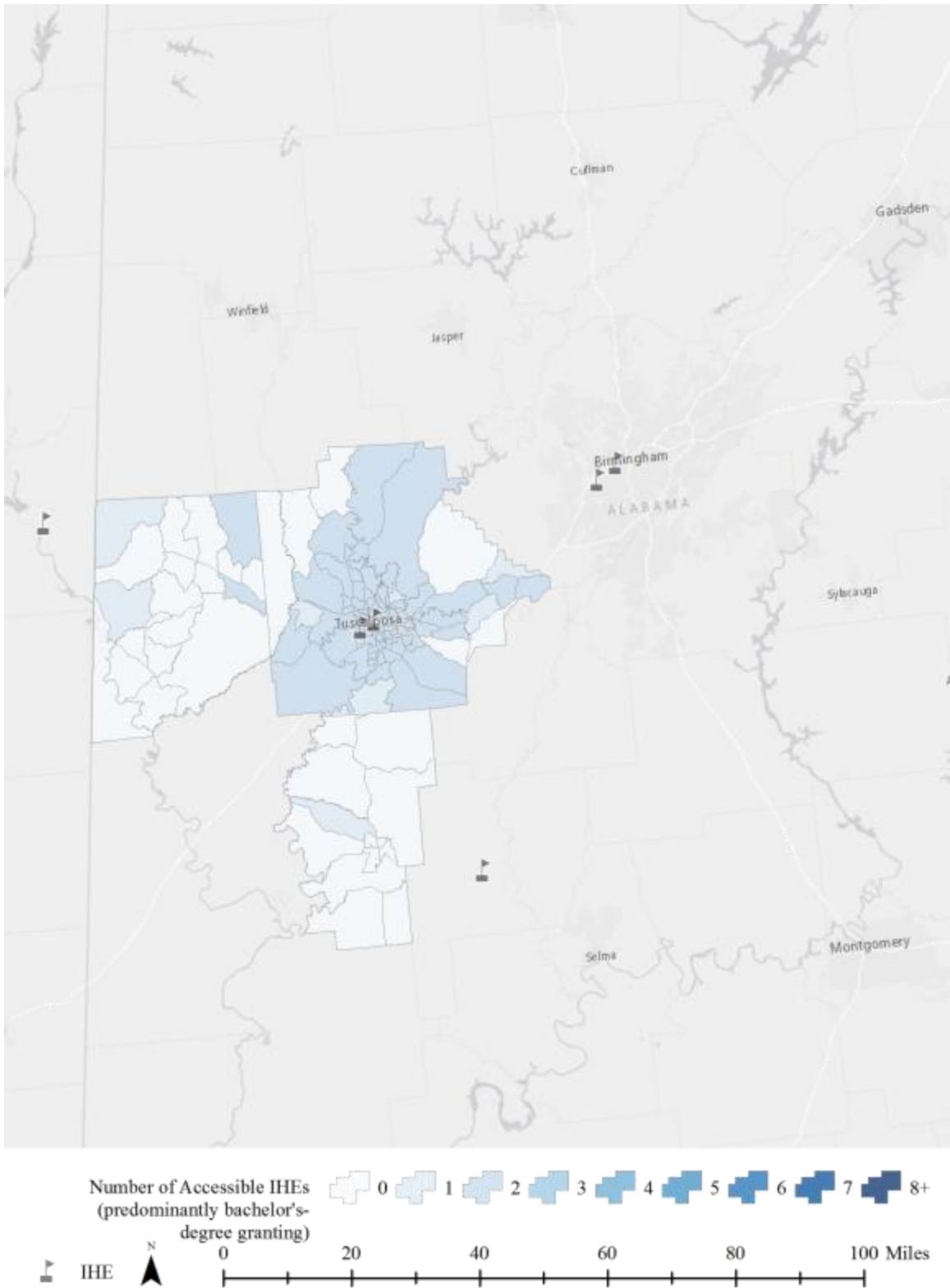


Figure 42: Number of accessible IHEs (predominantly bachelor's-degree granting) for the census block groups of the Tuscaloosa, AL MSA

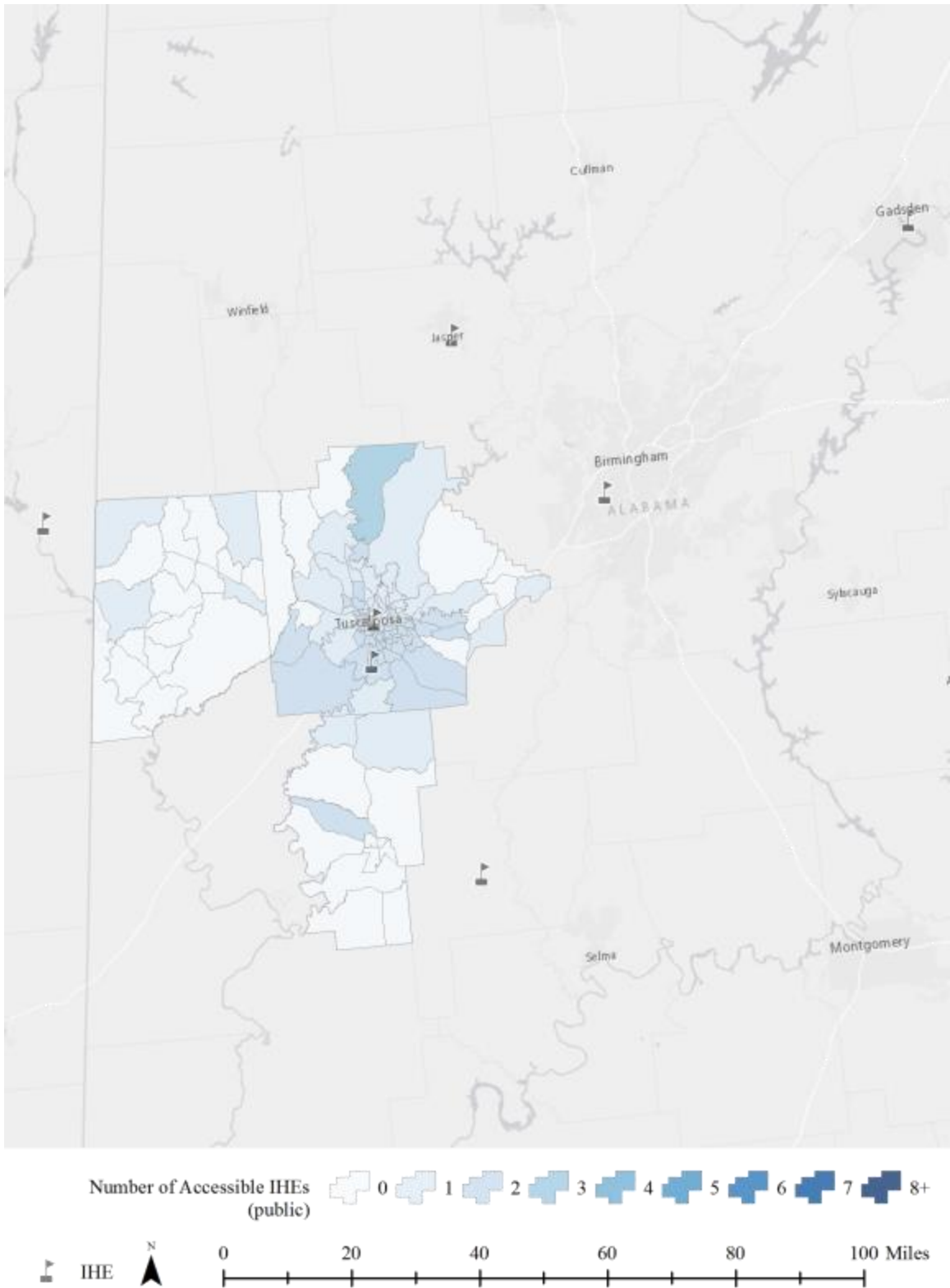


Figure 43: Number of accessible IHEs (public) for the census block groups of the Tuscaloosa, AL MSA

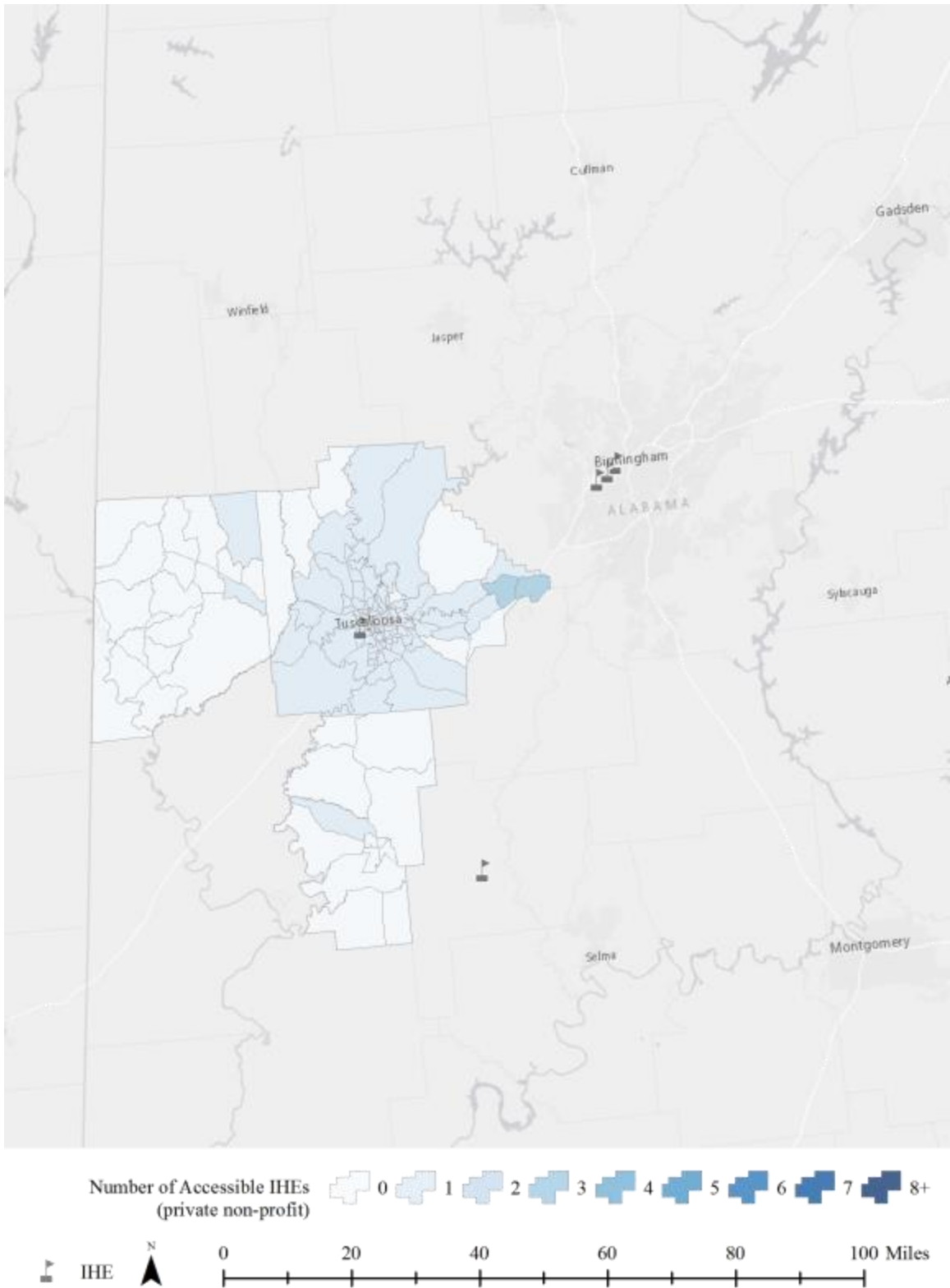


Figure 44: Number of accessible IHEs (private non-profit) for the census block groups of the Tuscaloosa, AL MSA

