SOCIAL MEDIA TO LOCATE URBAN DISPLACEMENT: ASSESSING THE RISK OF DISPLACEMENT USING VOLUNTEERED GEOGRAPHIC INFORMATION IN THE CITY OF LOS ANGELES

by

Bryan Schaefer

A Thesis Presented to the
FACULTY OF THE USC GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
MASTER OF SCIENCE
(GEOGRAPHIC INFORMATION SCIENCE AND TECHNOLOGY)

December 2014

DEDICATION

I dedicate the document to my family for the constant support, and to my instructors, teachers, friends, and the GIS community for their encouragement and inspiration during the project. Finally, to my wife who simply believed in the completion of this project which in turn reflected her belief in me.

ACKNOWLEDGMENTS

Thank you to Dr. Matthew Zook and the Floating Sheep team who supplied the Twitter data for this project. This project would not have been completed without the guidance of Dr. Darren Ruddell, Dr. Karen Kemp, and Dr. Daniel Warshawsky.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACS	American Community Survey
API	Application Programming Interface
BID	Business Improvement District
BPS	Bureau of Planning and Sustainability
CHAS	Comprehensive Housing Affordability Strategy
CSV	Comma-Separated Value
DOLLY	Digital OnLine Life and You
FRC	Federal Renewal Community
GIS	Geographic Information Systems
GPS	Global Positioning Systems
HUD	U.S. Department of Housing and Urban Development
HAMFI	HUD Average Median Family Income
MFI	Median Family Income
SEZ	State Enterprise Zone
TNI	Targeted Neighborhood Initiative
VGI	Volunteered Geographic Information

ABSTRACT

This project investigates gentrification-related displacement in the City of Los Angeles, California by introducing an analytic method that utilizes Volunteered Geographic Information (VGI). Data harvested from the social media network Twitter were analyzed and the results compared against an established method to assess risk of displacement that utilizes aggregated census data. Aggregated census data are problematic in displacement research due to spatial and temporal constraints. The purpose of this investigation is to advance research on displacement by introducing an alternative method to gain a better understanding of the dynamic nature of gentrification and displacement by leveraging spatially explicit real-time VGI data.

This study examined approximately one million randomly harvested geotagged Twitter posts (tweets) within the City of Los Angeles, from August 2013 to January 2014, to investigate patterns of displacement. The research employed two frameworks: 1) a traditional census-based Data Aggregation Method; and 2) an alternative VGI (Twitter) based method. The results indicate that although tweets consisting of words related to displacement were not densely located in census tracts that have a high risk of displacement, as recorded by the Data Aggregated Method, areas of Los Angeles that are going through or just finished revitalization projects did contain such tweets. If left unmonitored, these areas could soon gentrify and displace as indicated by their demographic change over the last twelve years. In other words, the VGI Method detected a signal for potential displacement. Further, the VGI Method shows that data from Twitter produced results that are comparable to an established method of locating demographic change and go beyond an aggregated method's spatial and temporal level of analysis.

CHAPTER ONE: INTRODUCTION

Once home to gritty industrial zones and overcrowded urban slums, many forgotten neighborhoods within the City of Los Angeles, California now contain desirable pieces of real estate (Figure 1). In neighborhoods such as Hollywood and Downtown Los Angeles, the opening of chic cafés and converted loft apartments are just some of the indicators of the recent social and economic shift. A mix of local and federal policies prompted the market change by allowing police to have a more aggressive presence within communities known for having high crime rates and encouraged private investments into assigned poverty zones. While some Angelenos view this change as positive, it comes with the potential of displacing current, and potentially, long-standing residents.

Figure 1: The County, City, and Downtown boundaries of Los Angeles, California.

Displacement is the voluntary or involuntary removal of a household from a neighborhood due to conditions that are beyond the household's control (United States. Department of Housing and Urban Development. Office of the Secretary, Eunice Grier, and George. 1978). This can occur as development sets off a spike in rent and a change in the economic makeup of a community. Existing residents are forced to relocate because they can no longer afford to live in a neighborhood under this type of transition. A community that desires improvement without concern for the obstacles that occur as property values rise and one demographic succeeds another, runs the risk of cannibalizing the long-term benefits of revitalization with the short-term consequences of gentrification.

Displacement due to new social environments and an increase in the local economy, like in Los Angeles, is one of the most studied aspects of metropolitan neighborhoods undergoing a transformation, or gentrification. Yet displacement due to gentrification throughout the United States is often difficult to analyze because of the data used in common displacement studies. Often displacement studies use methods that rely on aggregated census data. As a result, the methods introduce a spatial and temporal constraint. Methods using census-aggregated data can only provide general geographic locations of areas at risk of displacement while the data becomes less significant to that location over time.

Spatial limitations begin with how the U.S. Census Bureau publishes census data. To protect a citizen's identity the U.S. Census Bureau publishes its enumeration using aggregated data. Census aggregated data sets summarize individuals, or other census features, by count within a particular area. When Geographic Information Systems (GIS) incorporate census-aggregated data into a method to determine risk of urban displacement, it results in a loss of

spatial detail. The finest scale shared by the features of a census-aggregated method determines the geographic scale of that study.

Adding to the spatial constraint, the U.S. Census is a decennial census, publishing data once every ten years. With the last U.S. Census conducted in 2010, displacement studies using census-aggregated data will not have an updated data set until 2020. Projects can use the American Community Survey (ACS), a sub-sample of the U.S. Census, as a supplemental data set to reduce the temporal difference of data collection to current displacement studies. However, the U.S. Census Bureau publishes ACS information one to three years after issuing a survey. Prior to this study, there is no existing method to locate real-time displacement at many spatial scales including an individual's specific longitude and latitude location.

To remove the spatial and temporal constraints found in census-aggregated methods this study examined the use of real-time Volunteered Geographic Information (VGI) to investigate neighborhoods at risk of displacement in the City of Los Angeles. This study used a traditional census-aggregation method and an alternative VGI based method to explore the advantages of real-time VGI data. By using VGI, studies researching displacement would benefit from a data set created with short temporal gaps between data collection and conclusion, while allowing analyses at a range of spatial scales.

Michael Goodchild coined the term VGI in 2007 to describe using the internet to create and assemble geographic information provided by individuals voluntarily (Goodchild 2007). He believes that the data sets created from VGI provide a timely, inexpensive, geographic view of life at many levels. While most of the volunteers have little geographic information qualifications, broadband communications, Global Positioning Systems (GPS), and geotag

technologies enable and motivate people to continue to contribute information. In most cases, a researcher can collect behavioral trends and geographic locations within the normal activities of a person's day-to-day life.

The purpose of this study was to design an alternative method to track the risk of displacement using Goodchild's description of VGI. The study assembled a data set for the VGI Method from posts containing geotagged information on the social media network Twitter. The project compared results of the VGI Method to baseline maps produced using a census-aggregated method to highlight their strengths and weaknesses. This project chose Dr. Lisa Bates' census-aggregated method (herein referred to as the Data Aggregated Method) currently used by the City of Portland, Oregon to create baseline maps.

This project made use of the Data Aggregated Method and framework for two reasons: 1) Bates designed a repeatable method that this study could use in the City of Los Angeles and 2) it provided an established framework within a year of this study with results accepted by a government agency. This study chose to recreate the Data Aggregated Method in the City of Los Angeles due to the large amounts of VGI data posted on Twitter by the Los Angeles population at the time of this project.

Using VGI data and the framework outlined in this study is an altogether new and unique method in displacement research. Anticipated results utilizing VGI data would be to spatially and temporally locate areas at risk of displacement in Los Angeles neighborhoods.

1.1 Motivation and Purpose

The motivation for this work was to design an empirical method that collected and shared information on displaced population within a shorter timeline than methods that use aggregated data. Evaluating urban displacement using census-aggregated data weakens the impact of displacement studies. In many instances, city planners and government agencies are analyzing future investments using data that does not currently connect to the population. Future displacement studies must introduce an alternative methodology and framework to remove the spatial and temporal disconnect. The VGI Method designed in this study is able to overcome these limitations by analyzing real-time population locations linked to displacement using an online social network.

Friedrich Engels (1884) was one of the first social scientists to publish works on urban displacement when he discussed factory workers and their living conditions during the Industrial Revolution. Engels observed the displacement of the working class to the slums of factory towns while factory owners profited from new manufacturing methods (Clark 2005). In the 1960s, British sociologist Ruth Glass coined the term gentrification and linked it to displacement. She used the term to describe the process of the working class districts in London's West End taken over by the middle to upper class (Hammett 2003). The districts began to gentrify as the middle class redeveloped old cottages and demanded greater amenities with the businesses to supply them. However, once redeveloped, the working class could not afford the higher rent on expired housing leases. This displaced the working class from London's West End and led to a complete demographic change within the districts.

The works of Engels and Glass have led to today's premise that gentrification-related displacement occurs in urban areas where the infrastructure creates a scenario for profitable

development of an existing neighborhood (Slater 2013). Private and public investments are able to take advantage of low property values to establish or expand their brand because of such variables as unemployment and a reduction in the local manufacturing industry. This allows for middle class professionals and service industries to relocate within urban areas they would not have considered before (Hammett 2003).

The relocation of the working class in a city often results in challenges for local governments. On one hand, public and private investments for revitalizing a neighborhood can be positive. Papachristos (2011) discuss the example of lower crime rates in Chicago linked to new upscale coffee shops in working-class neighborhoods. Here the economy created an opportunity for the working class to live in safer neighborhoods with greater retail investments raising the social value of Chicago's south and west side communities. Existing residents were willing stay and pay for improved communities because private and public investors developed with concern for the working class (Conn 2003). City government setting aside space for subsidized housing and creating policies to preserve a neighborhood while encouraging economic growth allows for the coexistence of all classes to take advantage of the common desire for quality schools, cleaner streets, and improved communal capital.

When responsible development does not take place it can result in negative consequences. As an example, as argued in a study by Reese of Los Angeles's skid row, displacement caused homelessness (Reese et al. 2010). In this case, Los Angeles in the early 1990s had policies in place that created dispersal zones around Downtown Los Angeles. Policies encouraged private investors to tear down public housing and develop projects to attract higher income classes. However, there was no program in place to accommodate existing residents. This strategy created gentrification and displacement with some of the population forced to move

to Los Angeles's skid row. When a city does not have the right tools to evaluate gentrification, long-time working class residents will not be able to handle the drastic rise of property value and the economy that comes with it. This type of gentrification leads to working class displacement, homelessness, or the extinction of whole communities.

Communities are in a constant state of change. Whether change to an urban area caused by redevelopment projects is for the better or worse is up for debate within different levels of governments, advocacy groups, and academia. This study did not analyze the social and economic effects of gentrification and displacement. Instead, the study's aim was to identify the flaws of using a census-aggregated method at a decadal scale to monitor a community and introduce a framework to overcome it. Analysis of demographic changes and housing market conditions has dominated previous gentrification-related displacement research (Aggeler 2012; Clerval 2011; De Verteuil 2011; Reese, Deverteuil and Thach 2010). Methods used in this type of research contain a mix of predetermined temporal and spatial units to examine the present. The process involves overlaying variables in hopes that the aggregation of the data does not apply a bias. Yet, by definition a displace population is no longer in their initial location. To remove these constraints when observing the risk of displacement, this paper implements an alternative VGI Method to observe displacement at a finer spatial and temporal resolution.

Including Twitter in the framework generates a large amount of spatial data used to analyze human behavior. The data set created by Twitter is new to displacement research, but attempting to use this data is not completely new to the spatial studies. Researchers have used Twitter data to study individual responses to natural disasters and global events such as wildfires and earthquakes (Fujisaka, Ryong and Kazutoshi 2010). Volunteered georeferenced data in

these cases observed a person's geographic pattern during these disasters to help coordinate future response and relief resources.

To analyze and map the risk of displacement, introducing VGI data from Twitter will add a new element when discussing its effect on urban communities. This will shift the timeline and scale of studying displacement. Studies will no longer have to rely on the use of decadal U.S. Census or ACS data and its set boundaries. Research for gentrification-related displacement will be able to collect and examine data below a street level scale with the ability to update the data daily. Monitoring current neighborhoods using VGI will help government departments come to a better understanding of the impacts, costs and benefits, and relocation programs associated with displacement.

1.2 Thesis Organization

This study is organized into six chapters. Chapter Two reviews the background information for this study. This includes: 1) the definition of displacement; 2) introducing Bates' aggregated displacement study for the City of Portland, Oregon; and 3) key elements that influenced the use of Los Angeles, California as the study area. Chapter Three analyzes the Data Aggregated Method for locating census tracts at risk of displacement, and then introduces the alternative VGI Method that leverages geolocated tweets to investigate displacement. Chapter Four presents the results of Data Aggregated Method and VGI Method. Chapter Five includes the analysis of geotagged tweets as compared to Economic Development Zones in the City of Los Angeles and demographic change over a twelve-year period to help understand high-density tweet locations. Chapter Six discusses the use of the VGI Method and data from Twitter as a spatial tool to locate and track the risk of displacement. The chapter finishes looking ahead to future VGI displacement studies.

CHAPTER TWO: BACKGROUND

Glass coined “gentrification” and framed this problem as an urban social issue (Slater 2013). Her work focused on the social character change happening within working-class neighborhoods due to higher-class takeovers. Glass regarded gentrification as the new urban gentry displacing the working class through an increase in housing values and cost of living. Her observations became a trailhead for displacement studies.

Since the 1960s, Glass' definition of gentrification and displacement has expanded. To understand the evolution of displacement studies the beginning section of this chapter separates the definitions of gentrification and displacement. Having two distinct definitions was a key element to outlining the framework for a VGI Method for locating areas at risk of displacement. Later, the chapter reviews the Data Aggregated Method, as applied to the City of Portland, Oregon. The chapter finishes by discussing the City of Los Angeles as the project’s study area.

2.1 Displacement and Gentrification

This study employs a method for locating areas at risk of displacement, but when trying to define displacement it can be easy to use gentrification as a synonym for the process. These two terms have a close relationship and it is important to define and discuss both for this study so that it creates some division between the two. To start, it is common to think of gentrification as linked to changes in a city’s economy (Marcuse 1985). In this setting, increases in property value and demand create a pattern of existing working class residents under pressure of displacement while the higher classes migrate to the new areas of revitalization. However, this general definition of gentrification does not address: 1) what motivates the migration into gentrified areas, 2) if gentrification always suggests displacement, and 3) if not, what are the alternatives to displacement when an investment’s intended is revitalization.

To attempt to answer these questions, displacement studies have had to expand the definition of gentrification from Glass' original observations. Glass viewed gentrification as the new gentry displacing the working class. Yet, when defining the term today there is little need to refer to the original "gentry" and its definition of the non-aristocratic English land-owning class (Redfern 2003). Instead, definitions of gentrification and displacement have splintered from her original observation. The new "gentry" no longer consists of the white middle class migrating into cast-off urban areas (Vandergrift 2006; Keating 2000). At present, gentrification includes any race that can simply afford the increase in property value and cost of living, unaware of the pressures they place on existing residents and the long-standing local businesses.

Academic journals such as the *International Journal of Urban and Regional Research* reflect the new composition of gentrification:

> Simultaneous with the ongoing tendency towards outer-suburbanization, we can observe an influx of middle-class populations into certain inner cities. An upgraded residential environment coupled with an influx of population keen to live in the city centre gives rise to the phenomenon known as gentrification. (Pattaroni, Kaufmann and Thomas 2012, pg 1224)

Within this definition, Pattaroni describes gentrification involving a process of the middle-class entering into different urban communities without the aspect of race. In 2013, Smith and Williams take the definition of gentrification even further away from race and see it as a spatial issue being too complex for a strict definition. Instead, they note we should recognize the different processes and components that lead to individual communities at risk of gentrification:

> Gentrification is a visible spatial component of this social transformation. A highly dynamic process, it is not amenable to overly restrictive definitions; rather than risk constraining our understanding of this developing process by imposing definitional order, we should strive to consider the broad range of processes that contribute to this restructuring, and to understand the links between seemingly separate processes. (Smith and Williams 2013, pg 3)

These are just two definitions out of many that attempt to describe the continuing depth and complexity of gentrification. Yet, there is one universal component to this social transition; it is a process with a spatial outcome. The most discussed spatial consequences of gentrification, but rarely as a standalone topic, is displacement. Over the last 50 years gentrification literature has explored displacement enough where the two have almost become substitutes for each other. This had led to far less independent displacement publications when compared to gentrification.

Gentrification does not have to be a despicable process. There are examples in New York and Boston where residents are more likely to stay and attempt to afford gentrifying neighborhoods due to the increase in quality of life and economic opportunities it can provide (Conn 2003). The neighborhoods where this side of gentrification exist have polices in place to guide revitalization projects (Montagne 2014). These polices can include the addition of affordable housing and controlling the timeline of public and private economic projects. However, when these policies are not in place or misguided by city agencies, gentrification can cause displacement.

Although there is minimal literature on just the subject of displacement when compared to gentrification, there is a base understanding of the concept outlined by George and Eunice Grier:

> Displacement occurs when any household is forced to move from its residence by conditions which affect the dwelling or its immediate surroundings, and which: 1. are beyond the household's reasonable ability to control or prevent; 2. occur despite the household's having met all previously imposed conditions of occupancy; and 3. make continued occupancy by that household impossible, hazardous or unaffordable. (United States. Department of Housing and Urban Development. Office of the Secretary, Eunice Grier, and George. 1978, 8)

The difficulty in addressing displacement as a standalone subject is that researchers struggle to track and find relocated populations (Freeman 2005). This can occur as a study tries to differentiate between a displaced population and the average social transition of a neighborhood not related to gentrification. In these instances, gentrification literature commonly assumes displacement or addresses it as a substitute of each other. Despite this generalization, displacement is the preeminent outcome discussed within gentrification studies due to its symbolism as a class struggle (Vandergrift 2006).

Combating displacement by laws and responsible funding of city projects would assure affordable conditions remain an option for long-term residents. Future research on displacement should separate displacement from gentrification. Upon doing so, as the term gentrification becomes more ambiguous when discussing the process of social changes in an urban setting, literature can stop assuming the two are synonymous. This would allow displacement studies to focus on frameworks to resolve the issue of locating and tracking displaced populations without the hindrance of fitting it into gentrification definitions. One such example of a study undertaking the split between gentrification and displacement has come out of Portland, Oregon in 2013.

2.2 The Data Aggregated Method on Gentrification and Risk of Displacement

The Northeast communities of the City of Portland, Oregon have had displacement issues over the last twenty years; African-Americans have voluntarily and involuntarily been displaced due to demographic and economic change, public and private investments, and changes in the housing market (City of Portland, Oregon 2012). This resulted in splitting up low-income communities and the displacement of longtime residents and the small businesses they supported (Bates 2013). More recently, the City of Portland has adopted policies to improve inner North and Northeast neighborhoods (Theen 2014; City of Portland, Oregon 2012). These policies and private investments are starting to attract higher-income populations. Existing residents are beginning to see the early signs of displacement as some renters and homeowners are no longer able to afford the rising property values and cost of living.

To address the risk of displacement to longtime residents in North and Northeast Portland the city responded with a 25-year plan that focused on equity called the Portland Plan. The plan includes strategies developed by the city collaborating with local businesses, organizations, and residents to design guidelines for growth (City of Portland, Oregon 2014). This includes a three-part approach to help understand ways to reduce the risk of displacement. The first part includes guidelines for policies for affordable housing. The second is a business development tool kit. The third part includes methods for tracking populations and program evaluation.

Within the third part of the approach, the Portland Plan calls for methods that track social and economic changes throughout the city. These methods would evaluate the potential gentrification impacts of governmental policies. Portland's Bureau of Planning and Sustainability (BPS) contracted Dr. Bates to assess the vulnerability of displacement due to the

increased stress of gentrification (Planning and Sustainability 2014). Bates' research and method focused on residential displacement and included a review of national practices, tools, and programs that could help to reduce gentrification and displacement.