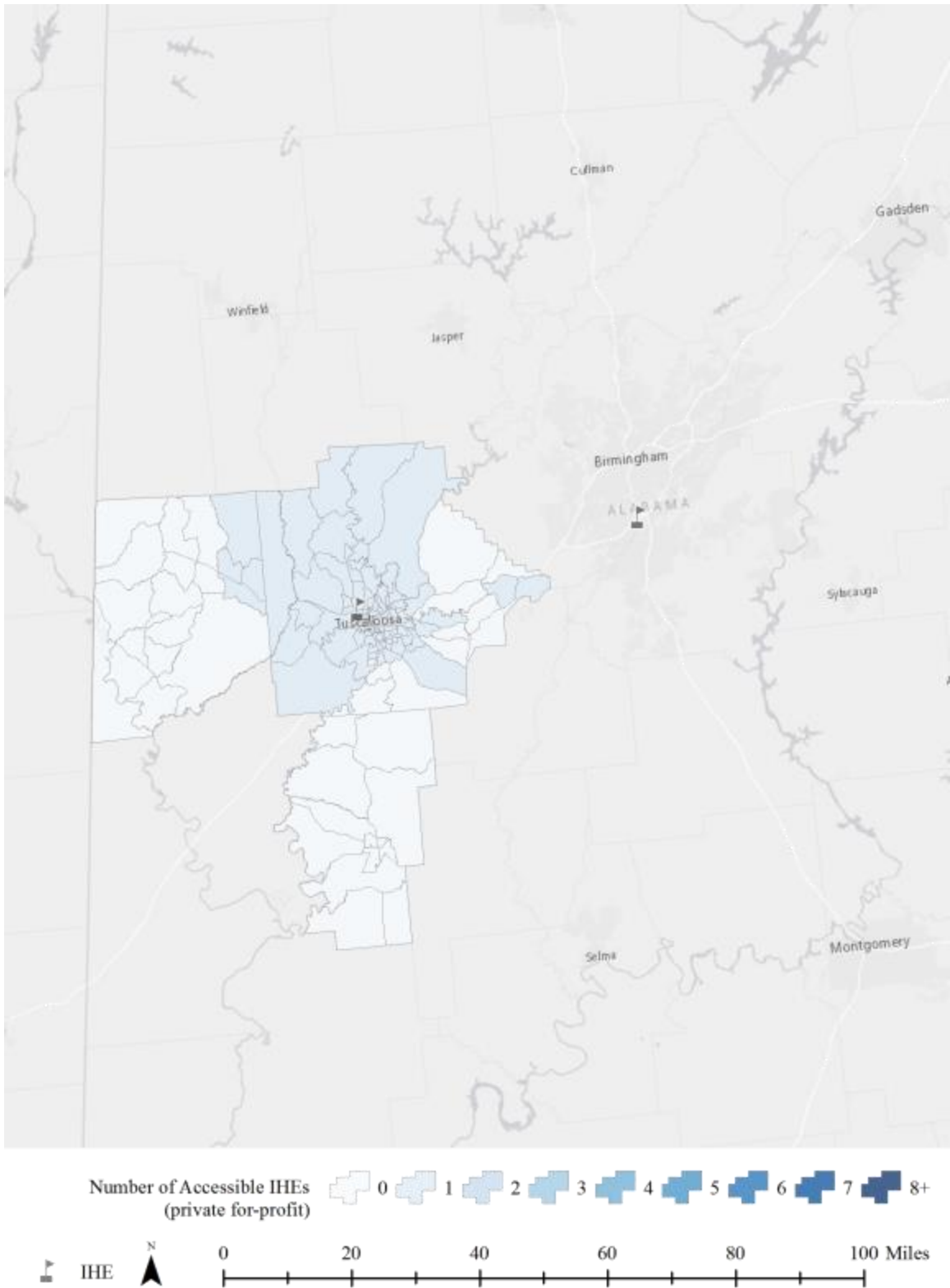


Figure 45: Number of accessible IHEs (private for-profit) for the census block groups of the Tuscaloosa, AL MSA

5.7: Riverside – San Bernardino – Ontario, CA MSA

The map of the median commute time to work for the census block groups of the this MSA (Figure 46), indicates longer commutes are taken by those living in the more urban census block groups of the MSA. There are pockets of shorter commutes in these more urban areas, but overall the extremely large, rural census block groups have a shorter commute. This perhaps is because of traffic within the urban areas. Using these median commute times to work as the maximum travel time for which an IHE is deemed accessible, the number of accessible IHEs per census block group was tabulated.

The number of accessible IHEs of any type is shown in Figure 47. There are a sizable number of accessible IHEs within the MSA, particularly concentrated around Riverside and Palm Deserts. However, outside the MSA, in the Los Angeles and San Diego areas, there are many more IHEs accessible to some census block groups of the MSA. Figure 48 through Figure 53 illustrate the number of accessible IHEs by IHE type. The number of accessible certificate-degree granting IHEs (Figure 48) is greater than the number of accessible associate's-degree granting IHEs (Figure 49) and the number of accessible bachelor's-degree granting IHEs (Figure 50).

Table 12 and Table 13 summarize the number of accessible IHEs (by type) for demographic and SES factors. Table 12 summarizes by calculating the median number of accessible IHEs. Table 13 summarizes by calculating the average number of accessible IHEs. These tables indicate the number of accessible predominantly associate's- and bachelor's-degree granting IHEs is approximately equal between races / ethnicities (with the exception of Asian populations having greater access) whereas the number of accessible predominantly certificate-degree granting IHEs vary considerably between races / ethnicities (minorities have greater access). Lower SES census block

groups have fewer accessible IHEs (of all types) compared to the higher SES census block groups, particularly among predominantly bachelor’s-degree granting IHEs.

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		22	18	23	31	18	24	11	30	5	32
Predominant Degree	Certificate	12	11	14	17	10	14	7	17	4	17
	Associate’s	4	4	4	5	4	4	2	5	1	5
	Bachelor’s	4	3	4	8	3	5	2	8	0	11
Institution Control	Public	5	4	5	6	4	5	3	6	1	6
	Private Not-For-Profit	4	3	4	6	3	4	2	7	1	8
	Private For-Profit	13	11	14	18	11	15	7	18	3	18

Table 12: Median number of accessible IHEs (by type) for demographic and SES factors of the Riverside – San Bernardino – Ontario, CA MSA

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		27.81	26.01	28.07	42.78	27.40	28.23	16.47	42.67	13.56	44.34
Predominant Degree	Certificate	15.50	14.51	15.65	23.65	15.28	15.73	9.53	23.44	7.89	24.24
	Associate’s	4.91	4.65	5.18	6.85	4.85	4.97	3.25	7.00	2.84	6.85
	Bachelor’s	7.40	6.85	7.25	12.28	7.27	7.53	3.69	12.23	2.84	13.25
Institution Control	Public	5.91	5.56	6.11	8.90	5.86	5.97	3.63	8.75	3.16	8.88
	Private Not-For-Profit	5.79	5.36	5.69	9.49	5.71	5.87	3.04	9.53	2.39	10.23
	Private For-Profit	16.10	15.09	16.27	24.39	15.83	16.38	9.80	24.39	8.01	25.23

Table 13: Average number of accessible IHEs (by type) for demographic and SES factors of the Riverside – San Bernardino – Ontario, CA MSA

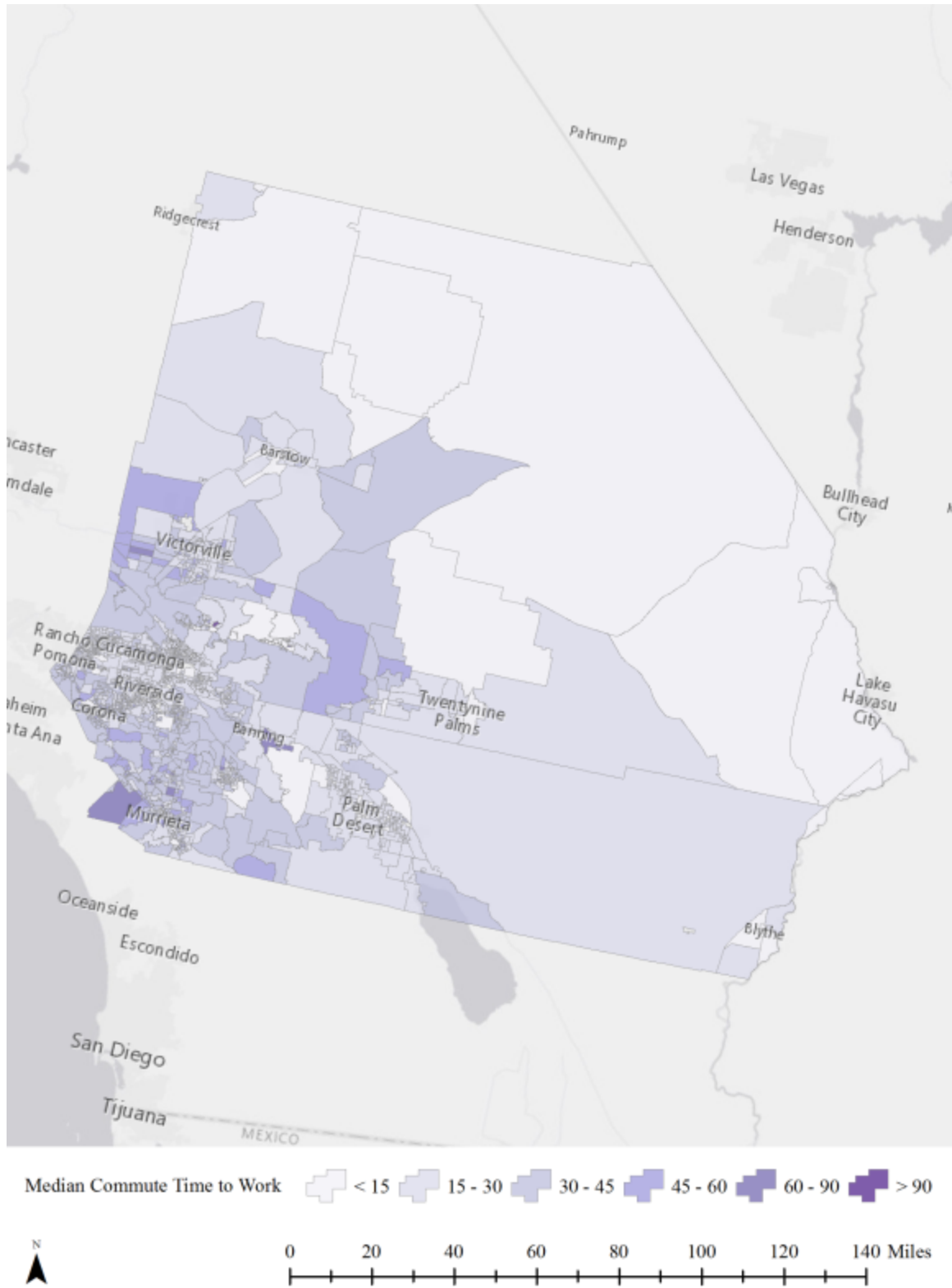
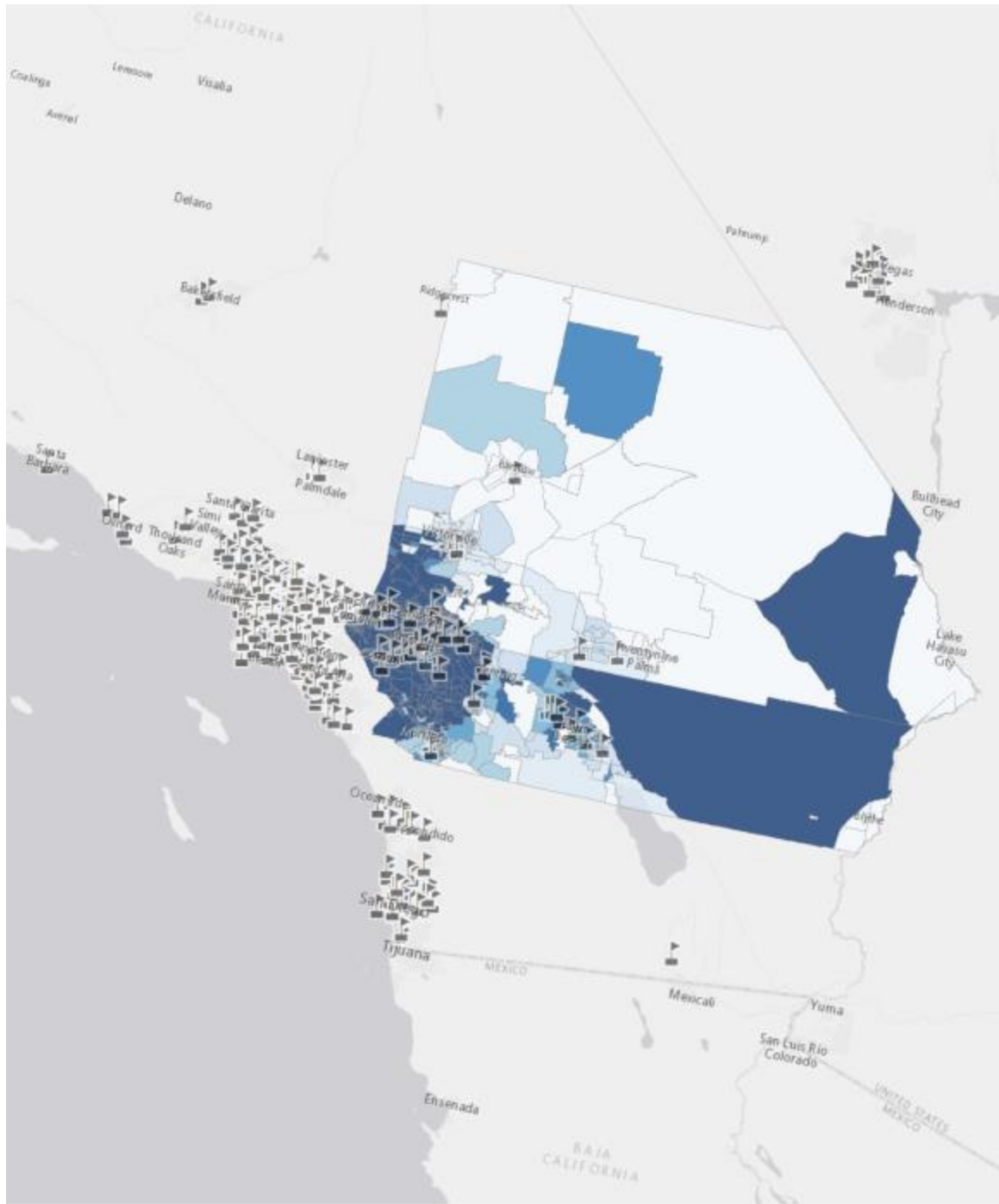


Figure 46: Median commute time to work for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA



Number of Accessible IHEs (all types)

0	1	2	3	4	5	6	7	8+
---	---	---	---	---	---	---	---	----

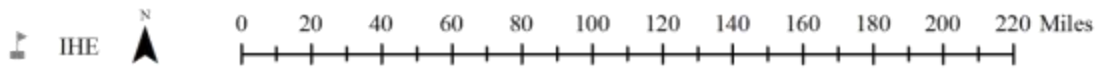


Figure 47: Number of accessible IHEs (all types) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA

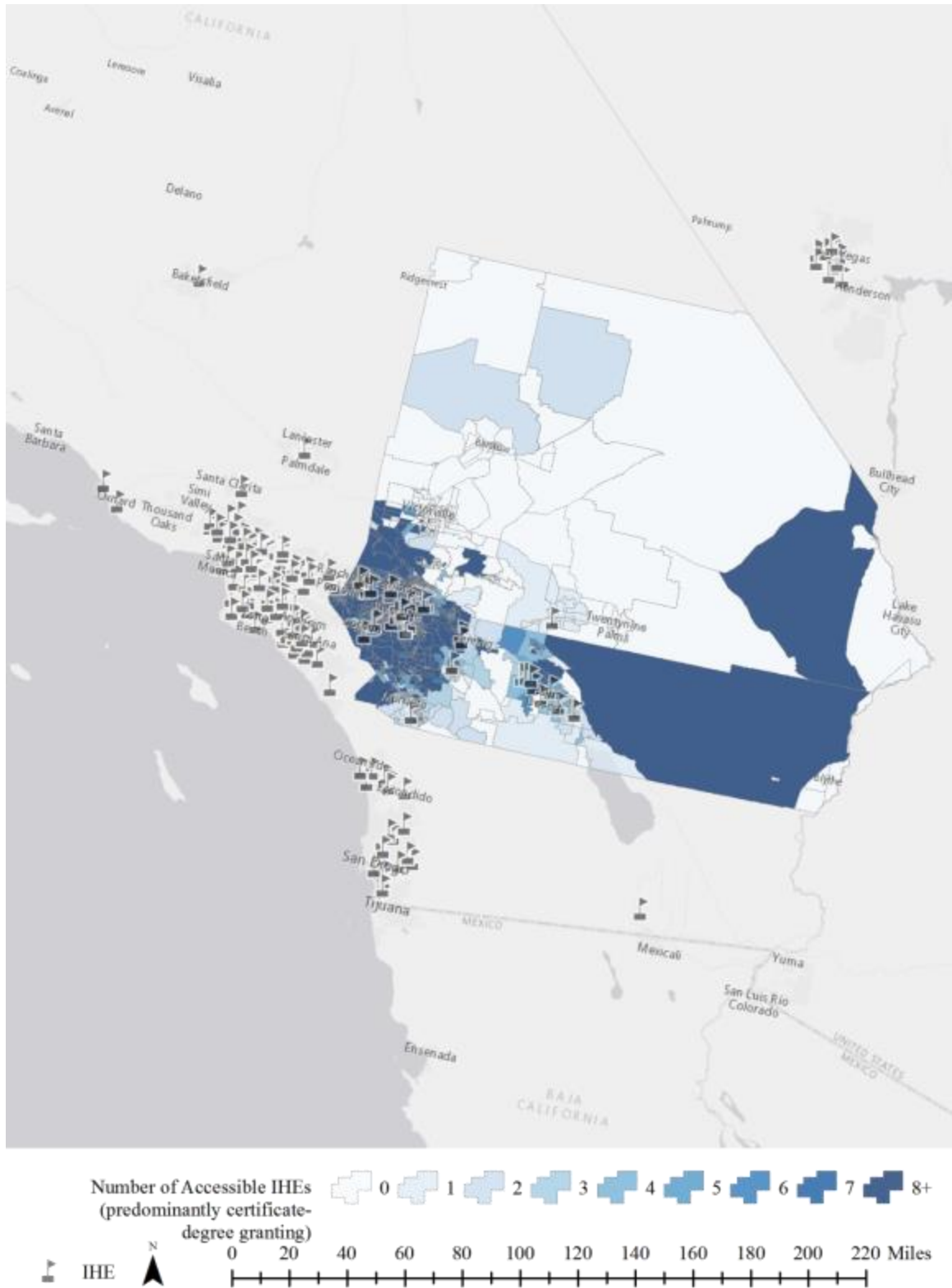


Figure 48: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA

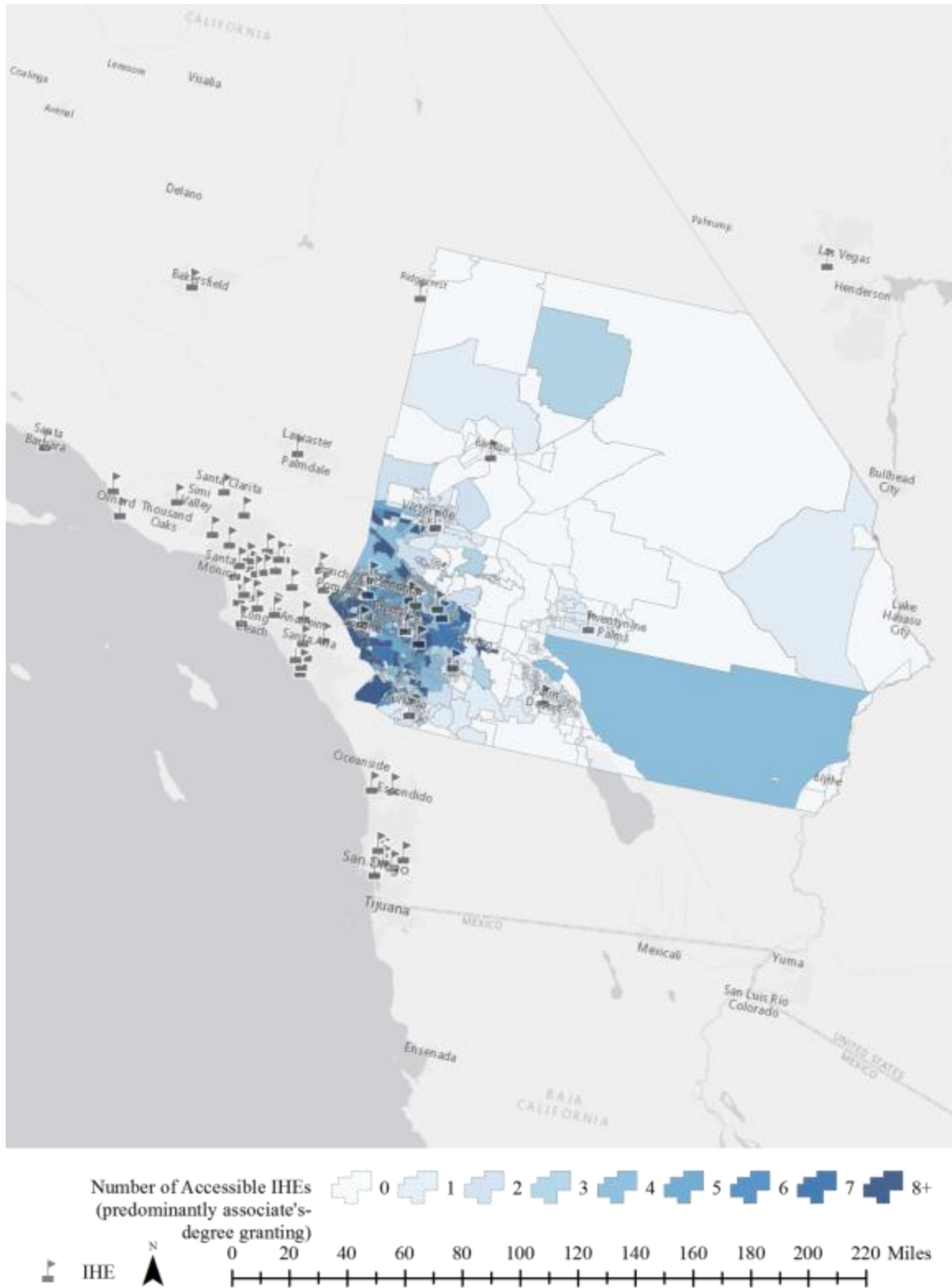


Figure 49: Number of accessible IHEs (predominantly associate's-degree granting) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA

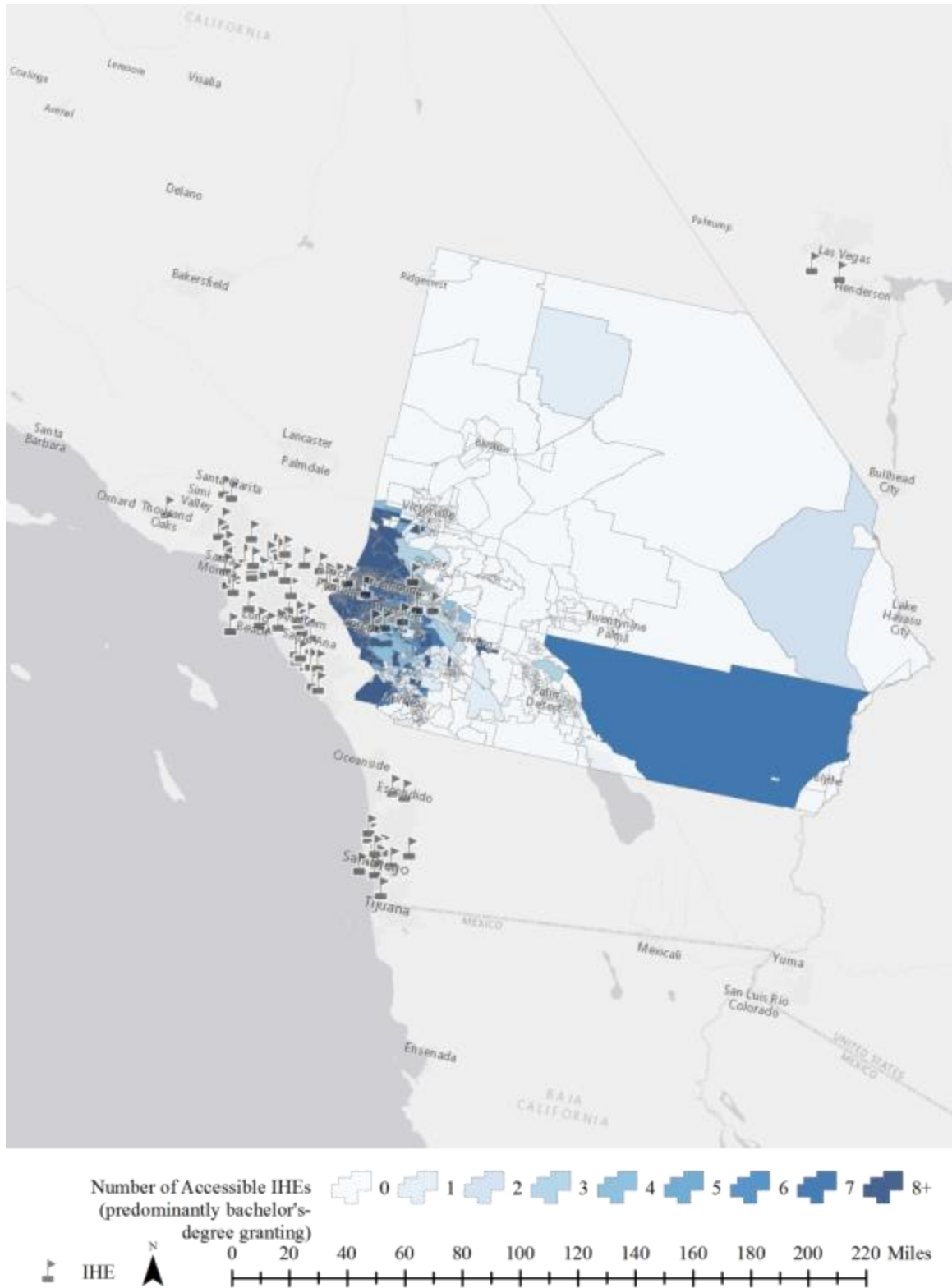


Figure 50: Number of accessible IHEs (predominantly bachelor's-degree granting) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA

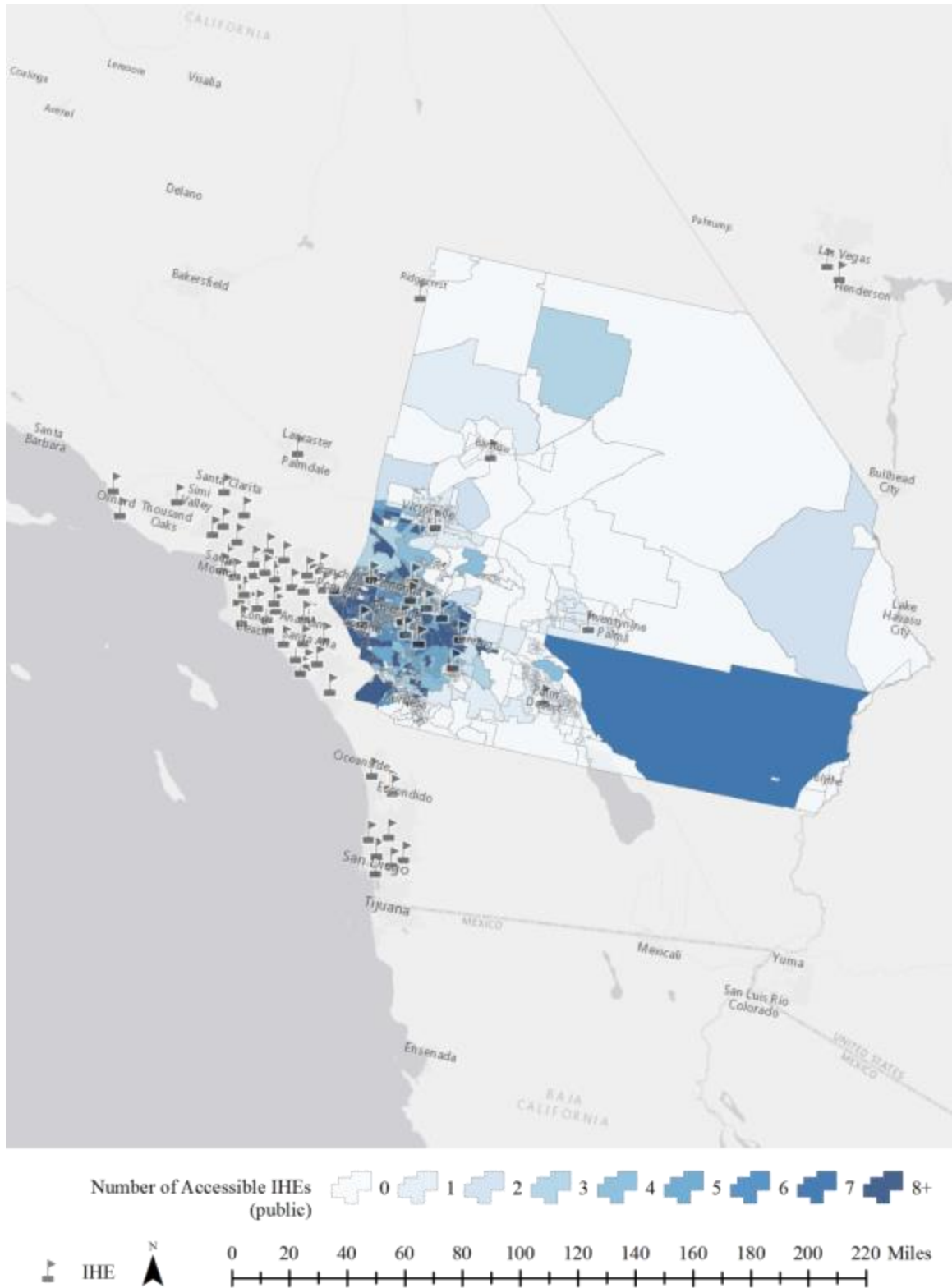


Figure 51: Number of accessible IHEs (public) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA

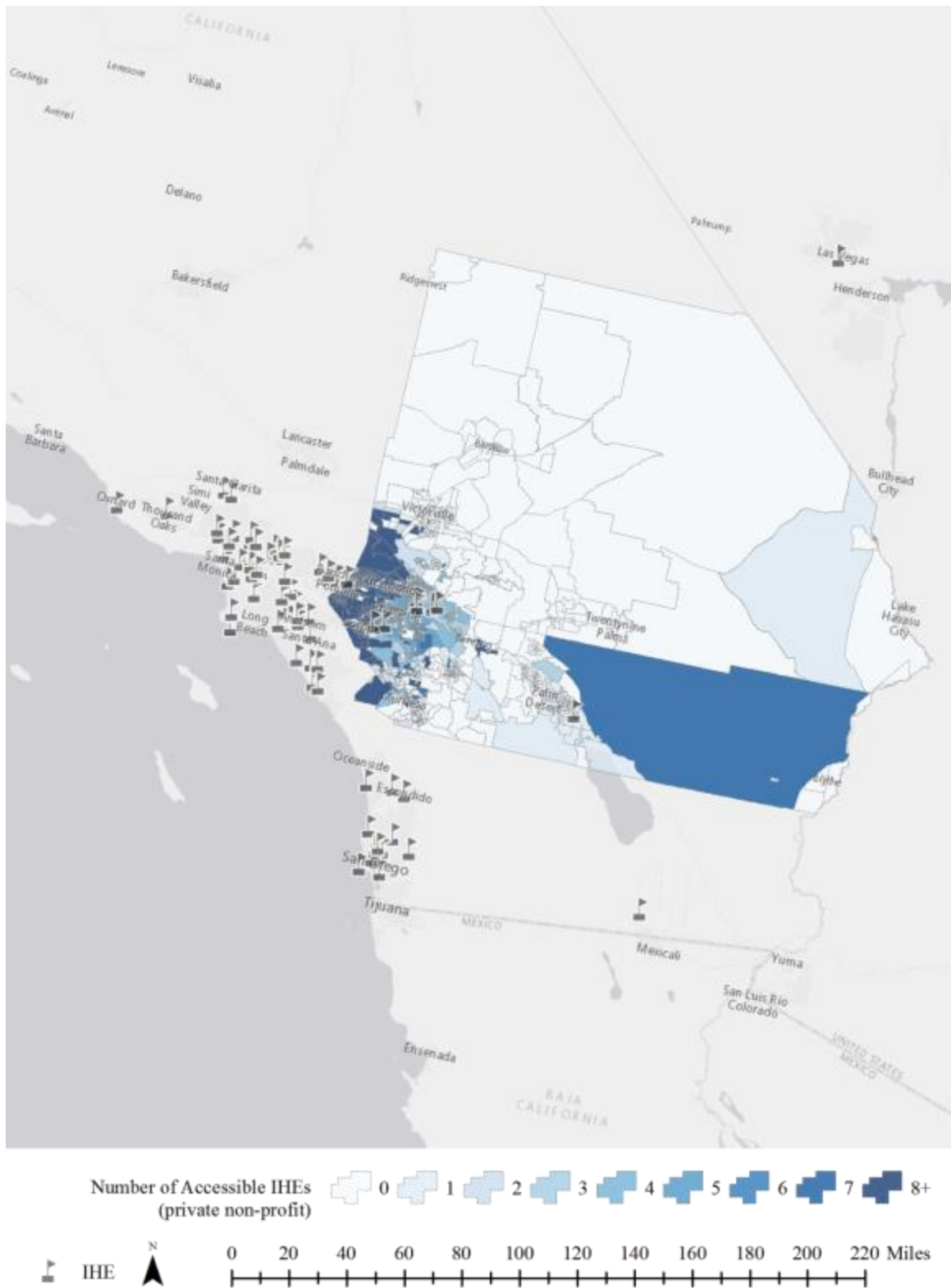


Figure 52: Number of accessible IHEs (private non-profit) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA

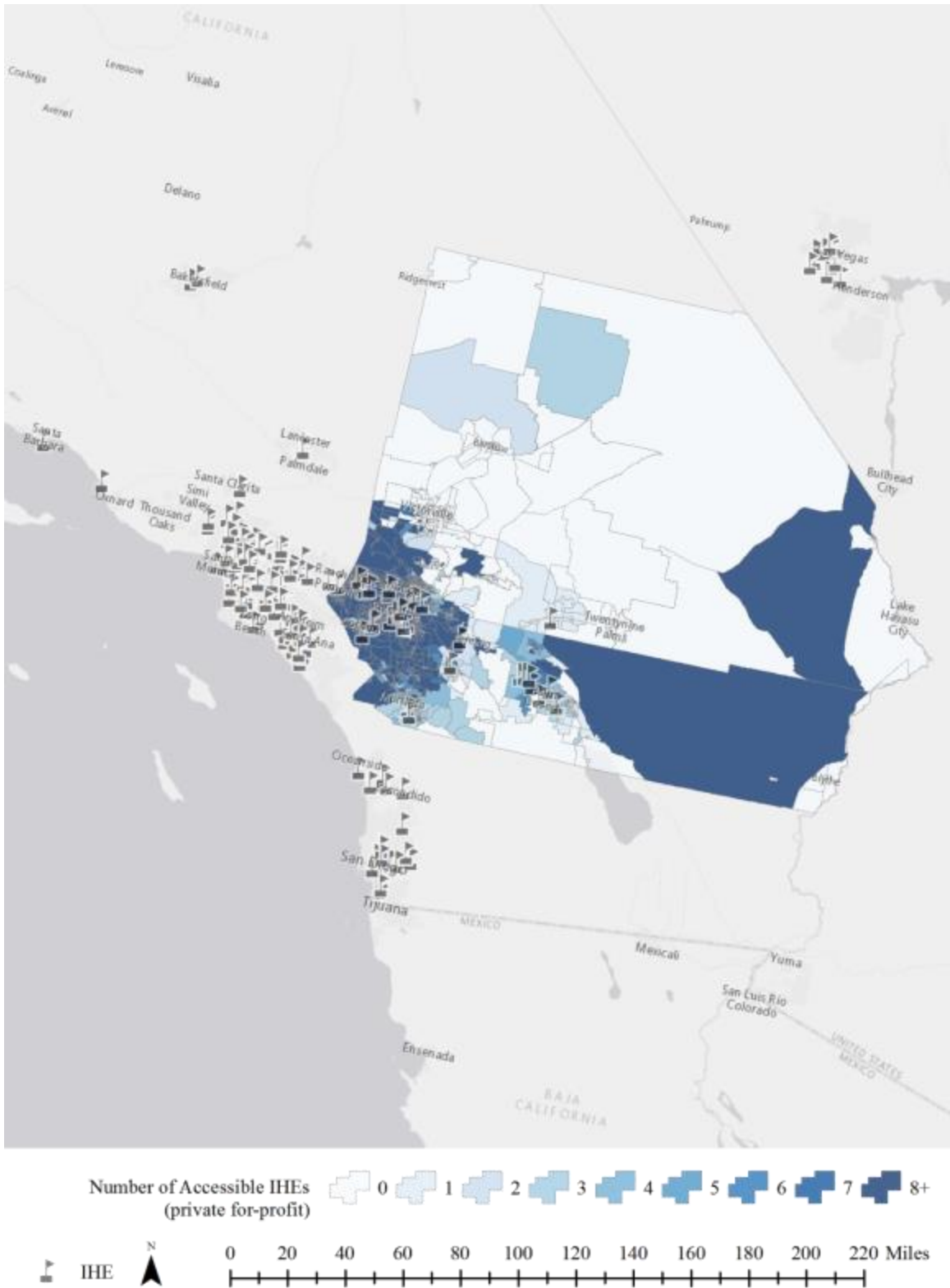


Figure 53: Number of accessible IHEs (private for-profit) for the census block groups of the Riverside – San Bernardino – Ontario, CA MSA