Bates' 2013 paper, *Gentrification and Displacement Study: implementing an equitable inclusive development strategy in the context of gentrification,* focuses on a framework to help Portland have a better understanding of displacement for policy strategies. To define gentrification and displacement, Bates defaults to the Portland Plan. The Portland Plan describes displacement related to gentrification as lower income households displaced due to the loss of housing taking place in under-valued neighborhoods, often with a demographic change to the residents and businesses (City of Portland, Oregon 2014). Bates expands this definition by adding that the only difference between revitalization and gentrification is the negative outcome of involuntary displacement.

The method outlined in the research used aggregated analysis as a way to set a value, or numeric weight, to census tracts by their changing demographics. The method relies upon U.S. Census and U.S. Department of Housing and Urban Development (HUD) for its demographic data. This resulted in locating census tracts that were vulnerable to displacement by evaluating four different demographic variables per tract. The four variables were: 1) percentage of household renters; 2) percentage of communities of color; 3) percentage of bachelor's degree; and 4) percentage of household incomes at or below 80% of the HUD Median Family Income (MFI). By using the census and HUD data sets to locate vulnerable areas of displacement, Bates understands that the product of her method had spatial and temporal constraints. Bates can only update both the variables and the product of the Data Aggregated Method when the U.S. Census and HUD release significant updates (Bates 2013).

Bates' report assigns a vulnerability score to each census tract that is at risk of displacement. The Portland Plan uses these scores to help outline best practices for minimizing the negative effects of gentrification. Yet she states in her report that it cannot answer questions that pertain to issues of health, connected neighborhoods, and Portland's equity goals (Bates 2013). Only policy makers and community input can answer these questions. Bates' intent for her study was to help the Portland Plan determine which neighborhood the city should assist first to combat issues of displacement.

2.3 City of Los Angeles

Unlike the City of Portland, the County and City of Los Angeles have no clear policies or methods in place to discuss how to limit the effects of displacement. Advocacy groups and not-for-profit organizations such as the California Community Foundation and Congress for the New Urbanism have provided outlines for private and public investments for revitalization projects (California Community Foundation 2014; Congress of the New Urbanism 2014). These outlines, however, are just suggestions with the hope that they influence Los Angeles policy makers, but they do not have any authority as a project goes through its design and build phases.

This project selected the City of Los Angeles for its study area for three reasons. First, it provided an opportunity to introduce a VGI Method to locate areas at risk of displacement into a major metropolitan area. Second, according to a 2012 report by Semiocast, the location of posts recorded monthly makes Los Angeles a top ten city in the world in terms of Twitter use (Semiocast 2012). Third, Los Angeles in recent years is reaching new levels of diversity as more census tracts are finding a higher mix of races. Seen in Figure 2, The Weldon Cooper Center has created a map that represents this diversity (Evans 2012).

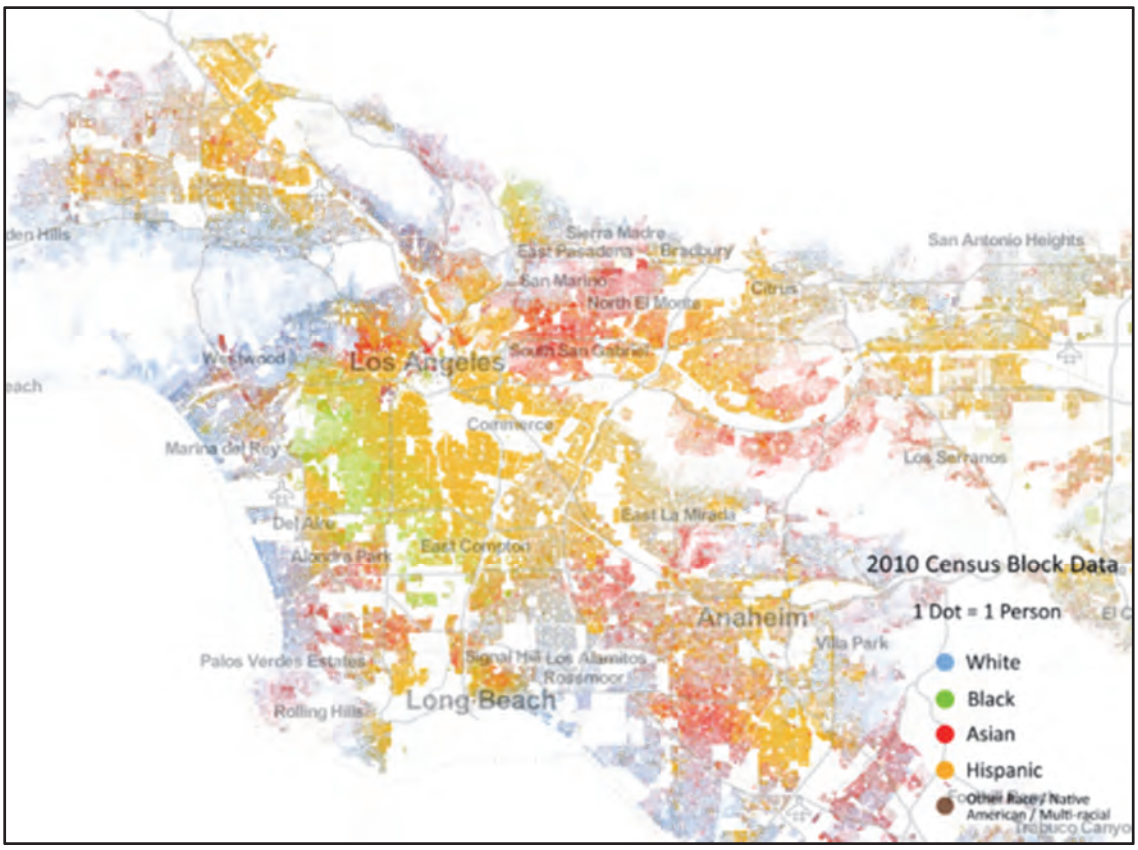


Figure 2: Racial dot map created by the Weldon Cooper Center showing the diversity and integration of races around Los Angeles County. Image Copyright, 2013, Weldon Cooper Center for Public Service, Rector and Visitors of the University of Virginia (Dustin A. Cable, creator).

Economy, geography, population, and general city statistics as seen in Table 1 show the diversity of Los Angeles. Los Angeles County is home to 9.963 million people; if it were its own state, it would be the ninth most populated state in the United States. Los Angeles has no majority population; instead, it hosts people from all over the world speaking over 180 different languages. There are 88 incorporated cities within the county ranging from cities like Vernon, with a population of just over a hundred people, to the City of Los Angeles, with a population estimated at 3.857 million. The County of Los Angeles is over 4,084 square miles and includes mountain ranges, a desert basin, and miles of coastline. Supported by tourism, commercial services, health services, sports and entertainment, and running the top port for trade and

transportation in the United States, the Los Angeles County economy ranks twenty-first in the world (Los Angeles County 2014).

Table 1: U.S. Census Bureau City of Los Angeles and City of Portland, Oregon demographic and geographic QuickFacts.

People QuickFacts	City of Los Angeles	City of Portland, Oregon
Population, 2012 estimate	3,857,799	603,106
Population, 2010 (April 1) estimates base	3,792,627	583,778
Population, percent change, April 1, 2010 to July 1, 2012	1.7%	3.3%
White alone, percent, 2010 (a)	49.8%	76.1%
Black or African American alone, percent, 2010 (a)	9.6%	6.3%
American Indian and Alaska Native alone, percent, 2010 (a)	0.7%	1.0%
Asian alone, percent, 2010 (a)	11.3%	7.1%
Native Hawaiian and Other Pacific Islander alone, percent, 2010 (a)	0.1%	0.5%
Two or More Races, percent, 2010	4.6%	4.7%
Hispanic or Latino, percent, 2010 (b)	48.5%	9.4%
White alone, not Hispanic or Latino, percent, 2010	28.7%	72.2%
Bachelor's degree or higher, percent of persons age 25+, 2008-2012	30.8%	43.1%
Homeownership rate, 2008-2012	38.0%	53.8%
Median household income, 2008-2012	$49,745	$51,238
Persons below poverty level, percent, 2008-2012	21.2%	17.2%
Geography QuickFacts	**City of Los Angeles**	**City of Portland**
Land area in square miles, 2010	468.67	133.43
Persons per square mile, 2010	8,092.3	4,375.2
Counties	Los Angeles County	Clackamas County Multnomah County Washington County
(a) Includes persons reporting only one race.		
(b) Hispanics may be of any race, so also are included in applicable race categories.		

This diversity allows Los Angeles to be influential on a global level in business and culture. Yet there are times when the diversity that has led to the city's success is strained. When populations from different backgrounds come across settings that create disharmony, tension can arise and cause communities to displace rather than share in the new cultural and economic wealth. Revitalization projects financed by public and private investments near the University of Southern California are examples of this growing disharmony. Investments in areas such as Jefferson Park have spurred economic revitalization while threatening to push current businesses and residents out of their community.

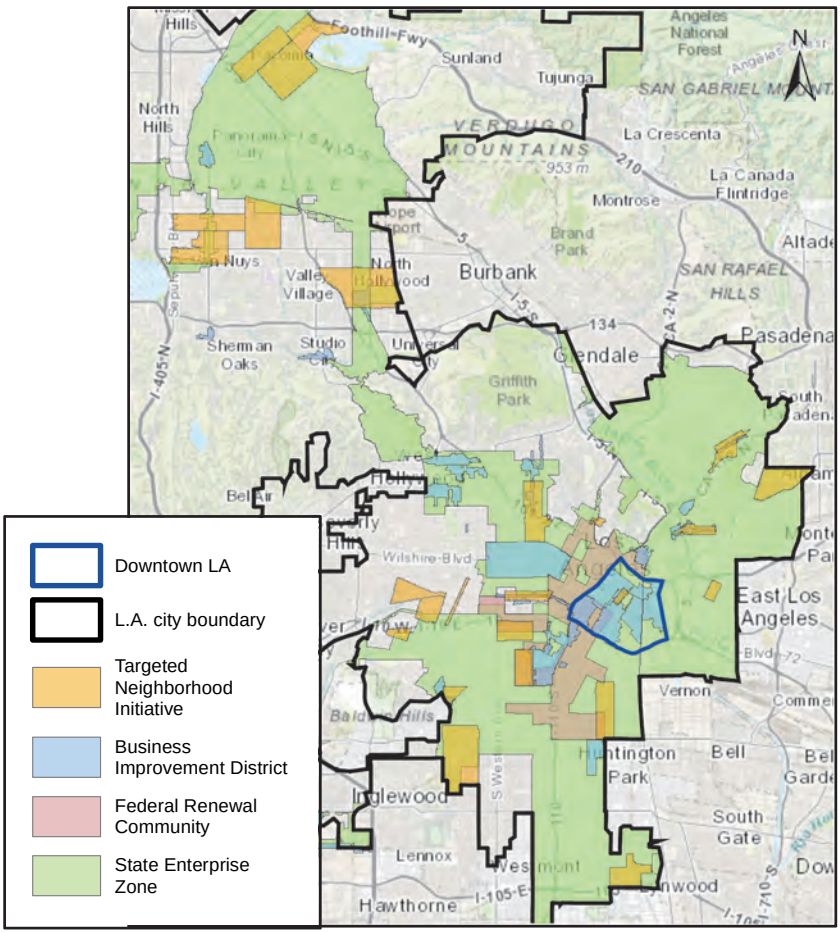


Figure 3: Economic Development Zones within the City of Los Angeles.

This is the challenge facing Los Angeles policy makers: balancing the revitalization projects, shown in Figure 3, which can create economic and social neighborhood value, while taking precautions not to displace existing residents. The VGI Method introduced in this study attempted to locate areas currently at risk of displacement in the overlapping economic development zones. By using the social media network Twitter, the results of this study were not limited by the spatial and temporal constraints found in aggregated methods like the one designed by Bates. Active data at a range of scales is an important tool in offsetting the potential of raising neighborhood economic value to displacement.

CHAPTER THREE: METHODS AND DATA SOURCES

This chapter details the frameworks used to compare how an established aggregated method locates areas at risk of displacement versus an alternative VGI Method. First, to complete the baseline results needed for comparison the chapter reviews the Data Aggregated Method recreated for the City of Los Angeles. Second, the chapter outlines Twitter as a VGI data source and the steps taken to execute the VGI Method.

3.1 Aggregated Data

Aggregating data is the process of grouping data at a spatial scale that is not as refined as the scale that it was collected at (Indiana University 2014). As an example, The U.S. Census collects data for every household in the City of Los Angeles, but releases the summarized information at a block level. Using aggregated census data always comes with a loss of spatial detail. In the case of using the finest level the U.S. Census provides, the census block might be as small as 250 feet by 250 feet, but is still a grouping of all data within that area.

For this study, aggregating data limited the Data Aggregated Method to the census tract level. This restraint happened for two reasons: 1) the U.S. Census releases data at scales to try to protect personal information and 2) one of the variables used in the Data Aggregated Method only scaled to the census tract level. Adding to the limit of scale, the U.S. Census is a decennial census, which only collects and releases data to the public every ten years. Even in cases where displacement research supplements ACS data for census data the information can still be years behind an active study. Breaking through the spatial and temporal constraints of using aggregated data for studying urban displacement is the focus of this project.

3.2 The Aggregated Method for Locating Census Tracts Vulnerable to Displacement

The Data Aggregated Method was used as a baseline for this study. Bates developed a method to analyze neighborhood change to help aid the Portland Plan with future policies. The method, she notes, is not absolute when predicting displacement, but does create a starting metric. Her displacement risk maps use only four data sets to reduce the amount of variables for a government to focus on and still create a clear typology. The Data Aggregated Method does not track population migrations from neighborhood to neighborhood; instead, Bates evaluates census tracts based on changes in income, educational attainment, race, and homeownership (Bates 2013).

Bates designed this aggregated method believing that demographic, socioeconomic, and housing market adjustments were indicators of neighborhood change caused by gentrification and displacement. Her framework is an extension of a method designed by Dr. Lance Freeman. Freeman's paper defined the process of gentrification as neighborhoods that have a higher than average increase in educational attainment and housing prices that start from below average incomes and housing values compared to the surrounding neighborhoods. These increases, combined with a change in racial demographics caused by white homebuyers, calculated by using Home Mortgage Disclosure Act data, have influenced several aggregated studies on gentrification and displacement (Bates 2013; Freeman 2005).

The method used by Bates to locate census tracts vulnerable to displacement reflects Freeman's definition. Bates' risk displacement model is grounded in four variables that she believes lead to displacement, which are: 1) A high percentage of housing units rented in a census tract compared to the citywide percentage. 2) A high number of communities of color out of the total population of a census tract, compared to the citywide percentage. A community of

color, as defined by Bates, is the total number of persons in a census tract that did not respond to the ACS as "No, not Spanish/Hispanic/Latino" and did not report "White" as their only entry when it came to race. 3) A high number of lower income households calculated against the percentage of family medium income of a census tract using the U.S. Department of Housing and Urban Development (HUD), compared to the citywide percentage. 4) A high number of persons without a college degree out of the total population 25 years and older in a census tract, compared to the citywide percentage. Bates argues that these variables contribute to the likelihood that residents will not withstand the increase in housing prices and revitalization that can turn into displacement due to gentrification (Bates 2013).

3.2.1 The Data Aggregated Method Framework

The Data Aggregated Method determines the risk of displacement for each census tract based on a score between zero and four. The method gives a numeric weight of one for each individual census tract variable that is over or equal to the citywide percentage of that variable. The citywide percentage of the four variables are: 1) 61% of households are renters, 2) 70% of the population are communities of color, 3) 71% of the population 25 years and older do not have a bachelor's degree, and 4) 34% of the households have incomes at or below 80% of the HUD adjusted MFI. The Data Aggregated Method adds the four variable scores together. Census tracts are at risk of displacement when they have a total numeric score of three or four.

3.2.2 The Aggregated Method Data and Preparation

The Data Aggregated Method used the 2006-2010 American Community Survey and the 2005-2009 HUD Comprehensive Housing Affordability Strategy (CHAS) files to calculate the tract and citywide percentages of the four variables. When recreating the method for the baseline maps of the City of Los Angeles this study updated all percentages and census tract

shapefiles using the 2008-2012 ACS and the 2006-2010 HUD/CHAS files. This study calculated the percentage of renters, bachelor's degrees, and communities of color unitizing files from the U.S. Census Bureau American Fact Finder website (http://factfinder2.census.gov/).

The American Fact Finder file provided the percentage of households that were renting per tract and citywide in Los Angeles. This project calculated the percentage of the population aged 25 years old and older that did not have a bachelor's degree by subtracting the percent of the population per tract and citywide that did have a degree from 100. The Data Aggregated Method calculated the percentage of the population classified as "communities of color" at the tract and citywide scale.

This project utilized HUD data from the U.S. Department of Housing and Urban Development website (http://www.huduser.org/). Household income was calculated at or below 80% of the HUD adjusted MFI using Table 8 of the Comma-Separated Value (CSV) CHAS data file. The method added all values from the rows that were less than 80%-100% and 100% and over the adjusted median family. That sum was divided by the total value per tract to find the variables' percent. The finest scale available given by HUD at the end of completing Bates' study for the Portland Plan and for this study was the census block level.

Using HUD/CHAS data as part of the data set highlighted some of the drawbacks of using aggregated methods. First, there was a large-scale limit to the method's outcome due to the HUD/CHAS data file. Table 8 of the CHAS file needed for this method only supplied data to the census tract level. This capped the scale of maps created to the census tract level regardless of the scale options of the other variables. Second, not all HUD/CHAS boundaries matched census tract boundaries because of their release dates. To make sure all variables were using the

same census tract limits, this study transferred HUD data to ACS 2012 tract boundaries. When HUD data were in more than one ACS census tract boundary, the rule classification scheme employed in this study utilized an equal distribution of the HUD percentage between the ACS census tract boundaries. This occurred thirty-one times when preparing the HUD data. When multiple HUD percentages were located in one ACS tract, the project averaged the HUD percentages within the ACS tract. Averaging the percentages occurred twelve times when preparing the data. Following Bates' study, in some cases the change to the census tract boundary was too insignificant to add this step before aggregating the data.

3.2.3 The Data Aggregated Method Applied to the City of Los Angeles

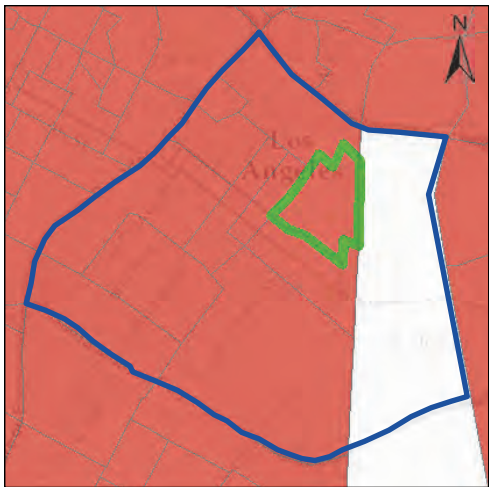
For the Data Aggregated Method, the project analyzed the risk of displacement for the City of Los Angeles utilizing a point system. Points were assigned using the citywide percentage of each displacement variable as a threshold and comparing each census tract percentage of that variable against it. If a variable's percentage was above, or equal to, the citywide percentage within a census tract the project labeled the tract as "high" and it received one point. If the percentage is below the citywide percentage the project labeled the variable "low" and the tract receives zero points. As an example, if the Downtown Los Angeles tract 2077.00 has a higher percentage of households that rent compared to the citywide percentage of Los Angeles, it gets one point. However, if the percentage of communities of color, for the same tract, did not exceed the variable's citywide percentage, the census tract would receive no points for that variable. The project assigned census tracts a total score between zero and four after the project added all four variable scores for each tract. A score of 0 to 2 labeled the census tract as having a low vulnerability of displacement while the project labeled a tract having at least three out of four as having a high risk of displacement. Table 2 lists all four variables, their citywide

percentages for the City of Los Angeles, and percentages for the City of Portland for comparison.