5.8: Sacramento – Roseville – Arden – Arcade, CA MSA

The map of the median commute time to work for the census block groups of the this MSA (Figure 46), indicates longer commutes are taken by those living in the more urban census block groups of the MSA. There are pockets of shorter commutes in these more urban areas, but overall the extremely large, rural census block groups have a shorter commute. This perhaps is because of traffic within the urban areas. Using these median commute times to work as the maximum travel time for which an IHE is deemed accessible, the number of accessible IHEs per census block group was tabulated.

The number of accessible IHEs of any type is shown in Figure 47. The majority of accessible IHEs are within the MSA although some accessible IHEs are located outside the MSA. Figure 48 through Figure 53 illustrate the number of accessible IHEs by IHE type. The number of accessible certificate-degree granting IHEs (Figure 56) is greater than the number of accessible associate's-degree granting IHEs (Figure 57) and the number of accessible bachelor's-degree granting IHEs (Figure 58).

Table 14 and Table 15 summarize the number of accessible IHEs (by type) for demographic and SES factors. Table 14 summarizes by calculating the median number of accessible IHEs. Table 15 summarizes by calculating the average number of accessible IHEs. These tables show the number of accessible predominantly associate's- and bachelor's-degree granting IHEs to be rather consistent between all the demographic and SES factors. There is variation in the number of accessible predominantly certificate-degree granting IHEs between the demographic and SES factors: racial / ethnic minorities and lower SES census block groups have greater access.

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		18	17	20	19	18	19	20	15	21	10
Predominant Degree	Certificate	12	11	14	13	12	13	13	10	14	6
	Associate's	3	3	3	3	3	3	3	3	3	2
	Bachelor's	3	3	3	3	3	3	3	2	3	2
Institution Control	Public	4	4	5	5	4	5	5	4	5	2
	Private Not-For-Profit	1	1	1	1	1	1	1	1	1	1
	Private For-Profit	13	11	14	13	12	13	14	10	14	6

Table 14: Median number of accessible IHEs (by type) for demographic and SES factors of the Sacramento – Roseville – Arden - Arcade, CA MSA

		Total MSA Population	White Population	Black Population	Asian Population	Not Hispanic / Latino Population	Hispanic / Latino Population	Bottom Quartile Median Household Income Population	Top Quartile Median Household Income Population	Bottom Quartile House Value Population	Top Quartile House Value Population
All IHEs		15.41	14.42	19.27	17.14	15.24	16.07	17.21	13.95	20.04	10.77
Predominant Degree	Certificate	10.14	9.47	12.79	11.25	10.02	10.57	11.34	9.16	13.29	6.87
	Associate's	2.68	2.52	3.25	2.96	2.65	2.78	2.96	2.46	3.39	2.05
	Bachelor's	2.60	2.43	3.23	2.93	2.57	2.72	2.91	2.34	3.36	1.84
Institution Control	Public	4.14	3.83	5.22	4.74	4.08	4.36	4.70	3.69	5.38	2.98
	Private Not-For-Profit	1.16	1.18	1.17	1.05	1.17	1.11	1.11	1.18	1.34	0.97
	Private For-Profit	10.12	9.41	12.88	11.36	9.99	10.60	11.40	9.08	13.32	6.82

Table 15: Average number of accessible IHEs (by type) for demographic and SES factors of the Sacramento – Roseville – Arden - Arcade, CA MSA

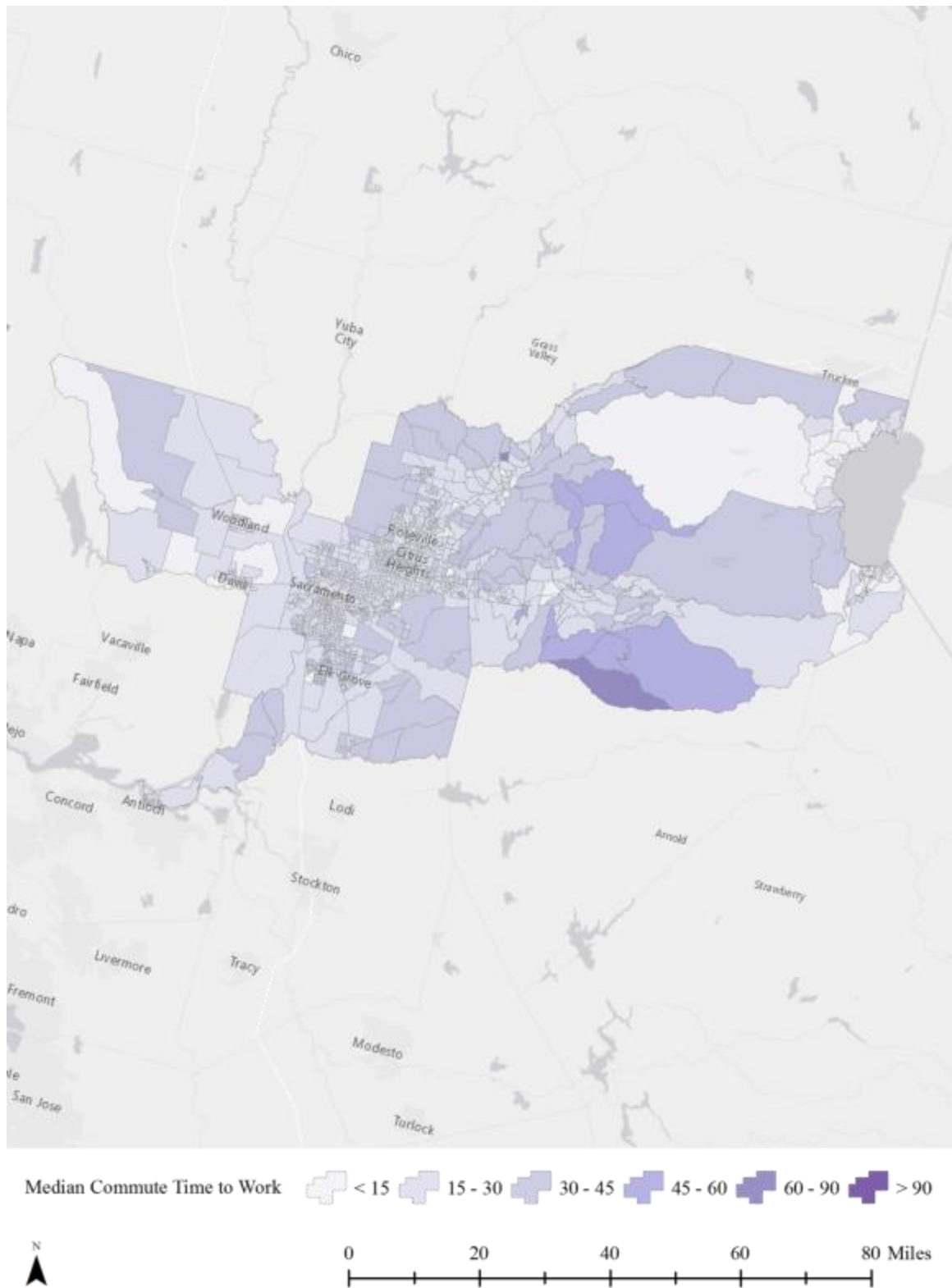


Figure 54: Median commute time to work for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA

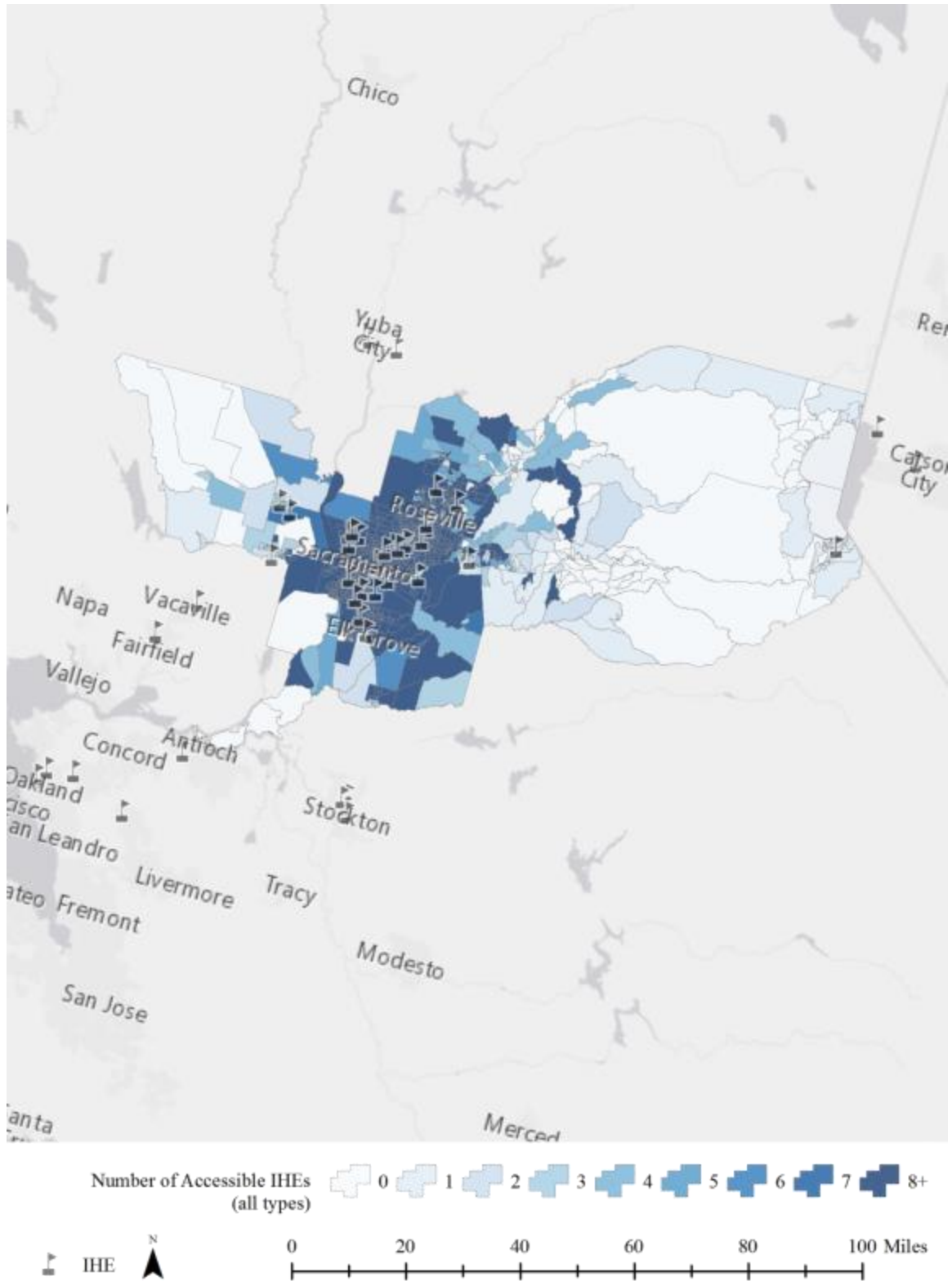


Figure 55: Number of accessible IHEs (all types) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA