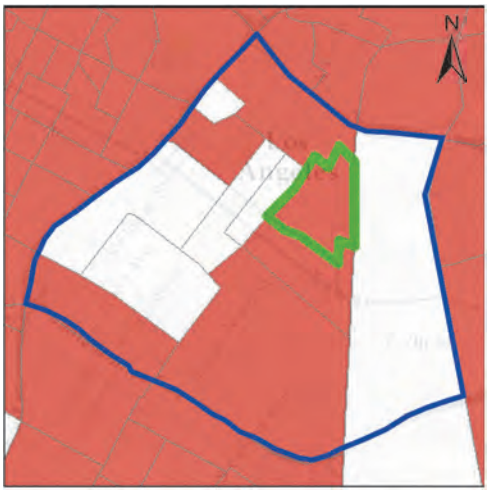
Table 2: Los Angeles and Portland, Oregon displacement indicators and percentage thresholds.

Displacement Indicators	City of Los Angeles percentage threshold	City of Portland percentage threshold
Housing units rented	61%	44%
Communities of color	70%	26%
25 years and older that do not have a bachelor's degree	71%	58%
Incomes at or below 80% of the HUD adjusted median family income	34%	47%

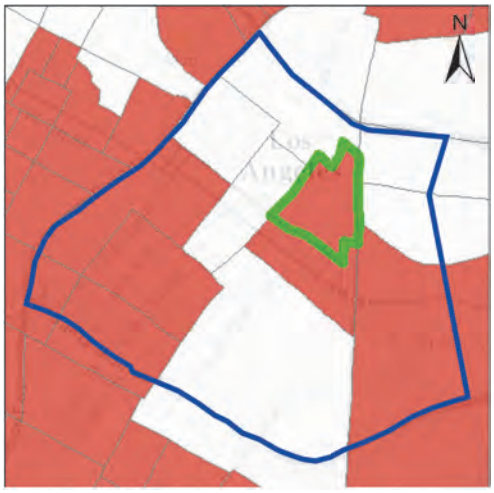
Figure 4 shows an example of how the Data Aggregated Method scored a single tract in Downtown Los Angeles. A green boundary highlights the tract. To start, 82% of the homeowners rented. This is greater than citywide percentage of 60%, so the variable receives one point. 86% of the populations are communities of color, which is greater than the citywide percentage of 70%, so the tract receives another point. 70% of households in this tract have incomes at or below 80% of the HUD median family income, which is higher than the citywide percentage of 34%; this variable adds another point to the total tract score. Finally, 77% of 25 years old and older in the census tract do not have a bachelor's degree. Again, this variable is higher than the citywide percentage of 71%. Adding these scores together, the census tract scored a four out of four and labeled as having a high risk of displacement.



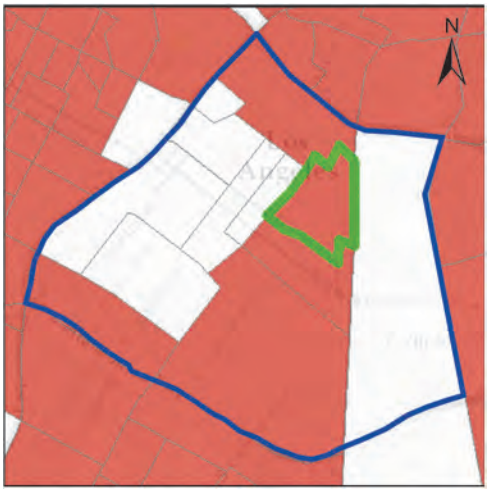
Household Renters
Over citywide % = 1 point



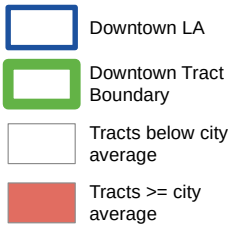
Communities of Color
Over citywide % = 1 point



Median Family Income
Over citywide % = 1 point



Without a Bachelor's Degree
Over citywide % = 1 point



The Data Aggregated Method analyzed the risk of displacement for the City of Los Angeles utilizing a point system. If a variable's percentage was above, or equal to, the citywide percentage within a census tract the project labeled the tract as "high" and it received one point. The example tract, highlighted in green, was "high" in all four variables and scored a four out of four. The tract was labeled as having a high risk of displacement.

Figure 4: An example of a census tract, highlighted in green, in Downtown Los Angeles scored for risk of displacement.

3.3 The VGI Method: An Alternative Displacement Framework

This study proposes an alternative method of extracting and mapping VGI data through social media. Unlike data released by the U.S. Census, which has clearly defined spatial and temporal parameters, Twitter data are real-time, geo-referenced data, which is ideal for displacement research. This study chose Twitter, a microblogging service, as the data set source for this method. Twitter provided a substantial amount of users, the ability to obtain data for this study on time, and the option for users to attach geolocations to blog posts.

3.3.1 VGI Framework

The VGI Method introduced in this study is grounded in five sequential steps, which are: 1) acquire a current geotagged data set from Twitter, 2) create a list of keywords related to displacement from current news articles and relevant websites, 3) isolate geotagged tweets from the data set that contain those keywords, 4) plot a kernel density map to analyze keyword tweet locations, and 5) compare those locations to the aggregated baseline maps. The expected outcome was that a high amount of displacement keyword tweets would be within aggregated census tracts at high risk of displacement displayed in the baseline maps.

3.3.2 Social Media Network Twitter

Twitter, Inc. launched in 2006. Twitter is an open platform for public expression and conversation through real-time microblogs. Twitter users can post from the Twitter website as well as mobile applications such as Vine and #Music for example (Twitter 2014). Twitter is accessible through its web portal or mobile applications on the iOS, Android, and Windows operating systems. Twitter also provides companies and individuals that apply for and get an OAuth consumer key access to its public Application Programming Interface (API). This gives developers the ability to embed Twitter into their own platforms. Its microblog posts, called

tweets, consist of only 140 characters, with as high as twenty percent of those tweets containing geolocation specific data linked to them (Weidemann 2013). Location data from tweets can create a spatial database for analysis at scales ranging from global to below the U.S. Census block level.

It is difficult to get a true sense of how many people are using Twitter or its popularity, but five recent articles about the platform can help. The Wall Street Journal in April of 2014 reported that there are 974 million existing Twitter accounts, yet 44% of those accounts have never sent a tweet (Koh 2014). This does not mean that those accounts are not active because they have not tweeted. Instead, they may be using Twitter to follow topics they find interesting. Jim Edwards for Business Insider in October of 2013 wrote an article stating 26% of teens view Twitter as the most important social network to join versus Facebook listed at 23% (Edwards 2013).

Tech Crunch reported in December of 2013, based on a Pew survey, that Twitter was the fourth most popular social network behind Pinterest (Lunden 2013). Yet, as reported by Mashable in April of 2014, Twitter users in the U.S. spend 86% of mobile social media time on Twitter, ranking it third among other social media networks. Lastly, in March of 2014, Pew released a study on the relationship between news and social media. Their report lists 16% of U.S. adults use Twitter, making it the third most used social network in the U.S. (Masta and Mitchell 2014).

Based on these reports and polls this study considers Twitter the third most used social media site in the U.S. in early 2014. This is always in a state of flux, as people's tastes change with the introduction of new ways to connect, growing social pressures to leave or join a

network, or companies updating their platforms in an attempt to gain new users. Yet, according to Twitter's IPO filing in September of 2013, the company handles 500 million tweets a day and looks to raise over one billion dollars in revenue in its first five years as a traded company (Kim 2013). Ranked third, with eight years of user growth and mobile activity, combined with an API that allows for longitude and latitude data harvesting, Twitter is an ideal social network to design a VGI Method for understanding displacement.

3.3.3 Floating Sheep: The DOLLY Project

There are two basic ways to collect Twitter data for spatial research. The first is to write your own API program that harvests profiles, geographic locations, and tweets. The second is to get Twitter data through a third party. This study acquired twitter data through Floating Sheep's DOLLY Project database, provided by Dr. Matthew Zook from the University of Kentucky. Using a third party to collect Twitter data allowed for more focus on the design of a VGI framework in the time given for this study. The Digital OnLine Life and You (DOLLY) Project is a data set archive of geolocated tweets created for real time research and analysis (Floating Sheep 2013). The Floating Sheep designed DOLLY for basic analysis of eight million geotagged tweets per day collected by the Floating Sheep team. The project has indexed and geocoded over three billion tweets since its launch in 2011. This has allowed for historic and real-time geolocation analysis.

Floating Sheep's system for harvesting tweets runs on a virtual private cloud operated by themselves and the computer service department in the College of Arts and Sciences at the University of Kentucky. The private cloud runs on 18 virtual servers with the option to add more as needed. To overcome the challenges they encountered when they started to create and index the large and ever growing database, Floating Sheep used a combination of Apache Cassandra, to

store the data, elasticsearch ELK Stack, to search the data, and Ruby scripts. This built a back end system designed for redundancy and availability that was then able to run if there was ever hardware failure. This is an important feature of the DOLLY project; tweets stream live and not saved indefinitely by Twitter Inc. If tweets are not analyzed and stored as users post them they could be lost immediately.

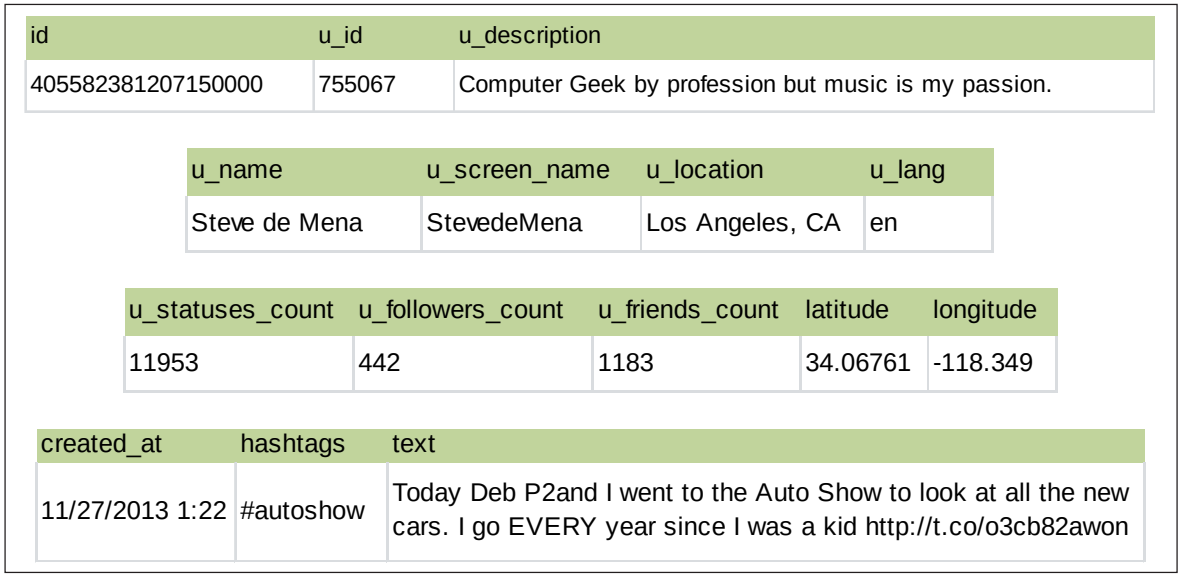
Supplied in January 2014, geotagged tweets from the DOLLY Project for this study included a CSV file containing a random ten percent sample of ten million tweets. The file consisted of tweets from August 2013 to January 2014 and centered on Downtown Los Angeles, California. After discussing the amount of tweets needed to design the VGI Method for this project, and then comparing it to the Data Aggregated Method, Floating Sheep believed the amount of one million tweets with geolocations would be a large enough data set to produce results within the given timeframe. This study agreed with the recommendation that the amount of tweets was satisfactory due to the DOLLY project's history of contributing to other spatial studies.

3.3.4 VGI Displacement Method Using Twitter

The alternative VGI Method for locating census tracts at risk of displacement through keyword geotagged tweets began by examining the CSV file from the Floating Sheep DOLLY project. Finding keywords in Twitter posts from the CSV file that relate to displacement and marking their locations was the basis of the VGI framework.

The CSV file contained 982,646 rows without duplicate Twitter posts. Each row contained a tweet's identification number, user identification number, user description, user name, user screen name, if the user gave a user location, user language, how many tweets they

have posted, how many followers they have, how many other users follow them, the tweet's latitude and longitude, the state, geo identification, when it was created, any hashtags, and finally the tweet itself. Figure 5 shows an example of one row from the CSV file.



id	u_id	u_description
405582381207150000	755067	Computer Geek by profession but music is my passion.

u_name	u_screen_name	u_location	u_lang
Steve de Mena	StevedeMena	Los Angeles, CA	en

u_statuses_count	u_followers_count	u_friends_count	latitude	longitude
11953	442	1183	34.06761	-118.349

created_at	hashtags	text
11/27/2013 1:22	#autoshow	Today Deb P2and I went to the Auto Show to look at all the new cars. I go EVERY year since I was a kid http://t.co/o3cb82awon

Figure 5: An example of one row of raw Twitter data from the Floating Sheep DOLLY Project CSV file.

To begin, the project plotted tweets using their latitude and longitude coordinates in ArcGIS 10.2. Figure 6 shows the map created by plotting all tweets. This map shows that the tweets provided by Floating Sheep did not cover the whole city; instead, they covered just over 882 square miles centered on the census boundary of Downtown Los Angeles. Although the tweets released by the Floating Sheep did not cover all of the City of Los Angeles, the amount and their locations were enough to design and compare the VGI Method to the aggregated baseline map. Figure 7 shows the project clipped tweets to the city limits of the City of Los Angeles using the U.S. census city limits shapefile. This created the same limit boundary as the Data Aggregated Method for analysis. Clipping tweets to the city limit reduced the number of tweets to 341,121.

Once the tweets were restricted to the city, the project selected keywords to locate posts related to displacement based on Los Angeles Times articles on gentrification and displacement (The Time editorial board 2014; Karlamangla 2014; Martelle 2014; Gerber 2014). Curbed Los Angeles's Gentrification Watch webpage (VOX Media 2014) and Urban Dictionary (Urban Dictionary 2014) also provided the project with words associated with the current language used by Twitter members. Reading Los Angeles Times, Curbed Los Angeles Gentrification Watch articles, and definitions from Urban Dictionary led to a 17-word list seen in Table 3. The project used the words to isolate tweets that had one or more of the keywords in the geotagged Twitter posts from the Floating Sheep CSV file. Although "revitalization", "displacement", and other variants of those words returned with no results, Table 3 includes them due to the project's subject of displacement.

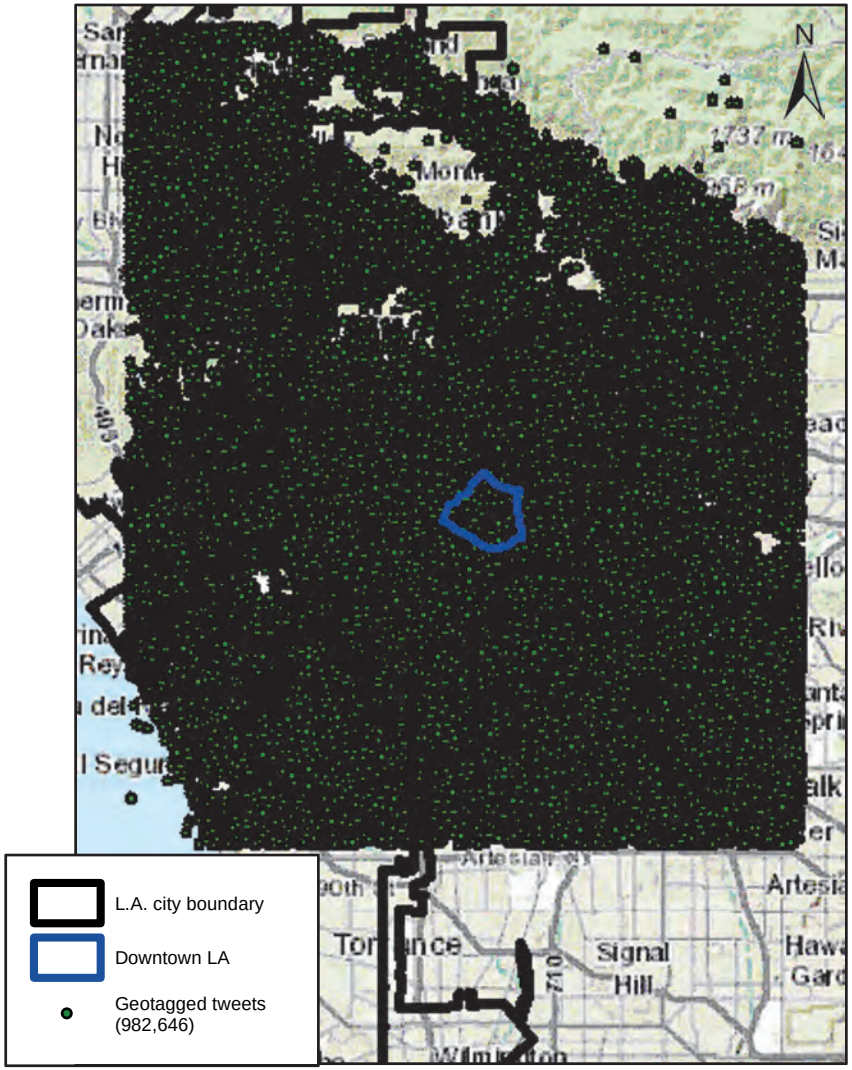


Figure 6: 10% of all geotagged tweets from August 2013 to January 2014 provide by the Floating Sheep's DOLLY Project.

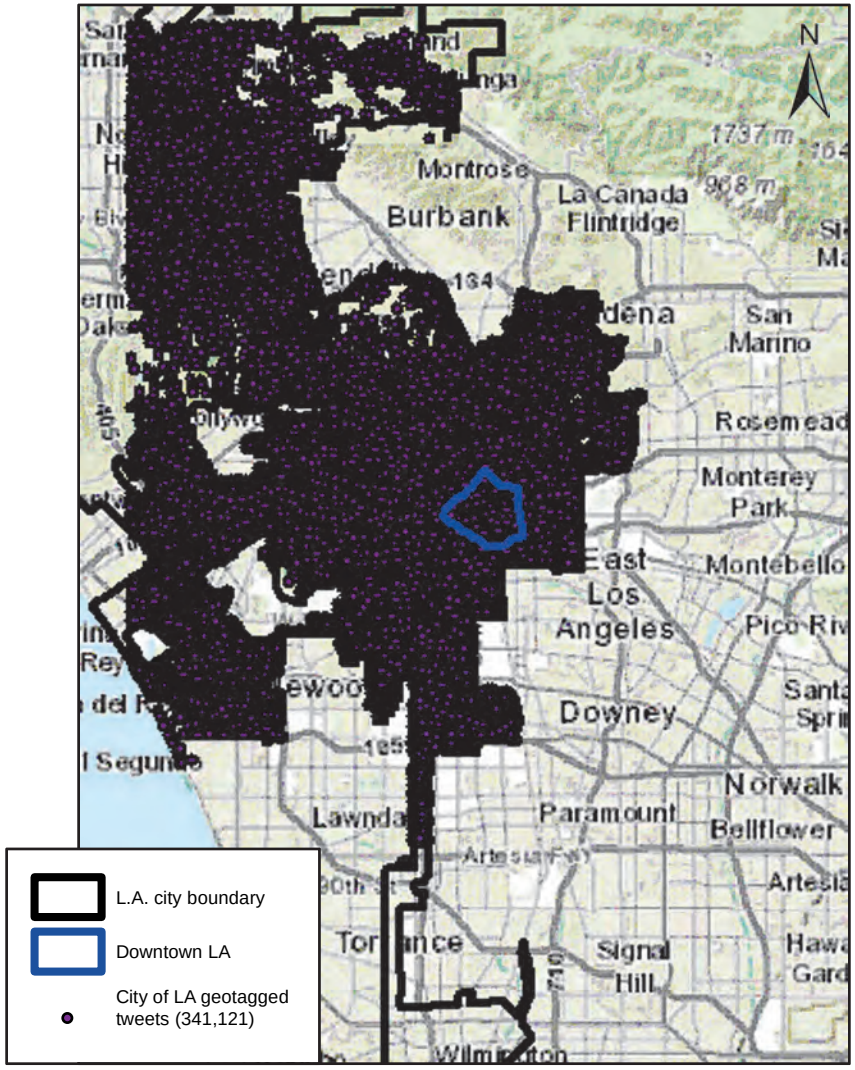


Figure 7: Geotagged tweets supplied by Floating Sheep's DOLLY Project clipped to the limits of the City of Los Angeles.

Table 3: The list of keywords and how many times they appear, including variants of that word, in the Floating Sheep CSV file of geotagged tweets in the City of Los Angeles.

Displacment keywords	number of tweets
Starbucks	1081
neighborhood	163
rent	161
hipster	159
district	98
retail	69
loft	58
urban	45
income	31
diversity	28
housing	15
renewal	9
gentri	7
eviction	3
yuppie	2
revitalization	0
displace	0

Tweets were isolated using a keyword search or variants of the word relevant to displacement from the rest of the tweets in the CSV file using an Excel wildcard search. Each word in each keyword tweet was given its own row on a separate sheet from the CSV file in Microsoft Excel. The project sorted all words using a single column PivotTable. A single PivotTable allowed for faster sorting and review of the 17 keyword variants, insuring the VGI Method included all relevant tweets. As an example, filtering the PivotTable using the keyword "loft" revealed the variant "loftjet". The project deemed "loftjet" to have no relevance to displacement. The procedure removed "loftjet" from the keyword list and by doing so removed the tweet containing it. Plotting keyword maps straight from a Microsoft Excel wildcard search of the word "loft" without filtering tweets in the CSV file would have included "loftjet" and affect the end results of the study . The project could have removed tweets that included variants

of “loft” when reviewing all tweets before plotting the final VGI maps. However, this study believed that filtering and removing tweets by using a single PivotTable at this stage of the method was more efficient.

Results were then reviewed to ensure that tweets that contained at least one of the words associated with displacement from the keyword list. This step reassured that no keyword tweets were duplicates or reposts. The project used these keyword tweets to create a kernel density map with directional distribution in ArcMap. To create the kernel density map the method used the Calculate Distance Band from the Neighbor Count tool in ArcMap to determine the average distance of tweet location to tweet location. The method used the distance as the search radius for the ArcMap Kernel Density tool. Figure 8 exhibits the geographic incidences and directional trend of the keyword tweets.

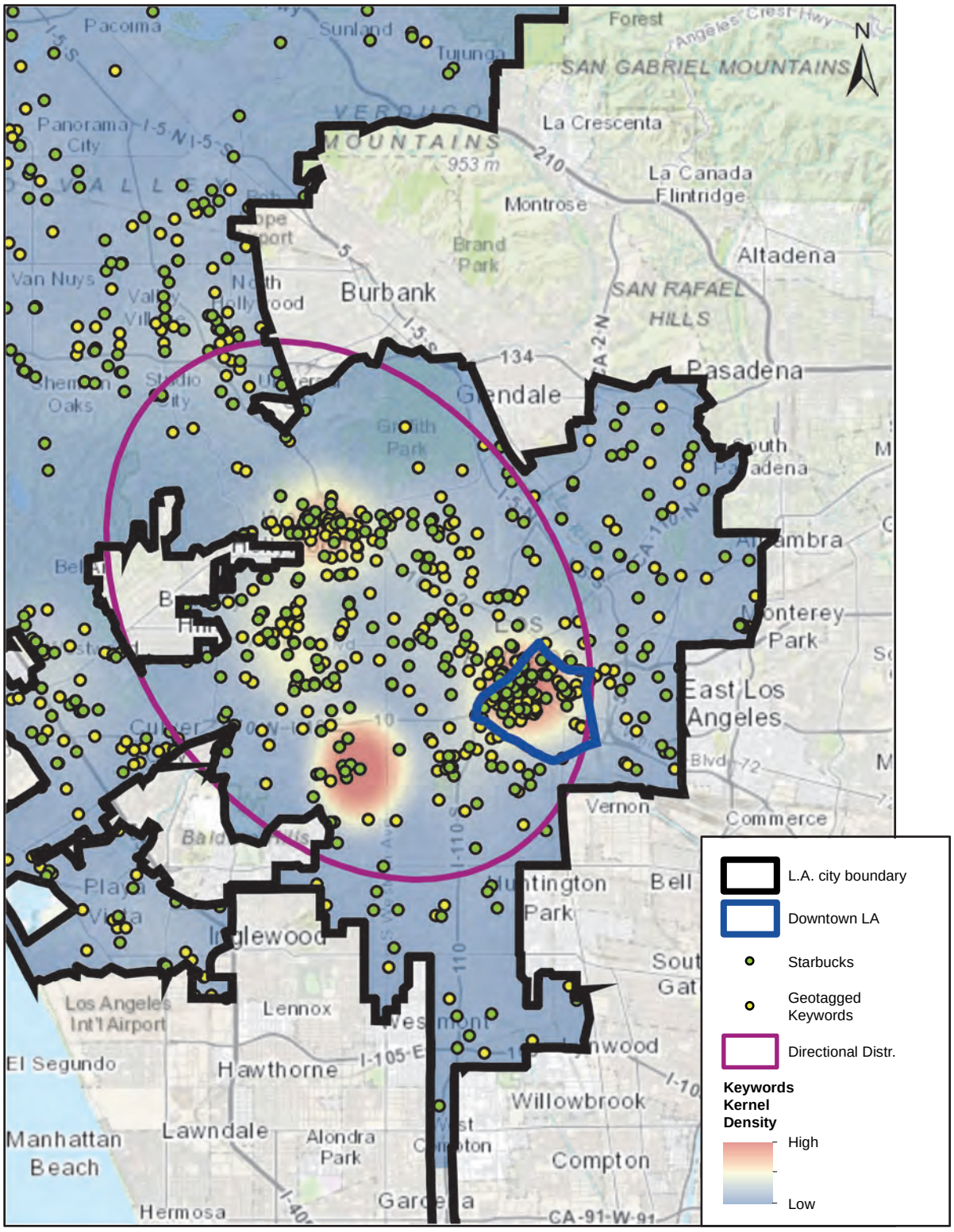


Figure 8: Kernel density map with directional distribution of keyword tweets.

3.3.5 Limitations Using Twitter

Twitter's VGI data advantage and benefit to spatial science research is its immediate access to geolocations when a project submits a query to Twitter's API. Twitter's weakness within this framework is its servers do not retain all posts. As of October 2013, Twitter indexes every tweet at the time of its post for only six to nine days depending on user activity. This means when a study needs VGI data before it executes its own query it will have to gather the data through other sources. Companies such as GNIP, Beevolve, and HootSuite can provide historic information at a range of prices, data, and timelines. An alternative is to request the data from not-for-profit projects like TweetTracker from Arizona State University or from Floating Sheep, as used in this study.

Relying on outside sources for Twitter data can lead to other limitations. Budgets for collecting data or the amount of data organizations release can limit a study's data sets when executing the VGI Method. As an example, this study only uses 10% of all tweets released in a period. The amount of data granted for this study is sufficient to create a VGI Method to compare against the Data Aggregated Method. Yet, the VGI maps produced for this study could vary if this project had access to all tweets from Floating Sheep within the same period.

In addition to Twitter not retaining its own posts and relying on third parties for historic data sets, using Twitter data comes with the risk that the site will be unavailable to harvest data at some point. Maintenance, software and hardware errors, website hackers, and other issues have the possibility of closing the site and access to Twitter's API. This would leave location gaps in data sets when trying to execute a VGI spatial study using Twitter. Even if the site is not completely offline, these issues have the potential to corrupt location data connected to tweets.

3.4 2000–2012 Demographic Change Related to Gentrification

The method designed for this study examined displacement through VGI supplied by Twitter users during a six-month period. If a high density of keyword-geotagged tweets is within a census tract at risk of displacement, reviewing the demographic history of that tract will give some understanding as to why it is at risk. Turning again to Bates' study for the Portland Plan, she notes the relationship of demographic change to gentrification. Analyzing increases in comparable variables to her vulnerability method can indicate that gentrification and displacement have occurred or are occurring in census tracts (Bates 2013). These variables include increases in homeownership, white populations, household income, and residents with college degrees. Bates believes that increases in these variables tell the story of wealthy, educated residents within a census tract raising property values and placing pressure on existing residents to displace.

To be able to compare the results of the VGI Twitter method with demographic change, this study divided percentages provided by the U.S. Census 2008-2012 ACS for homeownership, white population, the population of 25 years and older with a bachelor's degree, and the median household income by the same percentage variables in the 2000 Census for the City of Los Angeles. The percentage point difference from 2000 to 2012 established citywide thresholds for each variable. The same routine assigned a rate of change from 2000 to 2012 to each census tract within the city by using census tract percentages.

In some cases, the U.S. Census Bureau redrew a tract into two or more, or combined multiple tracts from 2000 to 2012. When the U.S. Census Bureau split tracts, the rule classification scheme in this study used equal distribution of the percentage from 2000 in the resulting tracts to calculate the rate of change. When the project combined tracts from 2000, it

used 2012 tract boundaries with the average percentage of the combined 2000 tracts to calculate the rate of change (Bates 2013). In some cases, the boundaries were determined as not having enough significant change to add this step when calculating the variable percentage point difference of a census tract. The study combined tracts eight times and split tracts twenty-six times after analyzing 1,119 tracts. To find the average household income in year 2000 dollars, the project adjusted for inflation to 2012 dollars before recoding percentages. The Bureau of Labor Statistics calculated that there was a 33% adjustment. This means $1.00 in 2000 had the same buying power as $1.33 in 2012.

To determine if a census tract had gone through enough significant gentrified demographic change over the twelve-year period, the project scored tracts similar to how Bates scored tracts vulnerable to displacement. If a census tract's rate of change for a demographic variable was above the citywide threshold for that variable then it received one point, if it was equal to or below the threshold, it received zero points. Again as an example, if the Downtown Los Angeles tract 2077.00 had a rate of change in homeownership that was above the City of Los Angeles's threshold from 2000 to 2012, it received one point. If the rate of change of the white population did not exceed the citywide threshold, the census tract would receive no points for that factor. Census tracts that had significant demographic change linked to gentrification had at least three out of four possible total points.

CHAPTER FOUR: RESULTS

This project examined the use of VGI data in an analytic method to locate displacement in the City of Los Angeles. Close to one million geotagged posts from the social media network Twitter were used in the VGI Method framework and the results compared against an established method that used aggregated census data. This chapter begins with a general description and format of the resulting maps used for analyses. The sections that follow present the tables and maps from analyses utilizing the VGI Method and the Data Aggregated Method frameworks.

4.1 Format of Resulting Tables and Maps

The aggregated maps used as a baseline for comparison in this study show the vulnerability scores for each census tract in the City of Los Angeles. A second set of aggregated baseline maps focused on Downtown Los Angeles. The VGI Method maps are at the extents of the Twitter data supplied by Floating Sheep and again around Downtown Los Angeles. The study shows census tracts within Downtown Los Angeles for all maps within a blue border. The following Data Aggregated Method maps show the four variables (households that rent, communities of color, lower income households, and lack of college degrees) mapped individually. The four maps show census tracts that are equal to or above (shown in red) and below (shown in white) citywide percentages for that variable. The final aggregated baseline map for comparing the VGI Method follows the four variable maps in the chapter. Unlike the binary outcomes of the variable maps where a tract either was or was not at or above the citywide percentage, the final aggregated map shows degrees of risk of displacement for each tract. A five color red gradient relates the five possible scores from the Data Aggregated Method. The study linked darker shades of red on the gradient with higher aggregated scores. This study considered a census tract to be at risk of displacement when a tract scored at least

three out of four possible points, meaning a tract had at least three out of four variables at or above that variable's citywide percentage.

4.2 Census-Based Variables to Locate Census Tracts Vulnerable to Displacement

4.2.1 Displacement Variable: Household Renter

The binary outcome of the Data Aggregated Method, Figure 9 and 10, show the vulnerability variable of the percentage of household renters within census tracts. The Data Aggregated Method gave census tracts greater than or equal to the citywide percentage of 60% one point. Tracts under the citywide percentage receive zero points. The census tracts that scored a point appear grouped in three locations: North Los Angeles, around Downtown Los Angeles, and South Los Angeles. Although other variables will have groupings in other locations and sizes, this example of finding variables clustered will carry throughout the variable maps for locating census tracts at risk of displacement.

Downtown Los Angeles for this variable showed that all tracts received a score of one, or over the citywide percentage, except for a single tract in east downtown. Like the City of Los Angeles, the groupings of census tracts continue throughout the other three variables within the downtown border. These groupings differ in location from map to map and cause the final baseline map scores of tracts downtown to appear scattered.

4.2.2 Displacement Variable: Communities of Color

Figures 11 and 12 map the displacement variable communities of color. Bates defines communities of color as a percentage of all residents within a census tract that are not non-Hispanic whites. Tracts that are greater than or equal to the citywide percentage of 70% scored one point. Groups of census tracts over the citywide percentage were in North Los Angeles and

around the downtown area to the southern end of the city. Tracts that are not over the citywide percentage for this variable were similar to the variable household renter, and located in West Los Angeles around Pacific Palisades, Santa Monica, and Beverly Hills. Downtown tracts that are under the citywide percentage include five tracts in West Downtown and again the single tract in the east.

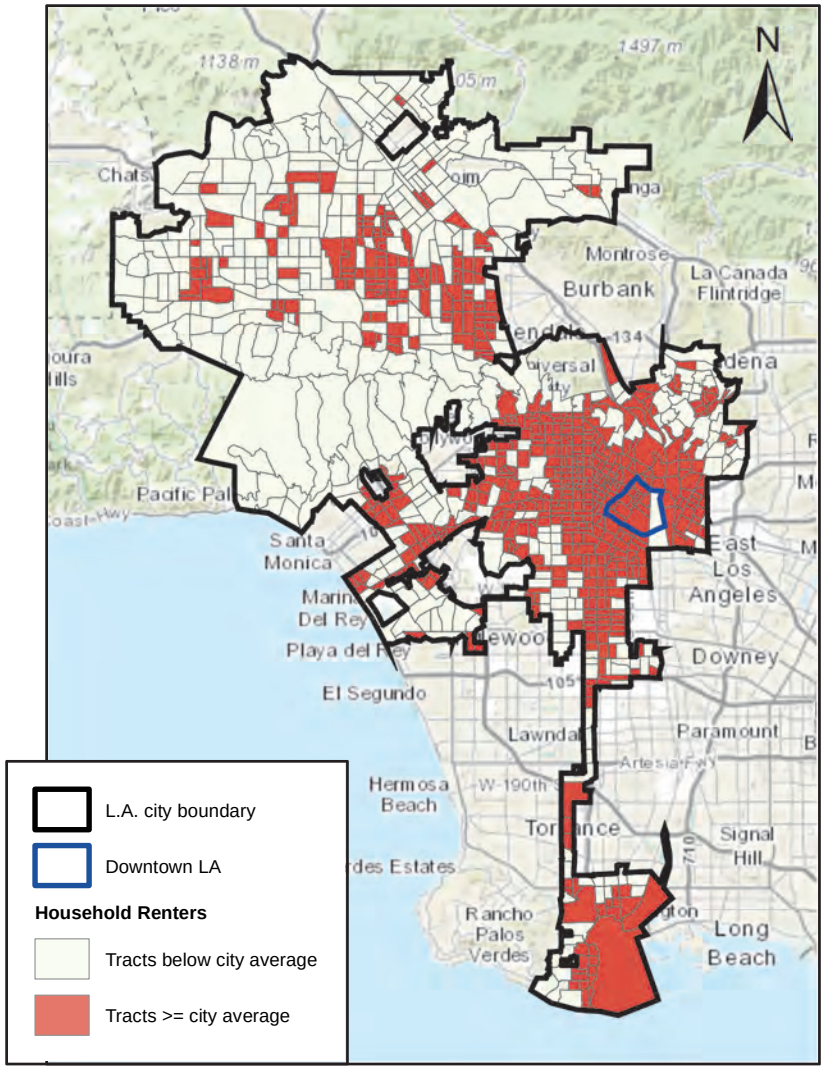


Figure 9: Household Renter >= 60%.

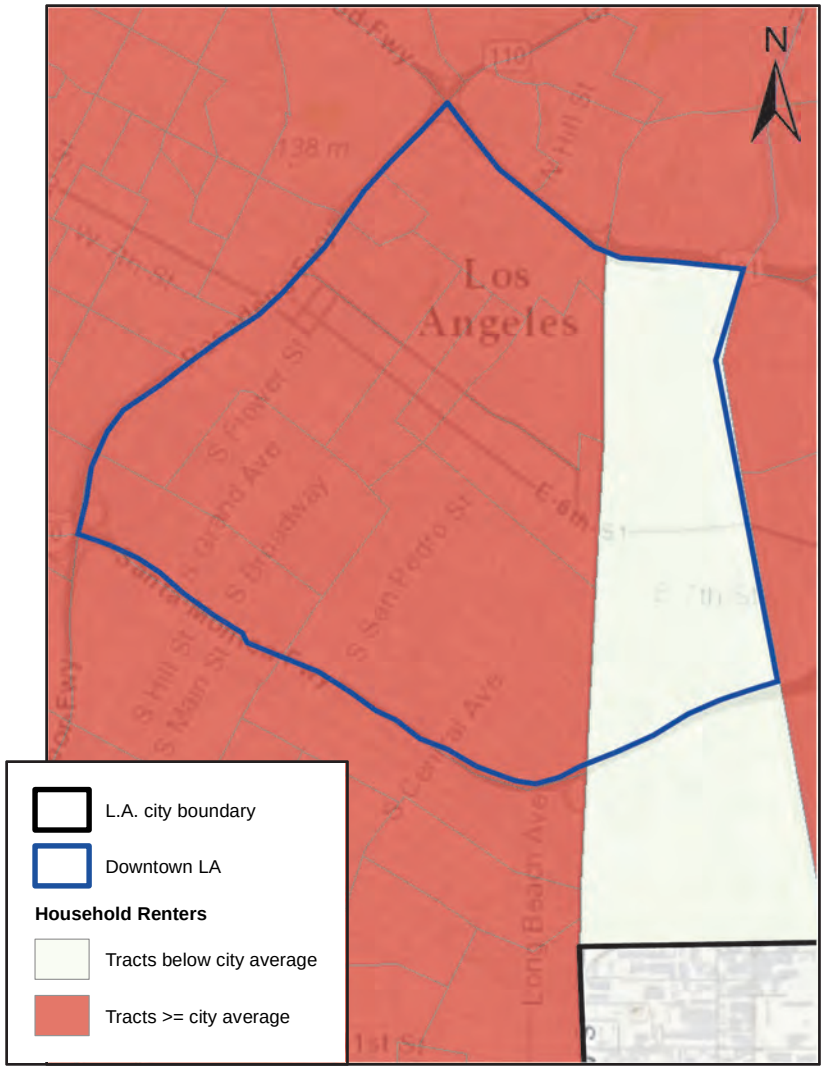


Figure 10: Household Renter >= 60% around Downtown Los Angeles.