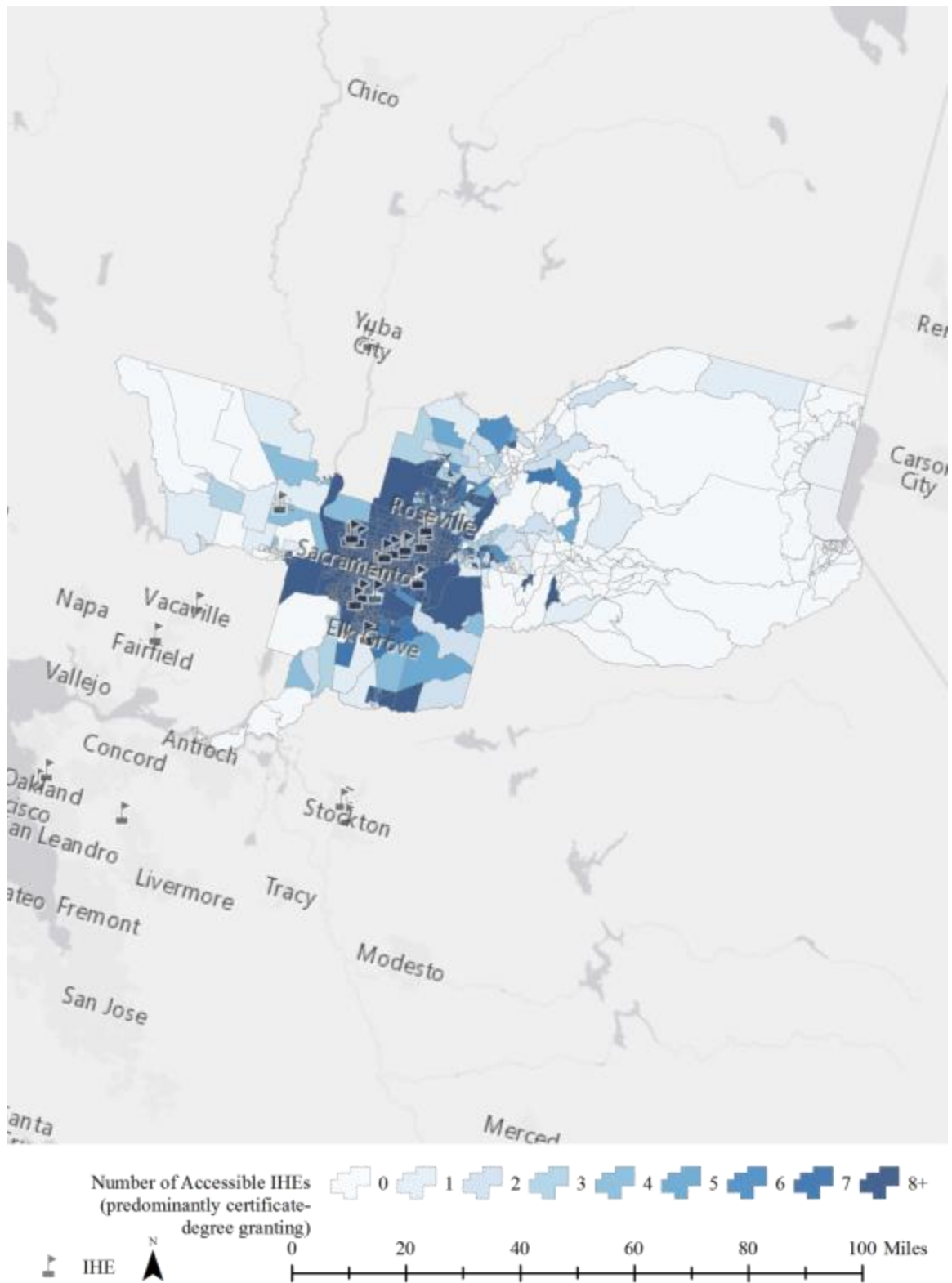


Figure 56: Number of accessible IHEs (predominantly certificate-degree granting) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA

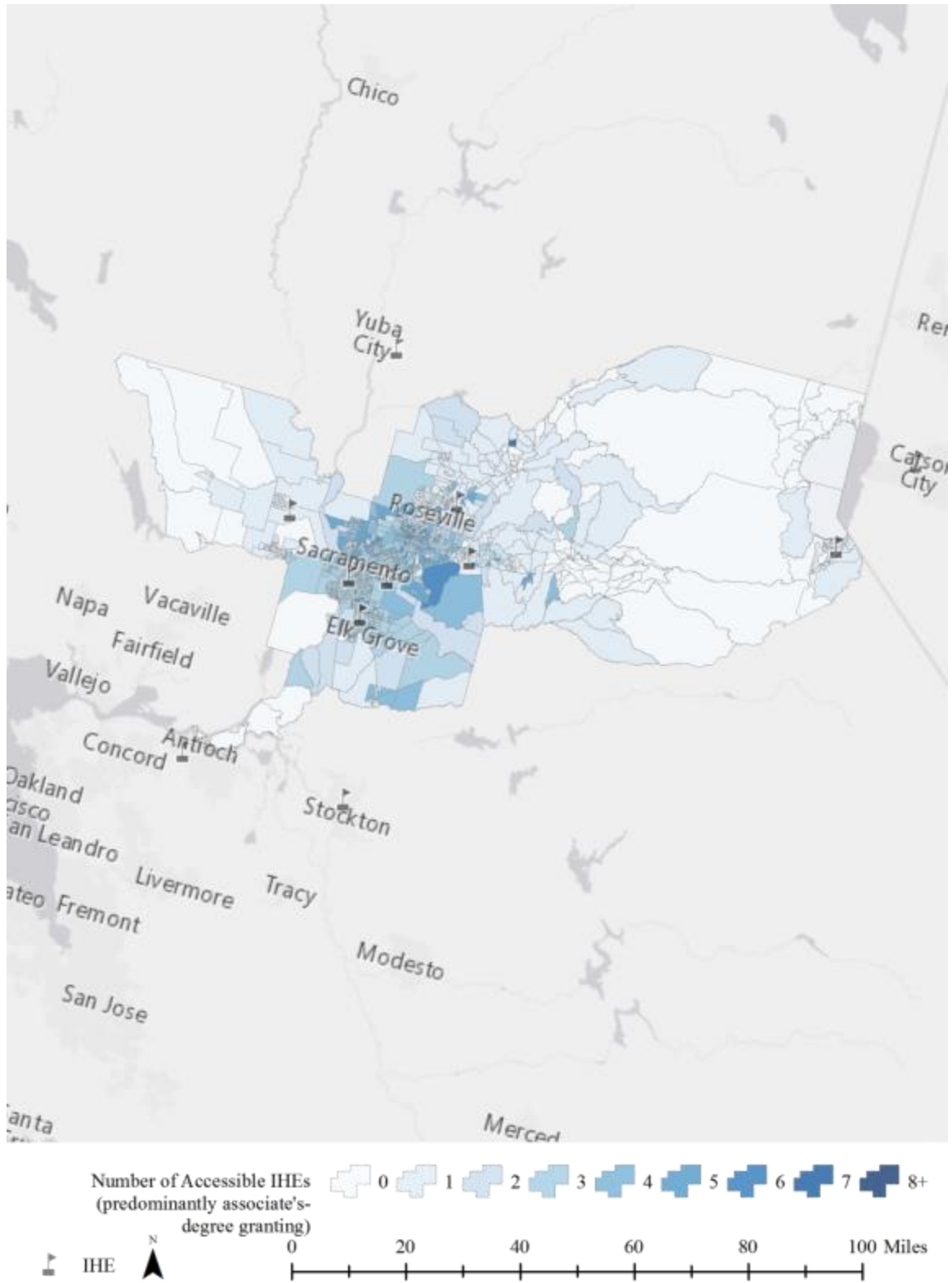


Figure 57: Number of accessible IHEs (predominantly associate's-degree granting) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA

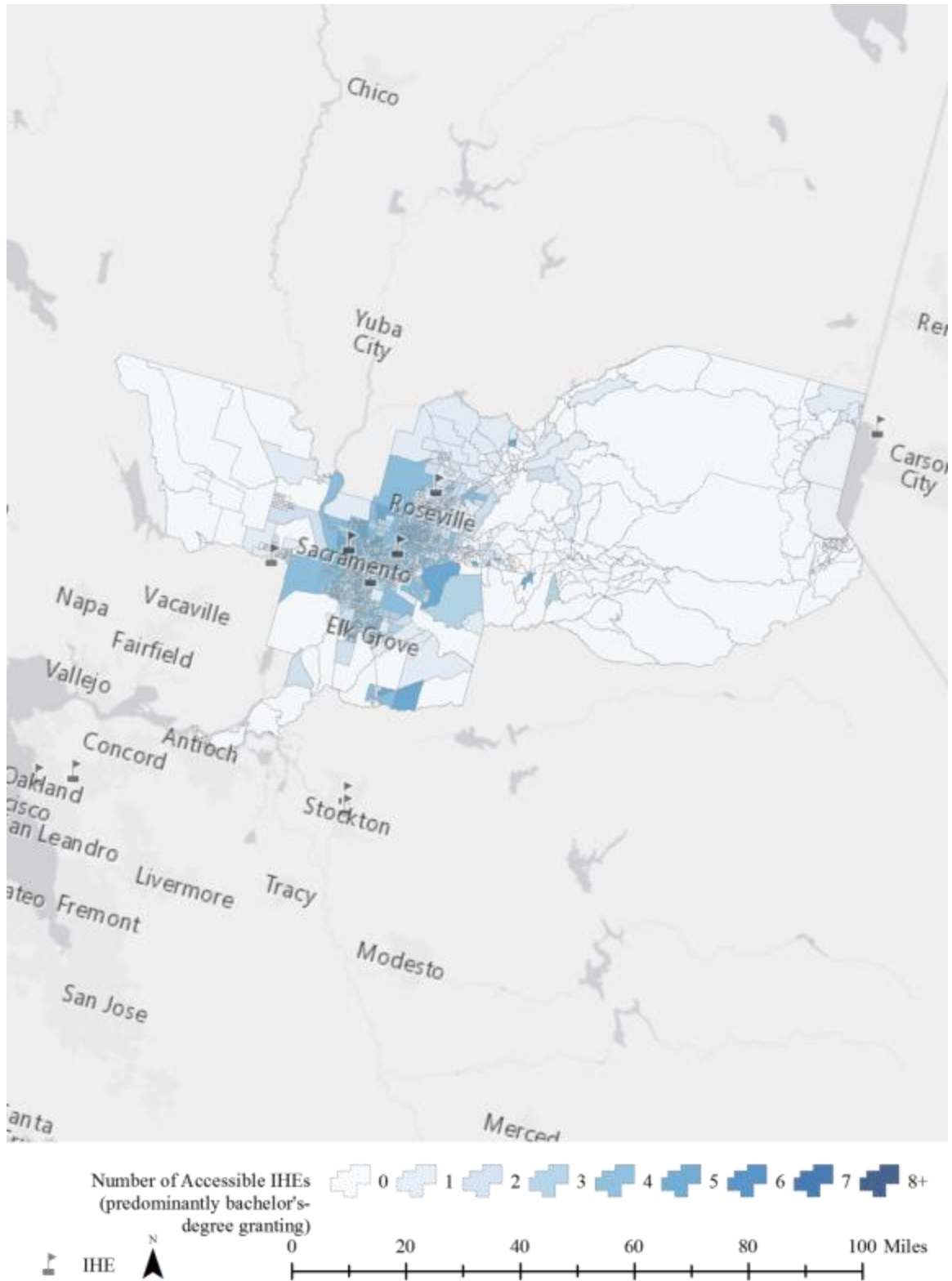


Figure 58: Number of accessible IHEs (predominantly bachelor's-degree granting) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA

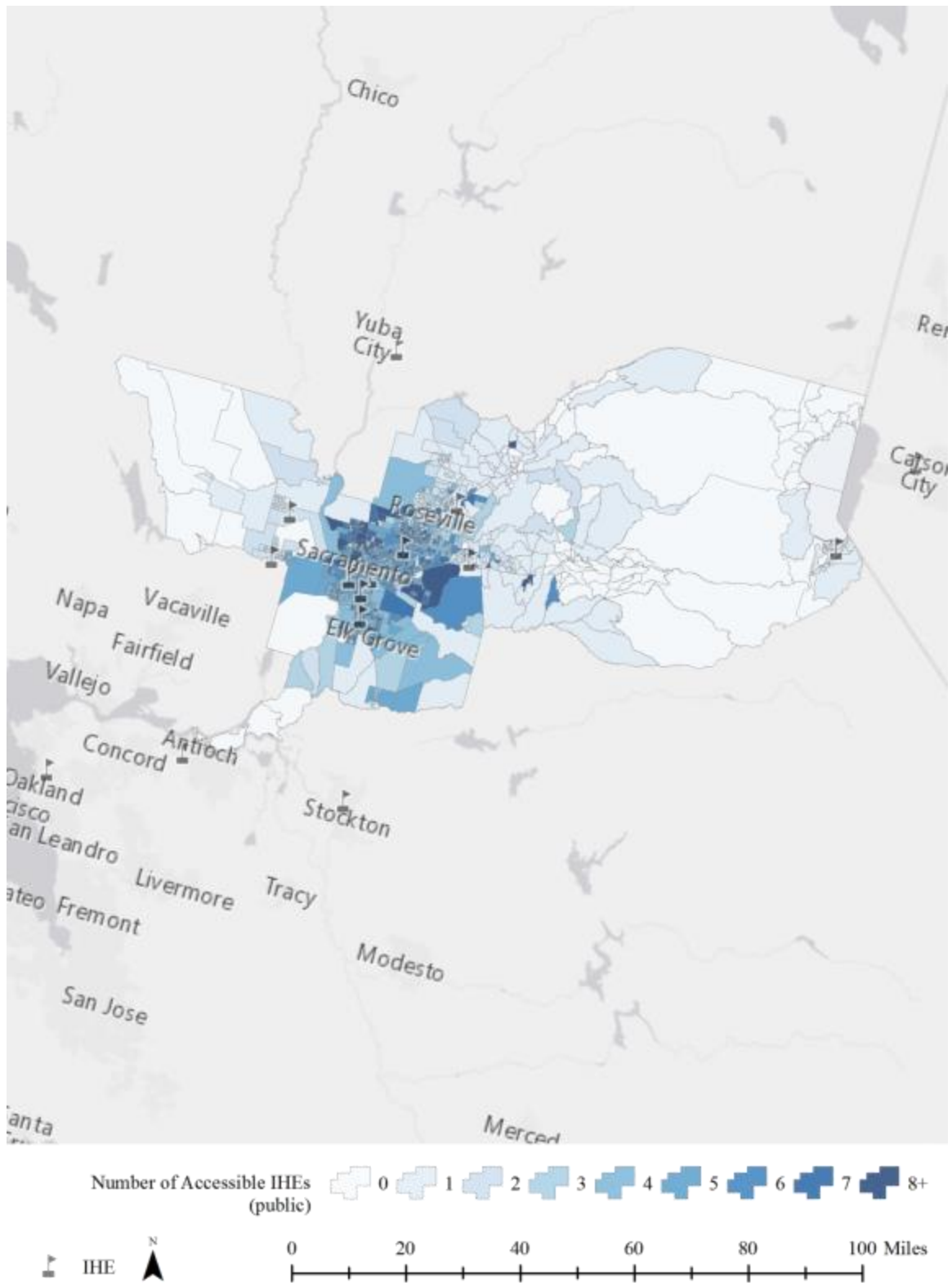


Figure 59: Number of accessible IHEs (public) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA

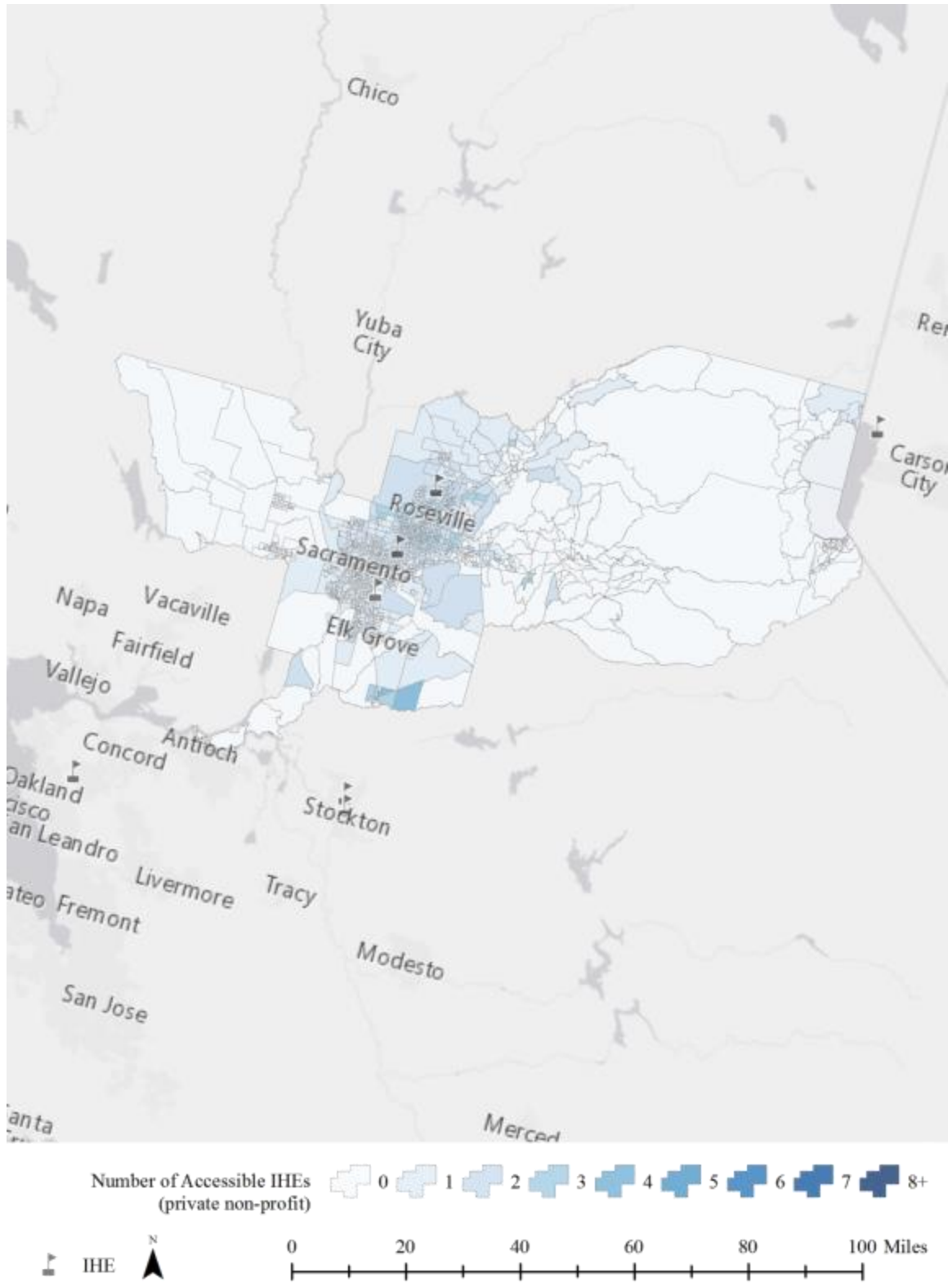


Figure 60: Number of accessible IHEs (private non-profit) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA

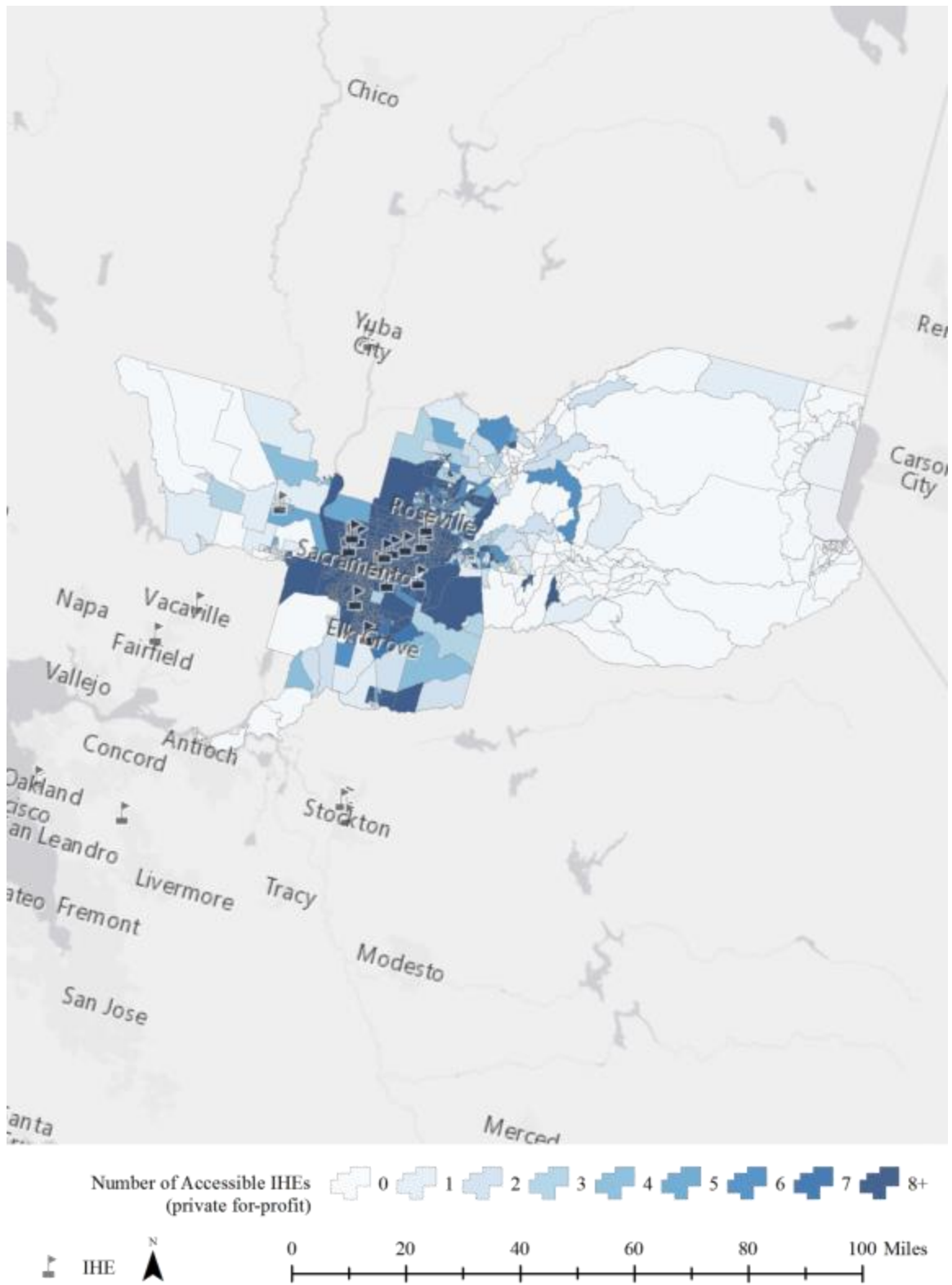


Figure 61: Number of accessible IHEs (private for-profit) for the census block groups of the Sacramento – Roseville – Arden - Arcade, CA MSA

Chapter 6: Conclusions

6.1: IHE access trends

In two of the three MSA pairs, the MSA with a higher percentage of college educated citizens had a higher number of accessible IHEs. This suggests that having more accessible IHEs may be correlated with higher percentages of college educated citizens. This could be the result of the local community having access to IHEs, graduating from them, and staying in the MSA or it could be caused by college educated citizens being attracted to the social and economic profile of MSAs with a high concentration of IHEs. In either case, this study does not attempt to prove causation.

Since the percentage of college educated citizens was calculated using the number of people 25 or older with a bachelor's degree or higher, it is appropriate to acknowledge that both the Dalton, GA MSA and the Vineland – Bridgeton, NJ MSA lacked access to any predominantly bachelor's-degree granting institution.

This trend was not observed in the pair of MSAs in California. However, in that particular case, the population of the Riverside – San Bernardino – Ontario MSA was almost twice the population of the Sacramento – Roseville – Arden – Arcade MSA. Therefore, the number of accessible IHEs is likely higher because of the larger population and not necessarily because of greater access. Despite these MSAs having similar demographic and economic compositions, their vast population difference makes directly comparing the two difficult.