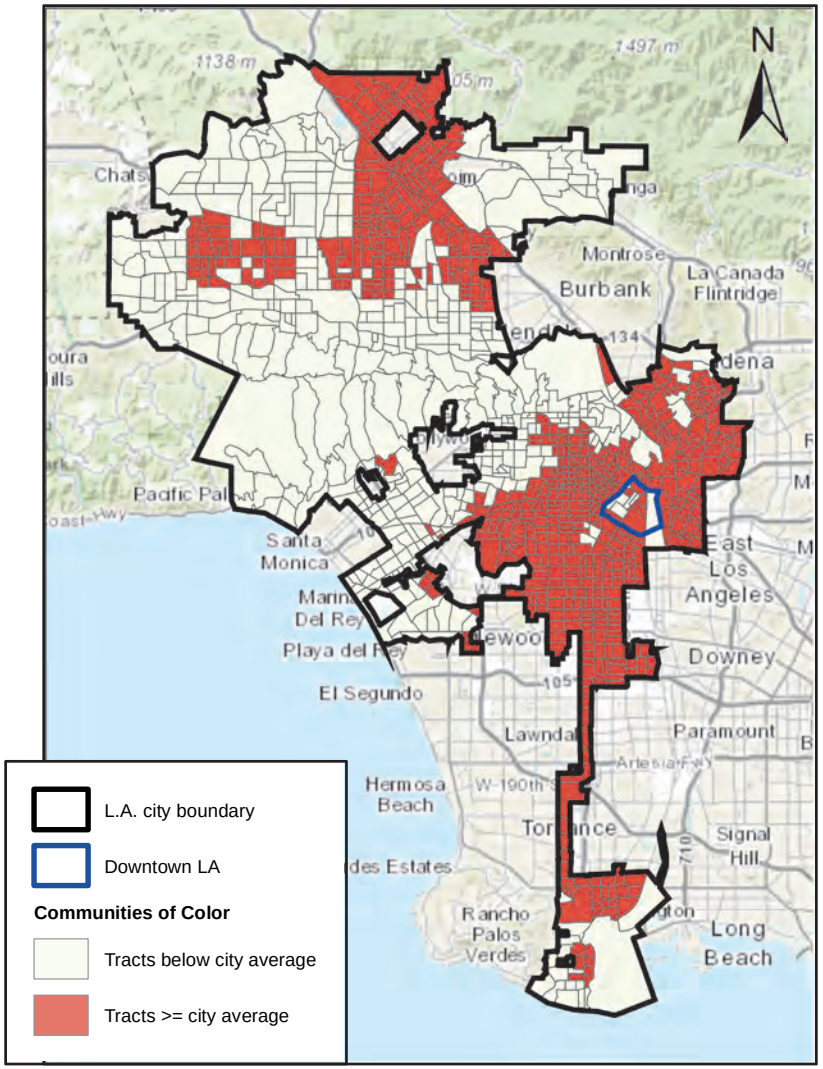


Figure 11: Communities of color >= 70%.

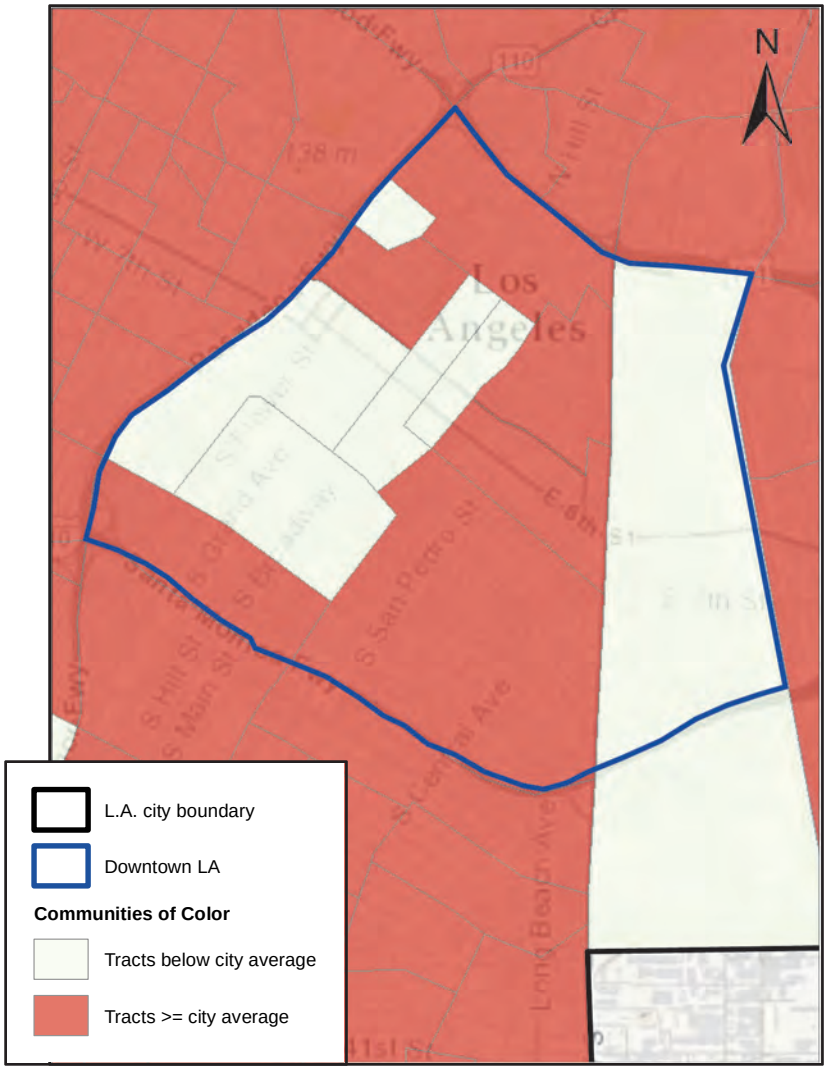


Figure 12: Communities of color >= 70% around Downtown Los Angeles.

4.2.3 Displacement Variable: Median Family Income

Figures 13 and 14 map the vulnerability variable of households at or below 80% of the median family income supplied by the HUD/CHAS data set. The project gave tracts that were equal to or above the citywide percentage of 34% a score of one. The project found census tracts that were above the citywide percentage in mid Los Angeles to the south end of the city. Unlike the other three variables, this data did not find any census tracts that were at or above the citywide percent in the northern part of the city. This led to there being no tracts that scored four points in North Los Angeles in the final baseline map.

Downtown, tracts that scored below the citywide percent ran through the middle of downtown. This variable contained census tracts that were the greatest contributor to downtown not having adjacent tracts with similar scores in the baseline map. Overall, the locations of the individual variable median family income had the greatest impact when the project added the scores of all four variables together to produce the Data Aggregated Method baseline map.

4.2.4 Displacement Variable: Bachelor's Degree

Maps shown in Figures 15 and 16 are census tracts at or below the citywide percentage of 25 years old and older that do not have a bachelor's degree. The study found groupings of census tracts over the citywide percentage in North and South Los Angeles similar to the communities of color. West Los Angles had only a small number of tracts that are over the citywide percentage for this variable. Combining the bachelor's degree variable with the variables of household renters and communities of color resulted in no census tracts in the baseline map having a risk of displacement score that was equal to four on the west side of the city. In fact, most tracts on the west side did not have a score that labeled them at risk of displacement in the final baseline map.

Downtown Los Angeles census tracts for the bachelor's degree variable mirrored the results of the communities of color variable. Both had census tracts in the east and west side of downtown that were below their citywide percentages. Overall, three out of the four variables had East Downtown census tracts with a score of zero. This resulted in East Downtown having a low risk of displacement in the final aggregated baseline map.

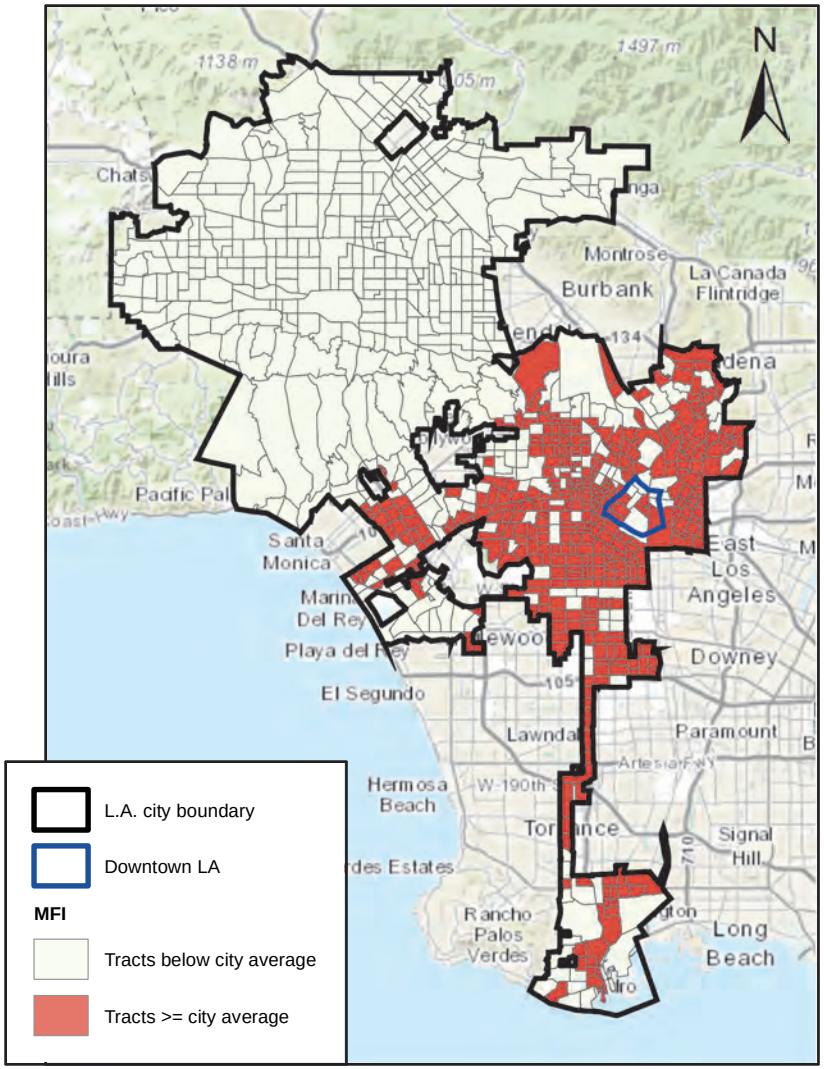


Figure 13: Households that have incomes at or below 80% of the HUD median family >= 34%.

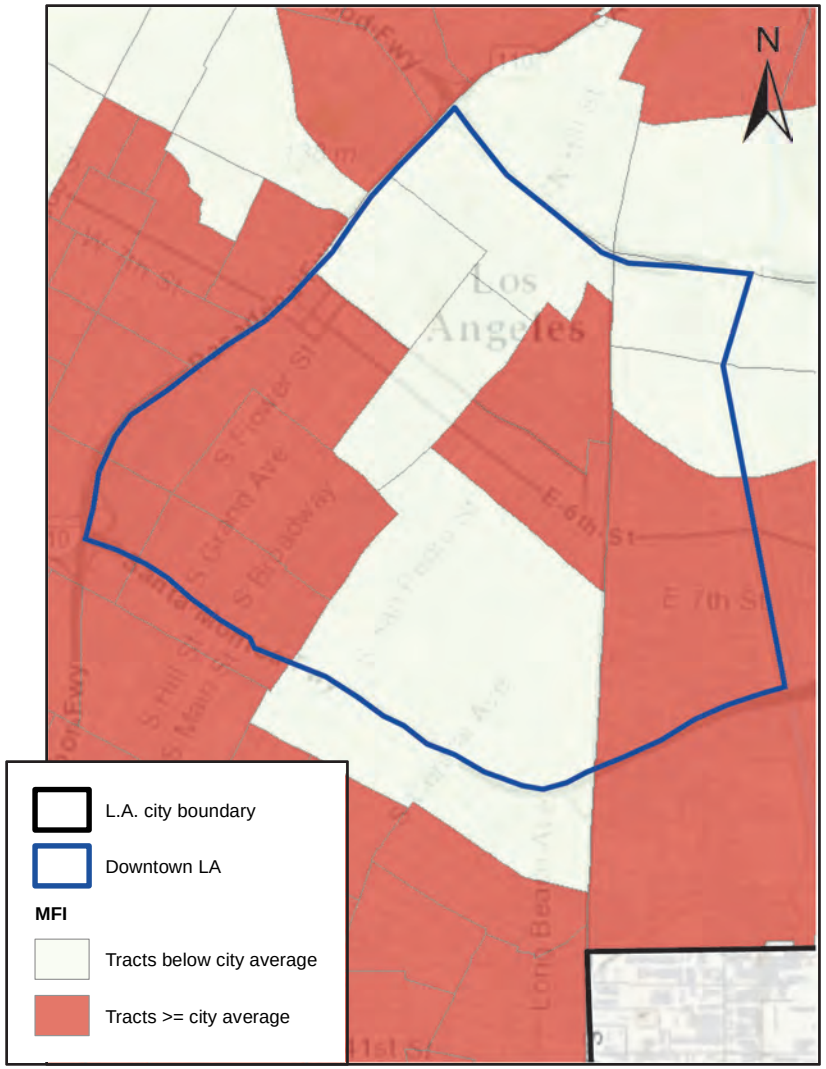


Figure 14: Households that have incomes at or below 80% of the HUD median family >= 34% around Downtown Los Angeles.

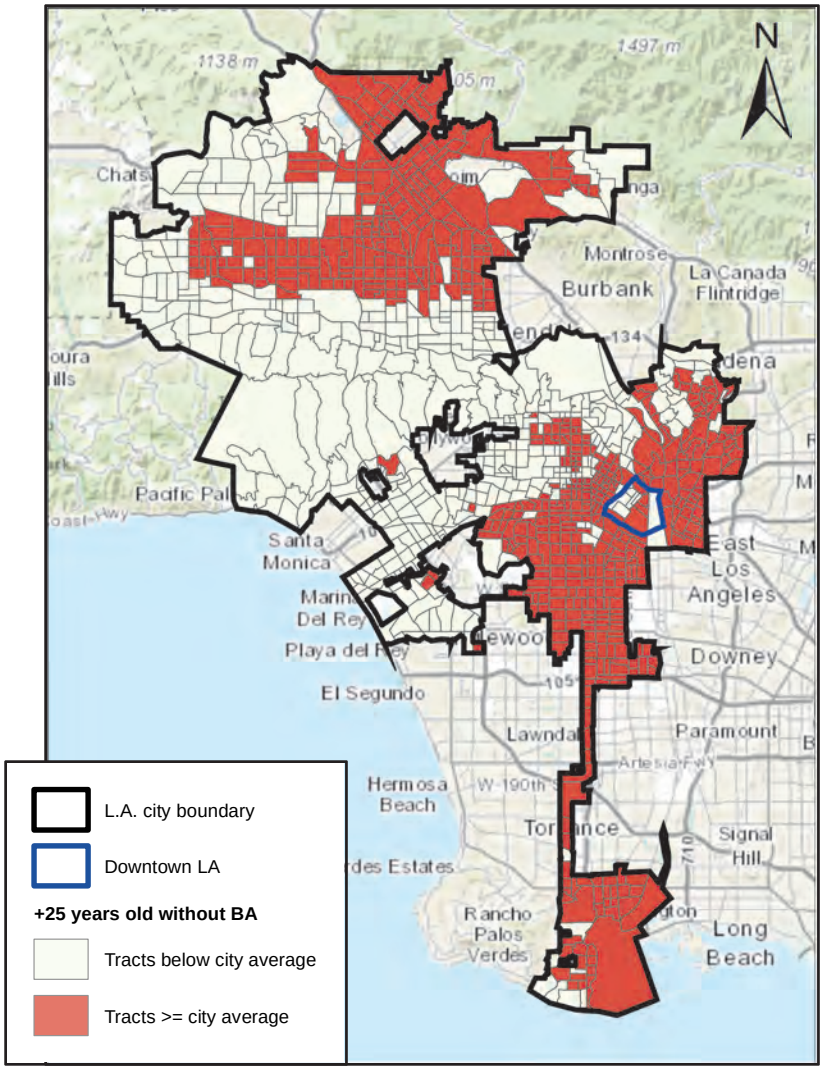


Figure 15: The percentage of the 25 years old and older that do not have a bachelor's degree >= 71%.

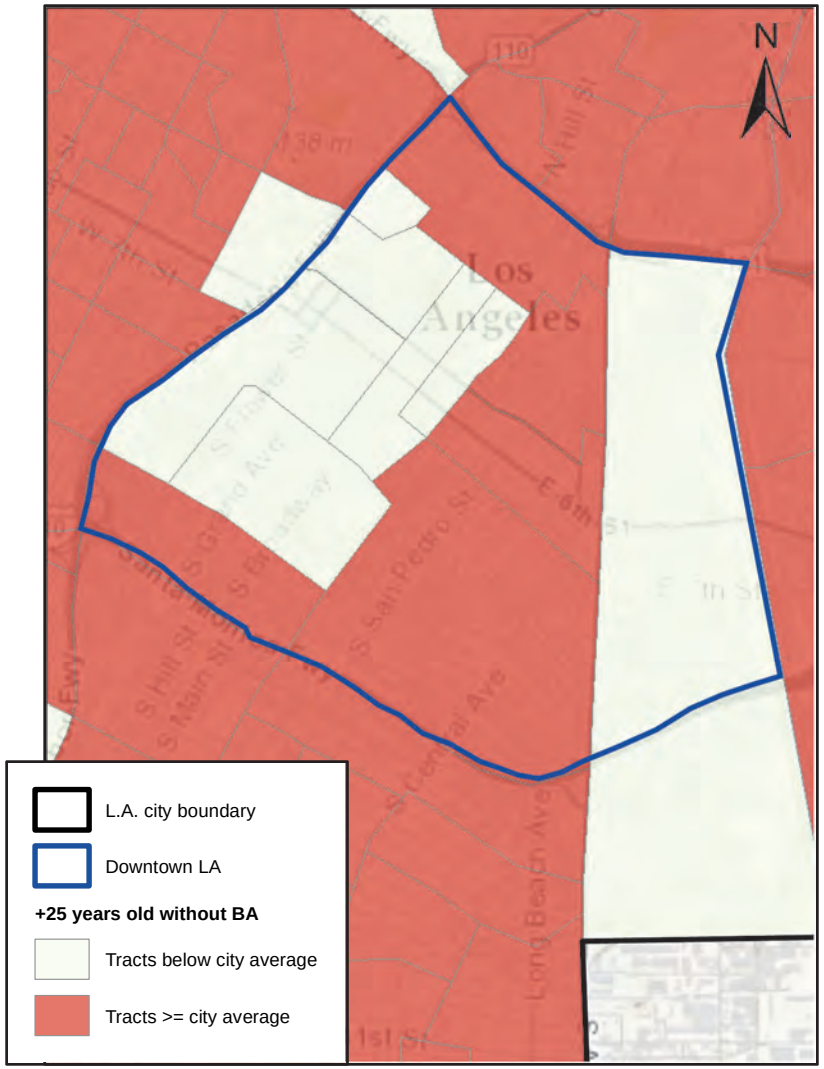


Figure 16: The percentage of the 25 years old and older that do not have a bachelor's degree >= 71% around Downtown Los Angeles.

4.3 Census Tracts at Risk of Displacement in the City of Los Angeles

Recreating the Data Aggregated Method used to locate census tracts that are at risk of displacement produced several maps for review and comparison to the VGI Twitter method. The final baseline map, seen in Figures 17 and 18, is the outcome of combining all the scores from each individual variable: household renters, communities of color, income below 80% of HAMFI, and bachelor's degrees. This method resulted in 461 census tracts having a high-risk of displacement total score out of 1,119 tracts throughout the city of Los Angeles.

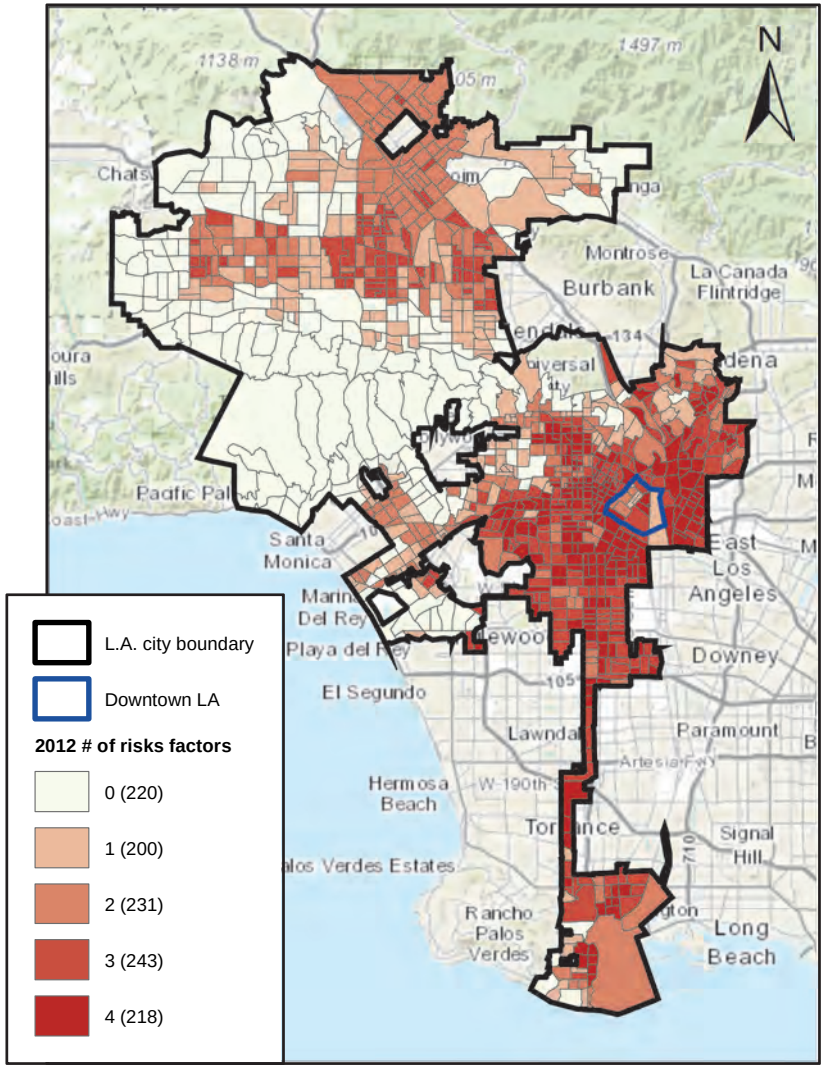


Figure 17: 2012 census tracts in the City of Los Angeles vulnerable to displacement. A score of 3 or 4 meant that a tract was at high risk of displacement.

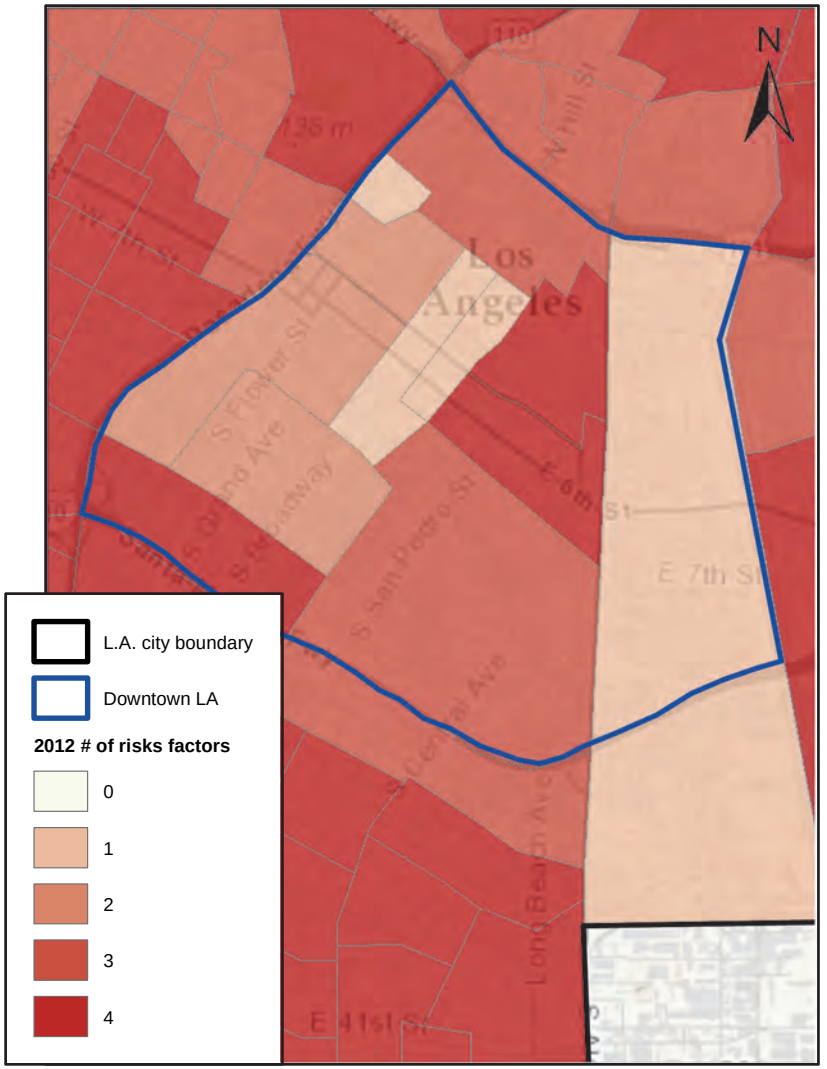


Figure 18: 2012 census tract total vulnerability scores around Downtown Los Angeles. A score of 3 or 4 meant the tract was at risk of displacement.

The Van Nuys area, a district in the northern part of the city, and Downtown Los Angeles had a high amount of tracts vulnerable to displacement. The largest grouping of census tracts that scored a four out of four using the Data Aggregated Method stretched along the mideast of the city, from Hollywood to South Los Angeles through Pico-Union, just west of Downtown Los Angeles. Although census tracts in North Los Angeles did have tracts at risk of displacement, no tracts scored a four out of four. This was attributable to no tracts being over the citywide percentage of median family incomes at or below 80% of the HUD. West and Northwest Los Angeles had the lowest scores on record. These areas included tracts that were in the Westridge-Canyonback Wilderness Park and Santa Susana Pass State Historic Park.

Downtown Los Angeles contained census tracts that ranged from a total aggregated score of one to four with no similar scored census tracts adjacent to each other within the downtown limits. This could be due to the downtown having multiple economic districts within its boundary rather than residential communities.

4.4 Geotagged Tweets Containing Displacement Keywords

The second method employed for investigating displacement in this study is the VGI Method. Figures 19 and 20 shows the results of the VGI Method over the final baseline map of census tracts at risk of displacement. Table 4 lists how many tweets contained the 17 keywords from the Floating Sheep CSV file within the City of Los Angeles compared to Downtown Los Angeles. In both the City of Los Angeles and downtown, the keyword “Starbucks”, and its variants, was the dominant word related to displacement. Citywide it appeared over six-and-a-half more times than the next keyword “neighborhood” and its variants. Downtown, “Starbucks” and its variants show up in tweets twice as much as the next keyword “rent” and it variants.

The study found geotagged tweets using displacement keywords from the Floating Sheep CSV file throughout Los Angeles. The first visual analysis of tweet locations showed groupings in both West Downtown Los Angeles and Hollywood, an area located northwest of downtown. Tweets in these groupings were in and around census tract areas that are at low risk of displacement designated by the Data Aggregated Method baseline map. The VGI Method map located a lower number of geotagged keyword tweets in West and North Los Angeles.

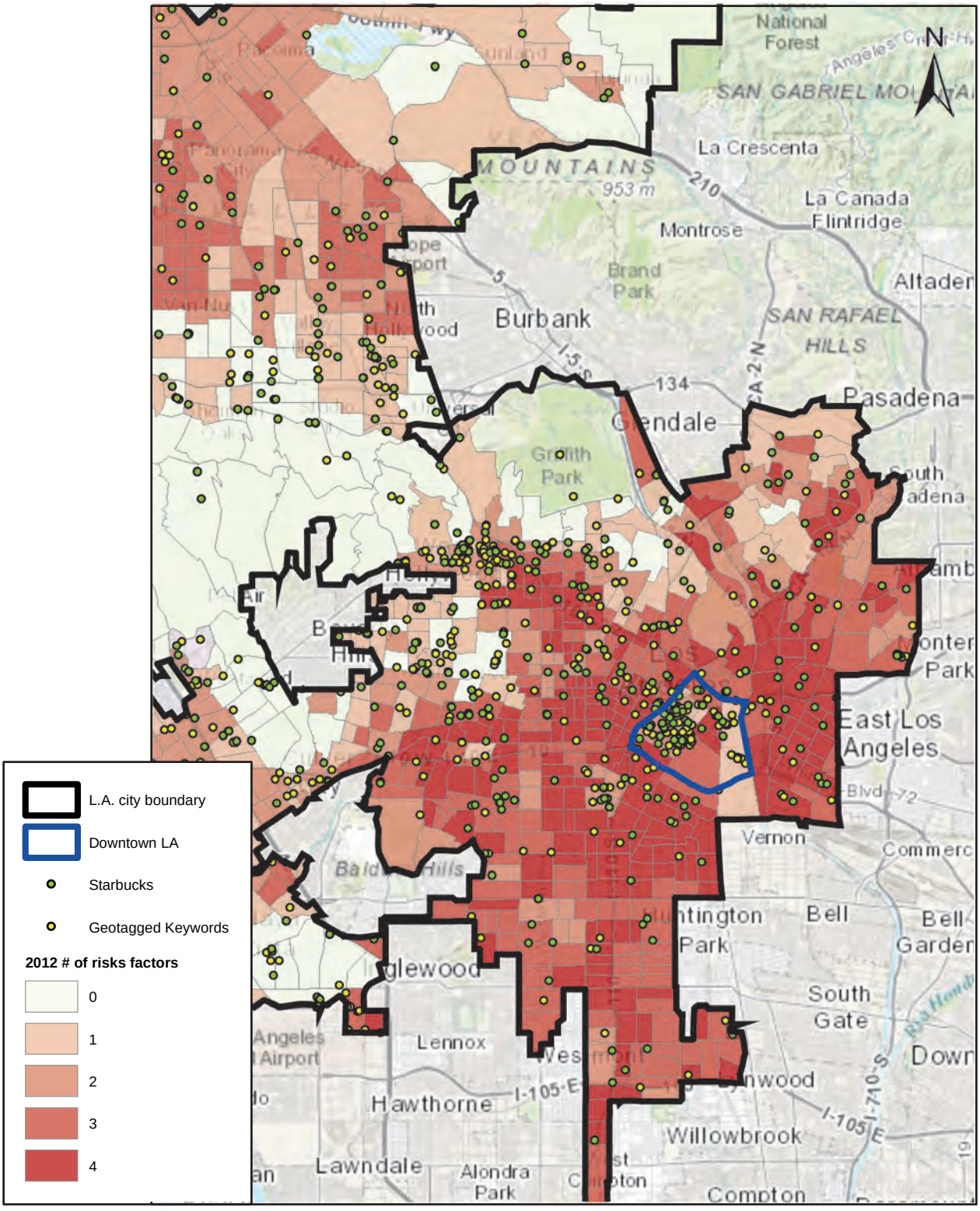


Figure 19: Geotagged tweets in the City of Los Angeles that contain keywords related to displacement.

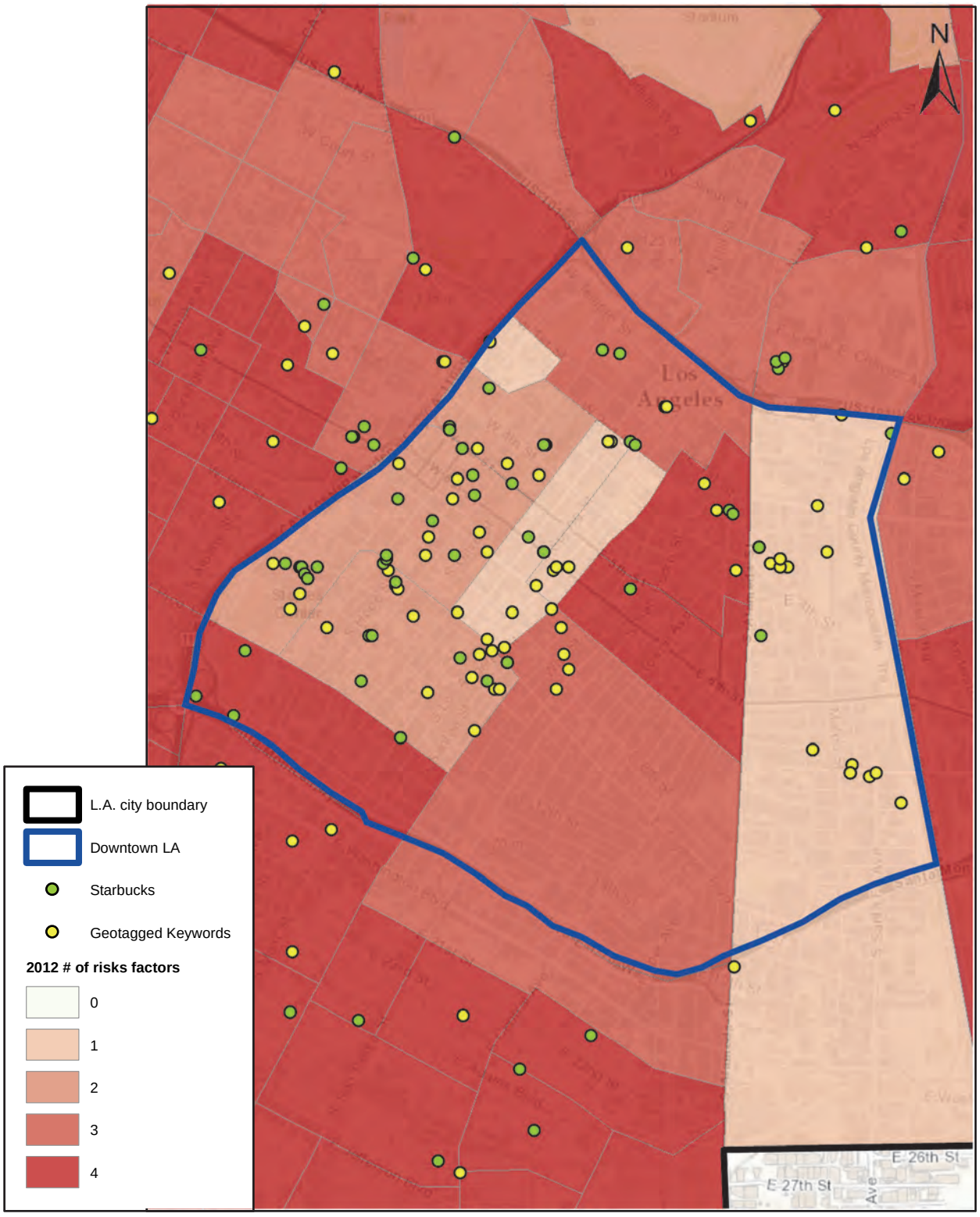


Figure 20: Geotagged tweets around Downtown Los Angeles that contain keywords related to displacement.

Table 4: Number of geotagged tweets that contain keywords in the City of Los Angeles and Downtown Los Angeles.

Displacment keywords in Los Angeles	number of tweets	Displacment keywords in Downtown Los Angeles	number of tweets
Starbucks	1081	Starbucks	114
neighborhood	163	rent	57
rent	161	district	54
hipster	159	loft	21
district	98	hipster	12
retail	69	neighboorhood	10
loft	58	retail	9
urban	45	housing	4
income	31	diversity	4
diversity	28	income	3
housing	15	urban	2
renewal	9	gentri	1
gentri	7	renewal	0
eviction	3	evictin	0
yuppie	2	yuppie	0
revitalization	0	revitalization	0
displace	0	displace	0

4.5 Kernel Density of Keyword Tweets

Understanding that more than one tweet can share the same longitude and latitude, the kernel density tool in ArcMap showed the concentration of tweets within the City of Los Angeles. In Figure 21, the study located keyword tweet densities along the census tracts that outline Central Los Angeles. Bunker Hill, Financial, South Park, Fashion, and Historic districts had the highest density of tweets around downtown. This coincides with the original comparison; however, because of the density analysis one more location had a high amount of tweets related to displacement within the data set.

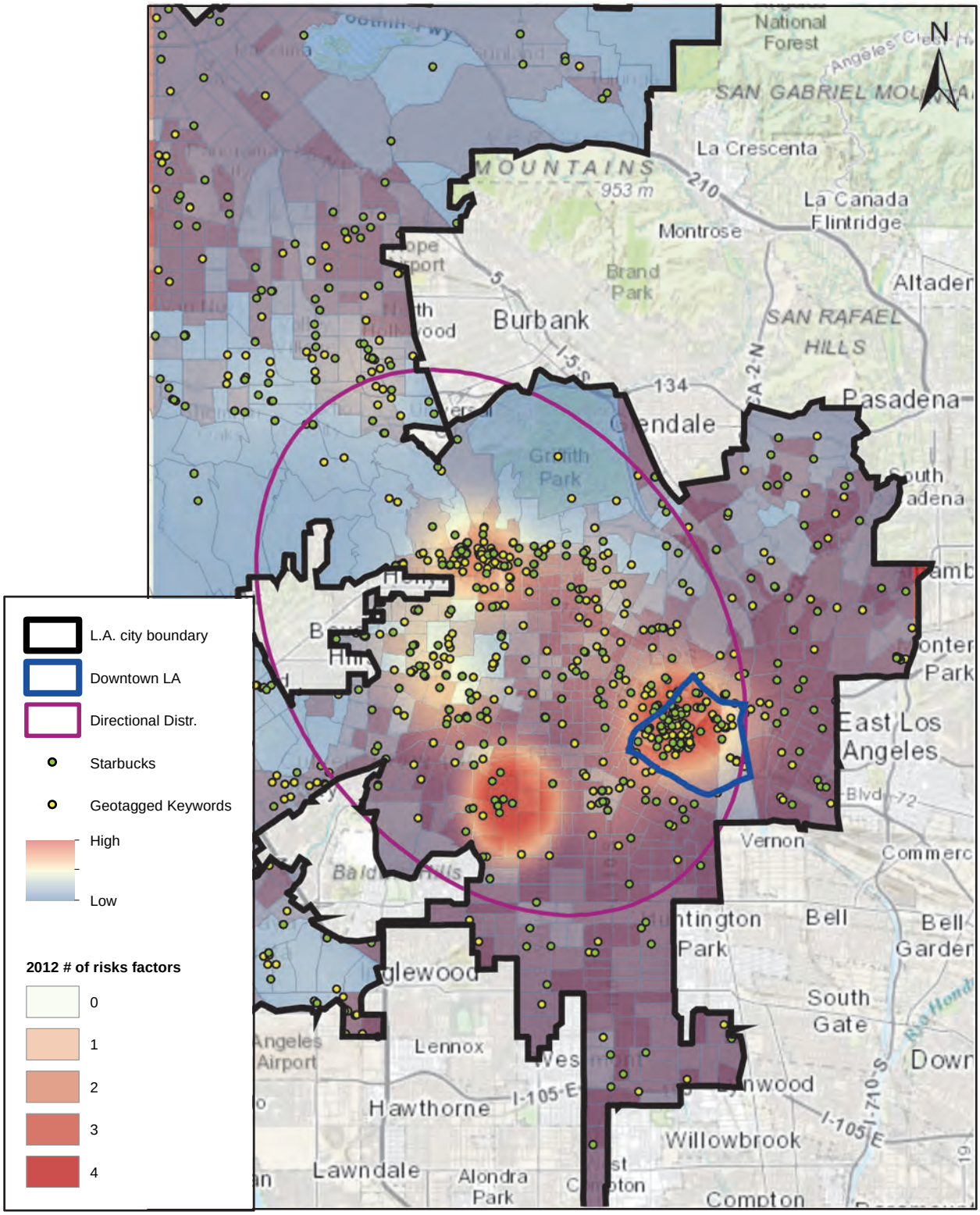


Figure 21: Kernel density map of keyword-geotagged tweets in the City of Los Angeles. The map located high-density areas that outline Central Los Angeles.