In four of the six MSAs, the lower SES census block groups had fewer accessible IHEs than the upper SES census block groups. This suggests unequal access to IHE may be correlated to SES. In the Vineland – Bridgeton, NJ MSA there was the same median number of accessible IHEs between the different SES classifications. In the Sacramento

– Roseville – Arden – Arcade, CA MSA there were more accessible IHEs in the lower SES census block groups than the upper SES census block groups. In this particular MSA, the number of accessible IHEs was higher in the lower SES census block groups because of private for-profit predominately certificate-degree granting IHEs.

Differences in IHE access based upon race were inconsistent. In many cases, minorities such as Black, Asian, or Hispanic / Latino had a greater number of accessible IHEs compared to Whites. However, this may be an artifact of rural areas of the MSAs being predominantly white and the minority communities are clustered within the core of the MSA where the majority of IHEs are. In both California MSAs, minorities have a greater number of accessible IHEs compared to Whites because of private for-profit predominately certificate-degree granting IHEs. If a more racially segregated MSA, for example St. Louis, MO, was analyzed, the measure of IHE access may have differ greater among races.

6.2: Limitations of analysis

When the median travel time for each census block group was computed, the ACS data included all forms of transportation. When the travel time from each census block group to its nearby IHEs was computed, the travel was assumed to be taking place in a personal vehicle. Incorporating other methods of transportation, particularly public transportation, uniformly across the country would not be feasible. Only with sufficient data about a transit system can service areas be generated (O’Sullivan, Morrison, and Shearer 2000). Further nuancing computing travel time via public transportation, the scheduling of the routes must also be accounted for (Lei and Church 2010).

However, commuting by other modes such as public transportation cannot be any faster than personal vehicle. Therefore, since the median travel time includes these other

modes of transportation, the median travel time is larger than it would be if everyone was commuting by personal vehicle. The labelling of IHE as accessible or inaccessible would not be as accurate for those not commuting by personal vehicle. As mentioned earlier, the effect of this is minimal in the MSAs studied because each MSA in this study had over 91% of commuters commuting by personal vehicle.

The travel time threshold differentiating accessible and inaccessible was the same for all IHE types. This is a likely an unrealistic assumption. Prospective students may be willing to commute further for a predominantly bachelor's-degree granting IHE compared to a certificate-degree granting IHE. Other intuitional factors besides institutional type may also impact a prospective student's willingness to commute. For example, a student may be willing to commute further to an IHE despite having a closer otherwise *equivalent* option if the further IHE is cheaper, has more academic programs, or less selective admission standards.

6.3: Future research considerations

The types of results presented here are only one way of presenting and analyzing the underlying relationships between college access and college proximity. These results could be re-organized for additional analysis and presentation. For example, the nearest IHE could be calculated for each census block group. This type of re-analysis could assess the accusation that private for-profit IHEs are opening up en masse in predominantly minority, low-SES communities (Dache-Gerbino, Kiyama, and Sapp n.d.). Another additional analysis methodology could create a measure of college access based upon the travel to time the nearest predominantly certificate-, associates-, bachelor's degree-granting IHEs. This measure could be weighted so preference to a particular type of IHE could be given. Yet another additional analysis methodology could create a

binary measure of access based on a set of criteria. For example, all census block groups within a 20-minute drive from any predominantly associate's-degree granting IHE and within a 30-minute drive from any predominantly bachelor's-degree granting IHE is considered to have college access. All other census block groups not meeting those criteria would be considered to lack college access. With the custom GIS-based routines developed for this study, the college proximity data calculated can be related to measures of college access in a nearly infinite number of ways.

Beyond just the college proximity data calculated for these six MSAs, the developed methodology could be used to compute access to IHEs nationwide. The sample of six MSAs composed about 1.8% of the census block groups within the United States. Assuming computation time continues to scale linearly, computing service areas for all IHEs would take about 22 hours, geometrically correcting the service areas would take about 31 hours and intersecting the service areas with the census blocks to compute the population-adjusted average travel time would take about 467 hours. Overall, these estimates suggest a total run time of about 22 days. Parallelization is already utilized to decrease computation time, but utilizing graphical processing units and/or distributed computing would further decrease computation time (Tischler 2016; Worboys and Duckham 2004).

Bibliography

- Alwitt, L. F., and T. D. Donley. 1997. Retail Stores in Poor Urban Neighborhoods. *Journal of Consumer Affairs* 31 (1):139–164.
- Boscoe, F. P., K. A. Henry, and M. S. Zdeb. 2012. A Nationwide Comparison of Driving Distance Versus Straight-Line Distance to Hospitals. *The Professional Geographer* 64 (2). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3835347/>.
- Brown, M. G., D. Y. Wohn, and N. Ellison. 2016. Without a map: College access and the online practices of youth from low-income communities. *Computers & Education* 92–93:104–116.
- Carnevale, A. P., S. J. Rose, and B. Cheah. 2013. *The college payoff: Education, occupations, lifetime earnings*. Georgetown University Center on Education and the Workforce. <https://repository.library.georgetown.edu/handle/10822/559300>.
- Carroll, S. J., and E. Erkut. 2009. *The Benefits to Taxpayers from Increases in Students' Educational Attainment*. RAND Corporation.
- Charreire, H., R. Casey, P. Salze, C. Simon, B. Chaix, A. Banos, D. Badariotti, C. Weber, and J.-M. Oppert. 2010. Measuring the food environment using geographical information systems: a methodological review. *Public Health Nutrition* 13 (11):1773–1785.
- Chen, Y., J. Yu, K. Shahbaz, and E. Xevi. 2009. A GIS-based sensitivity analysis of multi-criteria weights. In *Proceedings of the 18th World IMACS/MODSIM Congress*, 13–17. Cairns, Australia: Citeseer.
- Citro, C. F., and G. Kalton eds. 2007. *Using the American Community Survey: Benefits and Challenges*. National Academies Press.
- Clifton, J. 2013. *The Coming Jobs War*. Simon and Schuster.
- Cowan, C. D., R. M. Hauser, R. A. Kominski, H. M. Levin, S. R. Lucas, S. L. Morgan, M. B. Spencer, and C. Chapman. 2012. *Improving the Measurement of Socioeconomic Status for the National Assessment of Educational Progress: A Theoretical Foundation*. National Center for Education Statistics.
- Cummins, S., and S. Macintyre. 2002. “Food deserts”—evidence and assumption in health policy making. *British Medical Journal* 325 (7361):436–438.
- Dache-Gerbino, A. 2016. College Desert and Oasis: A Critical Geographic Analysis of Local College Access. *Journal of Diversity in Higher Education*.
- Dache-Gerbino, A., J. M. Kiyama, and V. T. Sapp. (Revise and re-submit). The dangling carrot: Proprietary institutions and lack of college choice for Latina students. Review of Higher Education.

- Esri. 2017a. How Grouping Analysis works. <http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/how-grouping-analysis-works.htm>.
- . 2017b. How Similarity Search works. <http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/how-similarity-search-works.htm>.
- Frenette, M. 2006. Too Far to Go On? Distance to School and University Participation. *Education Economics* 14 (1):31–58.
- Fry, R. 2017. Americans are moving at historically low rates, in part because Millennials are staying put. *Pew Research Center*. <http://www.pewresearch.org/fact-tank/2017/02/13/americans-are-moving-at-historically-low-rates-in-part-because-millennials-are-staying-put/>.
- Gibbons, S., and A. Vignoles. 2012. Geography, choice and participation in higher education in England. *Regional Science and Urban Economics* 42 (1–2):98–113.
- Gould, P. R. 1969. *Spatial , Resource Paper No. 4*. Washington DC: Association of American Geographers.
- Hendrickson, D., C. Smith, and N. Eikenberry. 2006. Fruit and vegetable access in four low-income food deserts communities in Minnesota. *Agriculture and Human Values* 23 (3):371–383.
- Hillman, N. W., and T. Weichman. 2016. *Education Deserts: The Continued Significance of “Place” in the Twenty-First Century*. Washington DC: American Council on Education.
- Hossler, D., and K. Gallagher. 1987. Studying Student College Choice: A Three-Phase Model and the Implications for Policymakers. *College and University* 62 (3):207–221.
- Jones, S. G., A. J. Ashby, S. R. Momin, and A. Naidoo. 2010. Spatial Implications Associated with Using Euclidean Distance Measurements and Geographic Centroid Imputation in Health Care Research. *Health Services Research* 45 (1):316–327.
- Kipp, S. M., J. K. Wohlford, and D. V. Price. 2002. Unequal Opportunity: Disparities in College Access Among the 50 States. <http://folio.iupui.edu/handle/10244/272>.
- Lei, T. L., and R. L. Church. 2010. Mapping transit-based access: integrating GIS, routes and schedules. *International Journal of Geographical Information Science* 24 (2):283–304.
- Lynch, M., J. Engle, and J. L. Cruz. 2010. *Subprime Opportunity: The Unfulfilled Promise of For-Profit Colleges and Universities*. Washington DC: The Education Trust.
- Ma, J., M. Pender, and M. Welch. 2016. *Education Pays 2016*. CollegeBoard. <https://trends.collegeboard.org/sites/default/files/education-pays-2016-full-report.pdf>.

- Manikandan, S. 2011. Measures of central tendency: Median and mode. *Journal of Pharmacology & Pharmacotherapeutics* 2 (3):214–215.
- Matisziw, T. C., and T. H. Grubestic. 2010. Evaluating locational accessibility to the US air transportation system. *Transportation Research Part A: Policy and Practice* 44 (9):710–722.
- McKenzie, B., and M. Rapino. 2011. *Commuting in the United States: 2009*. Washington DC: US Census Bureau.
- Minnesota Population Center. 2016. *National Historical Geographic Information System*. Minneapolis: University of Minnesota.
- Oakes, J. M. 2017. *Methods in Social Epidemiology* 2nd ed. John Wiley & Sons.
- Office of the Press Secretary. 2015. Empowering Students to Choose the College that is Right for Them. <https://www.whitehouse.gov/the-press-office/2015/09/12/fact-sheet-empowering-students-choose-college-right-them> (last accessed 13 April 2016).
- O’Sullivan, D., A. Morrison, and J. Shearer. 2000. Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach. *International Journal of Geographical Information Science* 14 (1):85–104.
- Powell, L., S. Slater, D. Mirtcheva, Y. Bao, and F. Chaloupka. 2007. Food Store Availability and Neighborhood Characteristics in the United States. *Preventive Medicine* 44 (3):189–95.
- Rosero-Bixby, L. 2004. Spatial access to health care in Costa Rica and its equity: a GIS-based study. *Social Science & Medicine* 58 (7):1271–1284.
- Scott, L. M. 2000. Evaluating Intra-metropolitan Accessibility in the Information Age: Operational Issues, Objectives, and Implementation. In *Information, Place, and Cyberspace*, Advances in Spatial Science., eds. D. G. Janelle and D. C. Hodge, 21–45. Springer Berlin Heidelberg.
- Singleton, D. A. D. 2012. *Educational Opportunity: The Geography of Access to Higher Education*. Ashgate Publishing, Ltd.
- Tischler, M. A. 2016. Accelerating Geospatial Modeling in ArcGIS with Graphical Processor Units. *International Journal of Applied Geospatial Research* 7 (4):41–52.
- Turley, R. N. L. 2009. College Proximity: Mapping Access to Opportunity. *Sociology of Education* 82 (2):126–146.
- US Census Bureau. 2016. Americans Moving at Historically Low Rates, Census Bureau Reports. *The United States Census Bureau*. <https://www.census.gov/newsroom/press-releases/2016/cb16-189.html>.

Walker, R. E., C. R. Keane, and J. G. Burke. 2010. Disparities and access to healthy food in the United States: A review of food deserts literature. *Health & Place* 16 (5):876–884.

Weinberg, Z. 1995. *No Place to Shop: The Lack of Supermarkets in Low-Income Neighborhoods*. Washington DC: Public Voice For Food & Health Policy.

White, C. 2013. Create a network dataset from osm data. <https://github.com/Esri/arcgis-osm-editor>.

Worboys, M. F., and M. Duckham. 2004. *GIS: A Computing Perspective* 2nd ed. Boca Raton, Fla: CRC Press.