The new high-density location centered near Arlington Ave, between Rodeo Road and West Jefferson Boulevard. Referring again to Figure 21, the kernel density location is west of Downtown Los Angeles. A single point in the center of this density area contained 111 geotagged tweets that had keywords for displacement. Of the 111 tweets, 33 contained the keyword “Starbucks” and its variants, with the other 78 a mix from the keyword list. The closest 11 points to the high-density point only contained one or two tweets per location. This is an extreme observational difference from the Data Aggregated Method. Using the Data Aggregated Method, whole census tracts are either vulnerable or not vulnerable to displacement. With VGI data from Twitter, it is now possible to examine below the U.S. Census block level for risks of displacement within the City of Los Angeles.

The geotagged tweets at Arlington Avenue are in the Los Angeles neighborhood of Jefferson Park. Single-family housing with two elementary schools, the Jefferson Library, and the developing Expo Line surround the location of the tweets. On May 1st 2014, Curbed Los Angeles reported about young professionals priced out of locations around the City of Los Angeles and are now considering Jefferson Park as an option for homeownership. Interest in the bungalows around the Arlington Avenue keyword location has raised the median home price to $450,000, an increase of 40.6% in the first quarter of 2014 (Barragan 2014).

Despite the recent increase in the housing market around Jefferson Park, the high-density point close to Arlington Avenue did raise some questions about either the Floating Sheep data set or using social media as a data source. For example, the closest two Starbucks to the 33 geotagged posts at the Arlington Avenue location were 0.9 miles to the west or 1.5 miles to the northwest. The study understands that tweets containing the keyword "Starbucks" do not have to come from within a Starbucks coffee shop, yet the amount of tweets containing "Starbucks" at

this one location were unexpected. The expected result was more dispersed keyword points caused by the two Starbucks coffee shops, multiple points of interest, and several metro stops close to this location.

When further researching the accuracy of harvesting Twitter geolocation for this study, several discussion boards on Twitter's developer website have posts about software errors, the developer's own code, or the data provider used during a tweet's check-in generalizing the tweet's location rather than to the user's specific longitude and latitude. Future expansion of this study would explore a tweet's geolocation in relationship to a point of interest written in a tweet. This would help to explain if Twitter or Floating Sheep is generalizing geotagged locations in high-density points like the one at Arlington Avenue when recreating the study. However, with Floating Sheep having a history of successfully contributing to other spatial studies (Zook and Poorthuis 2014; Shelton et al. 2014; Crampton et al. 2013), this study assumed that all 111 geotagged tweets posted from the Arlington Avenue location are true when analyzing the VGI Method results.

CHAPTER FIVE: DISCUSSION

This study reviewed economic development zones and recent demographic change of the City of Los Angeles to comprehend further the VGI Method results. First, this project analyzed the kernel density map of keyword tweets to four revitalization projects found around the city. Second, to compare the real time results of the VGI Method with historical data the Data Aggregated Method was used again to locate census tracts that had a high rate of demographic change over a twelve-year span. Dense areas of keyword tweets were consistent when compared to locations undergoing current and historical transformations. This project found a high density of tweets in overlapping revitalization projects and census tracts that had demographic change related to displacement.

5.1 Geotagged Keywords within Economic Development Zones

The three most dense keyword areas, Hollywood, west Downtown Los Angeles, and the Arlington Avenue location, were all in areas of at least one of four overlapping Los Angeles revitalization projects (Figures 22 and 23) (Laden 2014; Bitonio 2013). These projects include the 1) Targeted Neighborhood Initiative (TNI), 2) Business Improvement District (BID), 3) Federal Renewal Community (FRC), and 4) State Enterprise Zone (SEZ).

Each economic development zone has its own guidelines for stimulating business growth and community revitalization. TNI focuses on community participation rather than government involvement to improve a Los Angeles neighborhood. The City of Los Angeles gives grant funds directly to low income communities or individual residents for home rehabilitation. BID zones are areas where all businesses are required to pay an additional tax for services beyond what the City of Los Angeles provides (City of Los Angeles 2014). This can include private security patrols, marketing, and general maintenance and cleaning of public

areas. The FRC is a federal program to help stimulate economic development and job growth in lower-income areas of Los Angeles (U.S. Department of Housing and Urban Development 2014). The FRC's goal is to have businesses and developers invest into a community and create full-time employment for current residents. Finally, SEZ provides business and developers financial, technical, training, and tax incentive programs to improve the economy of Los Angeles (City of Los Angeles 2013). They believe these programs will increase local business and raise the standard of living in lower-income communities.

In the City of Los Angeles and Downtown Los Angeles, there are areas where economic development zones overlap each other. The high density of tweets located in Hollywood was in two overlapping zones, the SEZ and BID. In West Downtown Los Angeles the high density of keyword tweets were in all four zones. The Arlington Avenue location was not in an area of overlapping zones, but was in the middle of a SEZ.

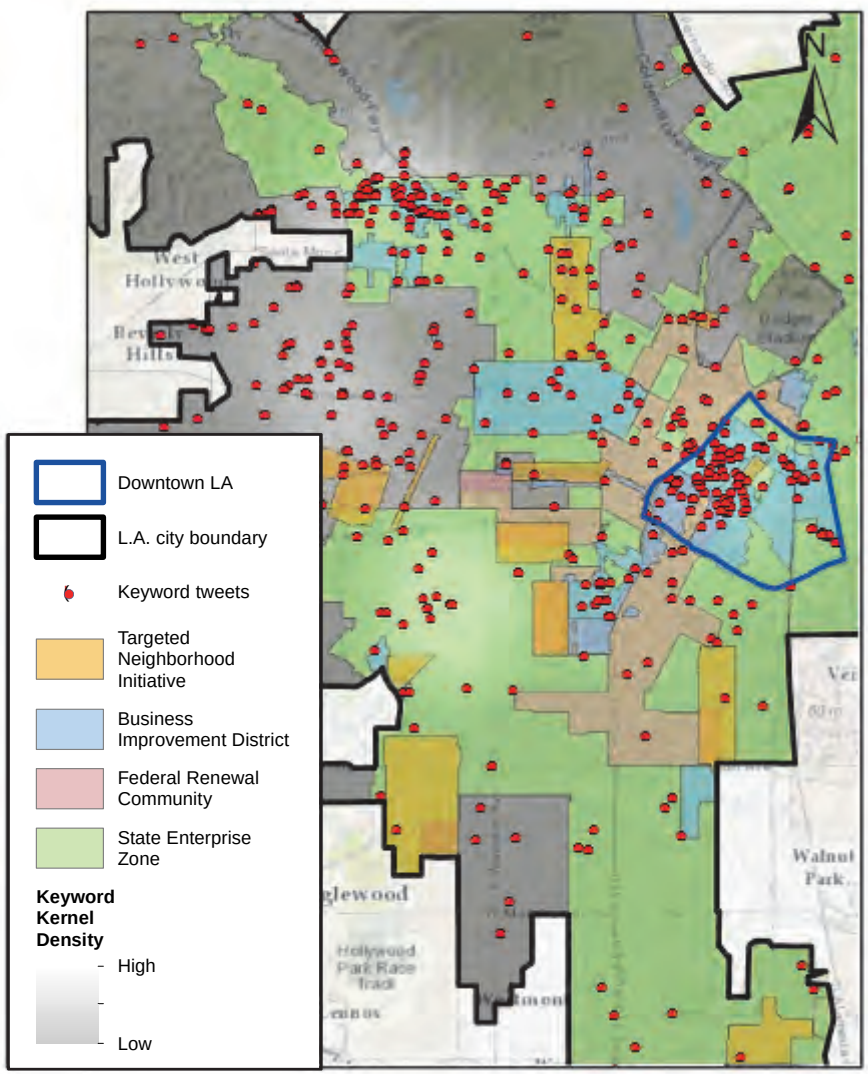


Figure 22: Keyword-geotagged tweets within the City of Los Angeles and Economic Development Zones provided by the Department of City Planning - 2014.

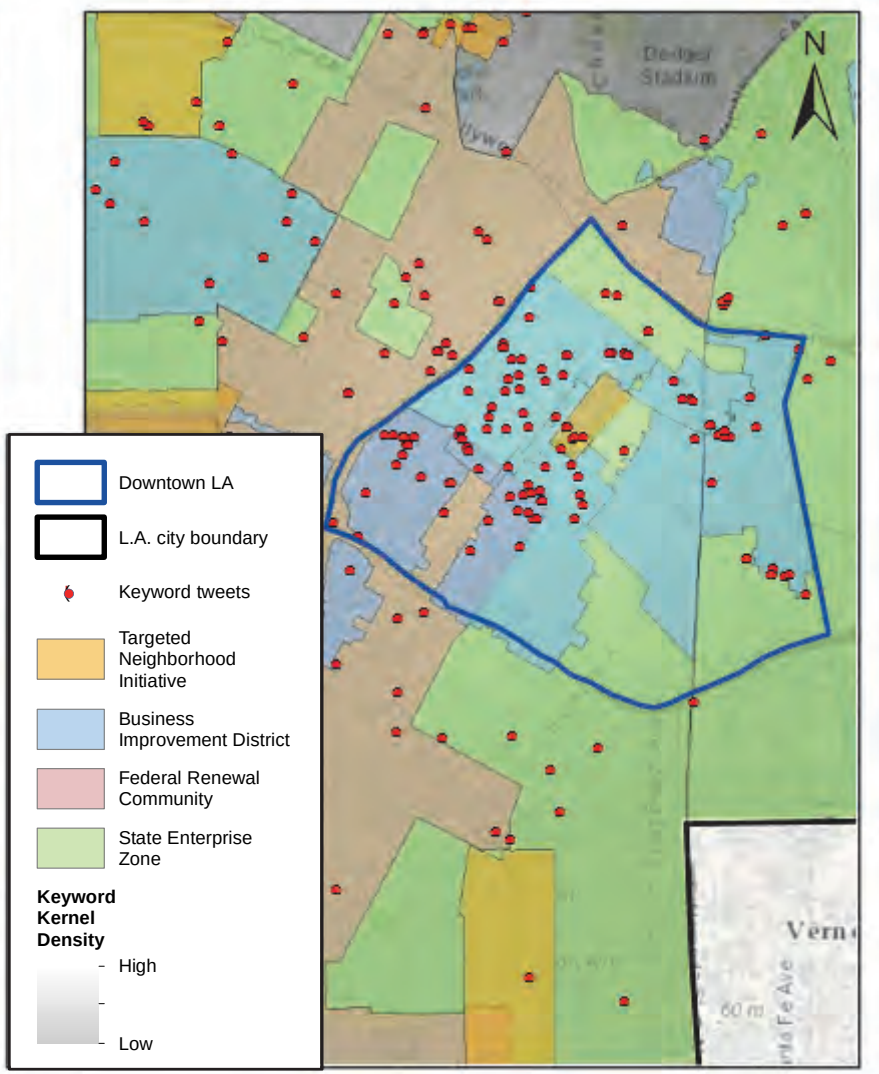


Figure 23: Keyword-geotagged tweets within Downtown Los Angeles and Economic Development Zones provided by the Department of City Planning - 2014.

5.2 Demographic Change in the City of Los Angeles

To give historical background to the census tracts that had a high density of keyword tweets within their boundaries, this study mapped demographic change over a twelve-year period. The study did this similarly to how Bates found tracts that had a high vulnerability to displacement. Bates noted that a significant rate of change of four variables within a census tract could lead to gentrification and later displacement (Bates 2013). These changes and variables are: 1) the increase in the percentage of the white population, 2) the increase in the percentage of bachelor's degrees, 3) the increase in the percentage of homeownership, and 4) the increase in median household income.

This study again used the Data Aggregated Method to give individual tracts a score of zero or one for each individual variable when compared to the variable's citywide percentage. Census tracts were then given a total score of zero to four after the variable scores were added together for each tract. A total score of at least three out of four possible points defined a census tract as having significant demographic change associated with gentrification. The map created from scoring these four variables, Figure 24, showed that dense areas of keyword tweets were in census tracts that had significant demographic changes over the last twelve years compared to citywide percentages. This was also true in examining Downtown Los Angeles in Figure 25.

The demographic change map shows less groupings of similar scores than the vulnerability map produced. Although visually scattered, the study found lower-scored tracts located in the mid to south areas of the city. This is in contrast to the vulnerability map shown in Figure 17, where census tracts in these areas were at risk of displacement. This outcome mirrors Bates' study done for the Portland Plan. Tracts that showed significant demographic change

were not predominantly in areas where tracts were vulnerable to displacement. The tracts were instead in areas that have gone through, or are going through, gentrification.

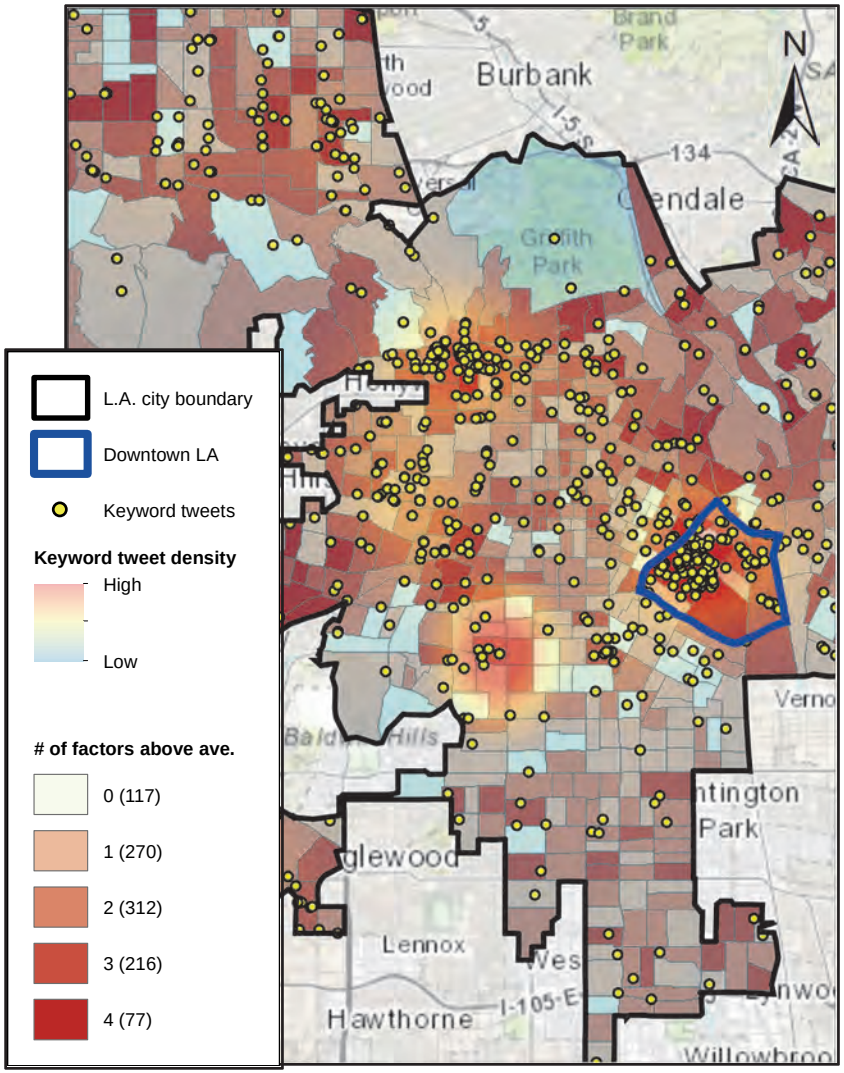


Figure 24: Census tracts from 2000 – 2012 in the City of Los Angeles that showed significant demographic change. A score of 3 or 4 meant that the tract had demographic change related to gentrification.

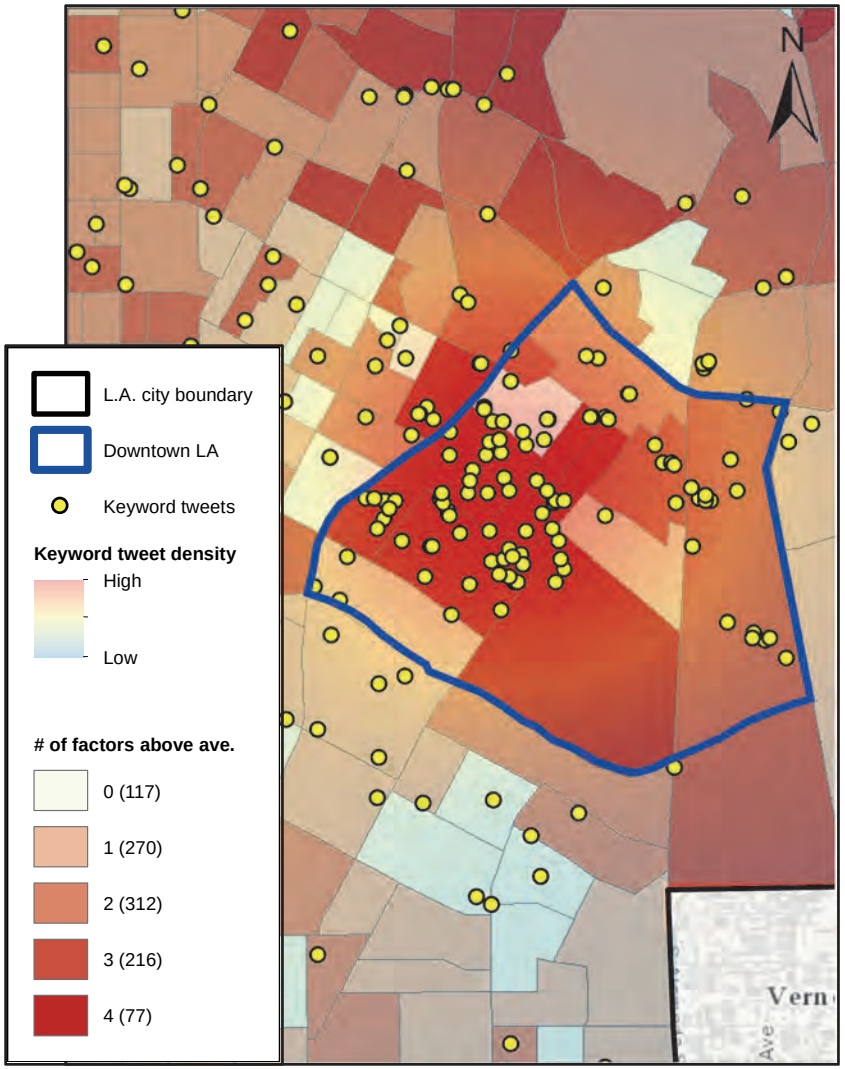


Figure 25: Census tracts from 2000 – 2012 around Downtown Los Angeles that showed significant demographic change. A score of 3 or 4 meant that the tract had demographic change related to gentrification.

CHAPTER SIX: CONCLUSION

This study investigated areas at risk of displacement within the City of Los Angeles, CA by employing two different research frameworks. The first framework, the Data Aggregated Method, recreated an established aggregated method used by the City of Portland to locate census tracts that were at risk of displacement in Los Angeles for use as a baseline for comparison. The Data Aggregated Method used four different census variables and a point system based on citywide percentages to determine if a census tract was at risk of displacement. The study recreated and plotted a final vulnerability map using the Data Aggregated Method for the City of Los Angeles.

The second research framework, the VGI Method, sorted the Twitter data supplied by Floating Sheep by keywords related to displacement. Recognizing that more than one tweet can be posted from the same location, the study calculated point density of the remaining keyword tweets using the Kernel Density tool in ArcMap 10.2. When the study compared the VGI kernel density map to the aggregated baseline map it revealed that geotagged tweets containing displacement keywords were not predominantly located in census tracts at risk of displacement.

To explain the locations of the geotagged tweets the study first went back and overlaid the VGI Twitter method map on top of current Los Angeles Economic Development Zones. The three high-density areas of keyword tweets, pointed out in this study, were all located in at least one of four zones. Los Angeles's hope for these zones is that they stimulate business growth and community revitalization. These overlapping zones could be contributing to areas at risk of displacement and need further investigation.

The study then went back to the Data Aggregated Method to examine demographic change in the City of Los Angeles. Again, the project gave each census tract a score based on four individual variables. Bates considered these variables triggers of gentrification related to displacement. The study found groupings of high-density geotagged tweets containing displacement keywords in census tracts having significant demographic change related to gentrification.

By overlaying the results of the VGI Method on top of the two aggregated baseline maps, this study reached three conclusions about a method that uses VGI versus aggregated data. First, aggregated census data for studying tracts at risk of displacement will always have spatial and temporal constraints when it comes to scale and refinement. Second, VGI data can break through the limitations of aggregated data as shown in the VGI Method used in this study. Third, the spatial data harvested from Twitter comes from a reactionary user base. Twitter in its own business summary describes itself as a public platform for self-expression and conversation. Mark Schaefer (2012) on describing Twitter users and their observations note that users are three times as likely to upload photos and post ratings and reviews, four times as likely to blog, and six times as likely to upload articles all related to topics around them than average internet users. This might explain why the Twitter data could not predict displacement, but instead the study found keywords in locations currently undergoing or finishing revitalization projects.

6.1 Review of Census Data for the Data Aggregated Method.

By using the Data Aggregated Method, with census variables, the baseline maps created for this study showed results that resemble Bates' initial findings. Recreating her method and examining the results reinforced that the study was successful in producing a baseline map for comparing the VGI Method and Data Aggregated Method frameworks. Yet, even with updating

the census, information used in the Data Aggregated Method still could only create outcomes that reflected census tracts from 2012, two years before this study concluded. This study had to assume the outcomes from 2012 data were still true throughout each scored census tract. Not until the VGI Method was executed could this study attempt to observe displacement near real time and below the census tract level.

The City of Portland contracted Bates to create a method to locate census tracts at risk of displacement to help form local policies. Bates' approach was binary, updatable, and used a small set of variables, making it repeatable. The chief weaknesses of this method is that the data used to inform the model are only produced once every ten years (US Census decadal data) for spatial boundaries delineated by the census. Accordingly, any city implementing this approach will be limited by these temporal and spatial constraints. These limitations make it difficult for the method to keep pace with active conditions that leads to gentrification and displacement. The Data Aggregated Method and the variables chosen are a place to start a conversation about displacement with the understanding that the data and maps might not be in harmony with the social and demographic status of any community. To stay in step with neighborhoods that are under threat of involuntary displacement, new tools such as the VGI Method designed for this this study need development.

6.2 Review of an Alternative VGI Method Using Twitter Data

The VGI Method using Twitter has its limitation. To collect a data set before an API program begins harvesting tweets the study must rely on a third party for VGI data. Another is maintenance, software errors, and website hackers create the possibilities of shutting down Twitter and its API resulting in a loss of data or generalizing geolocations. As an example, a longitude and latitude point plotted in ArcMap contained 33 tweets that have the word

"Starbucks" in the geotagged post. The method placed the tweets a minimum of 0.9 miles from the closest Starbucks store. If it is a location error, it might be due to Twitter generalizing tweet locations during website maintenance at the time Floating Sheep was harvesting some of those tweets. Ultimately, though, the greatest limitation of using VGI data from social media is that it is public participation based. If people stop using Twitter then the data set for this study does not exist.

Even with Twitter's limitation, VGI from social media outlets provide a large inventory of data that can be leveraged utilizing geospatial studies. With geo-tagged Twitter data, this study: 1) produced results that are comparable to an established method of locating demographic change related to gentrification and displacement; 2) executed a method that can go beyond an aggregated method's census block level of analysis; 3) improved the timeline for studying displacement to real-time analysis; and 4) introduced a repeatable method that can be implemented, updated, and/or refined depending on geographic location and research context.

When first examined, this study located tweets in census tracts that were at risk of displacement based on the Data Aggregated Method, but by running a density analysis, it showed that most geotagged tweets bordered the grouping of at-risk tracts in Central Los Angeles. The analysis also showed that when overlaid on an aggregated demographic history of the City of Los Angeles the study located tweets in census tracts that had significant demographic change related to gentrification. Therefore, while this new method did not locate displacement when compared to the Data Aggregated Method, it appears the alternative VGI Method can monitor gentrification. This discovery is especially true within Economic Development Zones.

After the study found that the final map and analysis of using a VGI data set was comparable to an established aggregated method when simply looking at their results, the study was then able to highlight why displacement and gentrification research should include the alternative method design in this study. This begins with returning to the definition of aggregated data and recalling that grouping data at a spatial scale can only be as refined at the scale that a study collects it at. This does not hold true for the VGI Method as shown by the high-density geotagged point near Arlington Avenue. During the examination of the kernel density map, this study was able to analyze the City of Los Angeles to Downtown to a single longitude and latitude location. This breaks through even the U.S. Census block level, the finest level the U.S. Census provides for studying displacement.

Next, the VGI Method in this study used Twitter data, a data stream that is ongoing and can be harvested on any day, at any time. This is in contrast to the Data Aggregated Method's reliance on census data that only released at set intervals. This moves the timeline of studying displacement and gentrification to real-time analysis. This study highlights this advantage when recreating the aggregated baseline map. Even after updating the data from the original Portland study from 2010 U.S. Census data to 2012, the results are still two years old when attempting to discuss current displacement. On the other hand, the VGI Method supplied data that was only four months old at the completion of the method.

Finally, the VGI Method using Twitter is repeatable, updatable, and has the flexibility to add local vocabulary defining regional displacement and gentrification. While the Data Aggregated Method relies on the same four variables, the VGI Method can monitor social and demographic trends. A study can select different displacement keywords for the VGI Method using Twitter data based on location. As an example, this study used local news articles and

related webpages to determine keywords related to displacement for Los Angeles. The method could later use different articles and other data sources to change the keyword list if repeated in another city, or even a suburb of that city. Understanding that there are regional dialects, observations, and variables to gentrification and displacement allows this study to personalize the VGI Method to any urban setting, unlike the Data Aggregated Method using census data.

6.3 Future Work

The alternative VGI Method introduced in this study shows how social media is another tool available to policy-makers to help combat inequity, such as gentrification and displacement of longtime residents. Future work should be applied to more cities and urban areas and for longer periods of study. As shown in this study by the location of a high density of keyword tweets in areas undergoing revitalization projects and have demographic change related to gentrification, studies could be performed as a validation of data models in cases where gentrification was believed to have occurred.

A larger data set would reduce any spatial errors that might have occurred by only using the supplied 10% of all tweets during the data collection period of the study. Again, Floating Sheep confirmed that 10%, around one million tweets, would supply enough data to design and evaluate a keyword method, but at this point it is unknown what maps the full data set would produce. A larger data set would also provide a chance to review Twitter profiles linked to posts. The data set used in this study does not supply enough profiles that give demographic information submitted by Twitter users.

This study, as mentioned above, relies on Floating Sheep data that introduces some limitation to the project. A future study would benefit from creating a program that harvested its

own tweets. This would allow studies to collect tweets that are relative to their project from Twitter's API. In this study, organizing and selecting tweets from one million posts took a significant amount of time. A faster way to harvest and sort a data set from tweets that contained keywords related to displacement used at the start of the method would create a shorter turnaround from data collection to map production.

If the study's own API program ran after the study, similar to the DOLLY Project, it would also create a historical data set just for gentrification and displacement studies. This would reduce the overall data set size compared to archiving all tweets, but over time still create a massive amount of information to study the long-term reaction to revitalization projects in the City of Los Angeles and other urban settings. Being able to store, maintain, and manage the study's own data set over a long period will produce a more detailed review of a VGI Method.

REFERENCES

Aggeler, H. *Background Economic and Housing Data.* Background Economic and Housing Data, City and County of Denver Housing Plan Task Force, Denver: City and County of Denver, 2012, 1-49.

Barragan, Bianca. "Young Homebuyers Rapidly Moving Into USC-Adjacent 'Hoods." *Curbed Los Angeles.* May 1, 2014. http://la.curbed.com/archives/2014/05/firsttimer_homeowners_looking_hard_at_uscadjacent_hoods.php (accessed June 17, 2014).

Bates, Lisa K. *Gentrification and Displacement Study: implementing an equitable.* City of Portland: City of Portland Bureau of Planning and Sustainability, 2013.

Bitonio, Joy. "Best Shopping And Dining In The Revitalized NoHo Arts District." *CBS Los Angeles.* May 9, 2013. http://losangeles.cbslocal.com/top-lists/best-shopping-and-dining-in-the-revitalized-noho-arts-district/ (accessed Apirl 25, 2014).

California Communty Foundation. *California Communtiy Foundation.* June 14, 2014. https://www.calfund.org/page.aspx?pid=1158&q=revitalization%20los%20angeles (accessed June 14, 2014).

City of Los Angeles. *Bisiness Improvement Districts.* 2014. http://clerk.lacity.org/BusinessImprovementDistricts/index.htm (accessed June 17, 2014).

—. *Economic Workforce Delevopment Department.* 2013. http://ewdd.lacity.org/business.html (accessed June 17, 2014).

City of Portland, Oregon. *Portland Plan.* 2014. http://www.portlandonline.com/portlandplan/ (accessed January 11, 2014).

City of Portland, Oregon. *The Portland Plan.* City Policy, Oregon: City of Portland, 2012.

Clark, Eric. "The order and simplicity of gentrification: a political challenge." In *The new urban colonialism*, by Rowland Atkinson and Gary Bridge, 261-269. New York: Routledge, 2005.

Clerval, A. "The spatial dynamics of gentrification in Paris: a synthesis map." *Cybergeo: European Journal of Geography*, 2011.

Congress of the New Urbanism. *CNU California.* June 14, 2014. http://www.cnu.org/taxonomy/term/1933 (accessed June 14, 2014).

Conn, Steve. "THAT DIRTY WORD: GENTRIFICATION IS ACTUALLY A GOOD THING." *Design Advocacy Group.* December 25, 2003. http://www.designadvocacy.org/sites/designadvocacy.org/files/2003_12_25_citypaper_cityspace_conn.pdf (accessed May 19, 2014).

Crampton, Jeremy, et al. "Beyond the geotag: situating 'big data' and leveraging the potential of the geoweb." *Catography and Geographic Information Science* 40 (2013): 130-139.

De Verteuil, G. "Evidence of gentrification-induced displacement among social services on London and Los Angeles." *Urban Studies* 48 (2011): 1563-1580.

Edwards, Jom. "Facebook Is No Longer The Most Popular Social Network For Teens." *Business Insider.* October 24, 2013. http://www.businessinsider.com/facebook-and-teen-user-trends-2013-10 (accessed April 11, 2014).

Engels, Friedrich, and David McLellan. *The Condition of the Working Clas in England.* Oxford, England; New York: Oxford University Press, 1820-1895, 1993.

Evans, Will. "Huffington Post." *Southern California Diversity May Be Reaching Its Peak: Report.* March 2, 2012. http://www.huffingtonpost.com/2012/03/02/southern-california-diversity_n_1316349.html (accessed May 6, 2014).

Floating Sheep. *floatingsheep.org.* 2013. http://www.floatingsheep.org/p/dolly.html (accessed April 17, 2014).

Freeman, Lance. "Displacement or succession? residential mobility in gentrifying neighborhoods." *Urban Review* 40, no. 4 (2005): 463-491.

Fujisaka, T, Lee Ryong, and Sumiva Kazutoshi. "Exploring urban characteristics using movement history of mass mobile microbloggers." *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications.* ACM, 2010. 13-18.

Gerber, Marisa. "With gentrification, Echo Park gang members move outside their turf." *Los Angeles Times.* February 3, 2014. http://articles.latimes.com/2014/feb/03/local/la-me-echo-park-gang-20140126 (accessed April 23, 2014).

Goodchild, Michael F. "Citizens as sensors: the world of volunteered geography." *GeoJournal*, 2007: 211-221.

Hammett, Chris. "Gentrification and the Middle-class Remaking of Inner London, 1961-2001." *Urban Studies* 40, no. 12 (2003): 2401-2426.

Indiana Univeristy. "GIS, Census Data, and Maup." *Department of Geography - Indiana Unveristy.* 2014. http://www.indiana.edu/~gisci/courses/g438/lectures/gis_census.html (accessed June 17, 2014).

Karlamangla, Soumya. "Money to L.A.'s 'Promise Zone' could displace poor, experts say." *The Los Angeles Times.* January 10, 2014. http://articles.latimes.com/2014/jan/10/local/la-me-ln-promise-zone-poor-20140110 (accessed April 23, 2014).

Keating, Larry. "Redeveloping Public Housing: Relearning Urban Renewal's Lesson's." *Journal of the American Planning Association* 66, no. 4 (2000): 384-397.

Kim, Susanna. "ABC News World News." *Twitter's IPO Filing Shows 215 Million Monthy Active Users.* October 3, 2013. http://abcnews.go.com/Business/twitter-ipo-filing-reveals-500-million-tweets-day/story?id=20460493 (accessed April 17, 2014).

Koh, Yoree. "Report: 44% of Twitter Accounts Have Never Sent a Tweet." *The Wall Street Journal.* April 11, 2014. http://blogs.wsj.com/digits/2014/04/11/new-data-quantifies-dearth-of-tweeters-on-twitter/ (accessed April 15, 2014).

Laden, Tanja M. "North Hollywood Arts District: An Insider's Guide." *LA Weekly.* Febuary 17, 2014. http://www.laweekly.com/publicspectacle/2014/02/17/north-hollywood-arts-district-an-insiders-guide (accessed April 25, 2014).

Los Angeles County. *Geography & Statistics.* June 14, 2014. http://www.lacounty.gov/wps/portal/!ut/p/b0/04_Sj9CPykssy0xPLMnMz0vMAfGjzOItDCDAyN830MzAyMfHy9U9yNzYwM9MvyDbUREA__aTcw!!/ (accessed June 14, 2014).

Lunden, Ingrid. "73% Of U.S. Adults Use Social Networks, Pinterest Passes Twitter In Popularity, Facebook Stays On Top." *Tech Crunch.* December 30, 2013. 73% Of U.S. Adults Use Social Networks, Pinterest Passes Twitter In Popularity, Facebook Stays On Top (accessed April 15, 2014).

Marcuse, Peter. "Gentrification, Abandonment, and Displacement: Connections, Causes, and Policy Responses in New York City." *Urban Law Annual; Journal of Urban and Contemporary Law* 28, no. 4 (1985): 195-240.

Martelle, Scott. "Saving cities with a kinder, gentler approach to gentification." *Los Angeles Times.* March 7, 2014. http://articles.latimes.com/2014/mar/07/news/la-ol-proposition-13-taxes-gentrification-affordable-housing-20140307 (accessed April 23, 2014).

Masta, Katerina E., and Amy Mitchell. "8 Key Takeaways about Social Media and News." *Pew Research Journalism Project.* March 26, 2014. http://www.journalism.org/2014/03/26/8-key-takeaways-about-social-media-and-news/ (accessed April 15, 2014).

Montagne, Renee. *Gentrification May Actually Be Boon To Longtime Residents.* January 22, 2014. http://www.npr.org/templates/transcript/transcript.php?storyId=264528139 (accessed June 14, 2014).

Papachristos, Andrew, Chris Smith, Mary Scherer, and Melissa Fugerio. "More Coffee, Less Crime? The Influence of Gentrification on Neighborhood Crime Rates in Chicago." *City and Community* 10, no. 3 (2011): 215-240.

Pattaroni, Luna, Vincent Kaufmann, and Marie-Paule Thomas. "The Dynamics of Multifaceted Gentrification: A Comparative Analysis of the Trajectories of Six Neighbourhoods in the Île-de-France Region." *International Journal of Urban and Regional Research* 36, no. 6 (2012): 1223-1241.

Planning and Sustainability. *Gentification and Displacement Study.* 2014. http://www.portlandoregon.gov/bps/62635 (accessed February 5, 2013).

Redfern, P A. "What makes gentrification 'gentrification'?" *Urban Studies* 40, no. 12 (2003): 2351-2366.

Reese, Ellen, Geoffrey Deverteuil, and Leanne Thach. "'Weak-Center' Gentrification and the Contradictions of Containment: Deconcentrating Poverty in Downtown Los Angeles." *International Journal of Urban and Regional Research* 34, no. 2 (2010): 310-327.

Schaefer, Mark. "The Tao of Twitter: Changing Your Life and Business 140 Characters at a Time." *Huffington Post.* August 22, 2012. http://www.huffingtonpost.com/mark-schaefer/business-benefits-of-twitter_b_1823247.html (accessed October 2, 2014).

Semiocast. *Twitter reaches half a billion accounts.* July 30, 2012. http://semiocast.com/en/publications/2012_07_30_Twitter_reaches_half_a_billion_accounts_140m_in_the_US/ (accessed Aprol 18, 2014).

Shelton, Taylor, Ate Poorthuis, Graham Mark, and Zook Matthew. "Mapping the data shadows of Hurricane Sandy: Uncovering the sociospatial dimensions of 'big data'." *Geoforum* 52 (2014): 167-179.

Slater, Tom. "Gentrification of the City." In *The New Blackwell Companion of the City*, by Gary Bridge and Sophie Watson, 571-585. Oxford, UK: Wiley-Blackwell, 2013.

Smith, Neil, and Peter Williams, . *Gentrification of the City.* Routledge: Allen & Unwin, 2013.

The Times editorial board. "San Francisco's bus wars are a proxy fight against gentrification." January 24, 2014. http://articles.latimes.com/2014/jan/24/opinion/la-ed-google-bus-san-francisco-20140123 (accessed April 23, 2014).

Theen, Andrew. "Portland Mayor Charlie Hales talks road woes, jobs, urban renewal area, owning policies that caused displacement." *The Oregonian.* March 14, 2014. http://www.oregonlive.com/portland/index.ssf/2014/03/portland_mayor_charlie_hales_t_3.html (accessed April 12, 2014).

Twitter. *About.* April 17, 2014. https://about.twitter.com/company (accessed April 17, 2014).

U.S. Department of Housing and Urban Developmen. *Welcome to the Community Renewal Initiative.* June 17, 2014. http://portal.hud.gov/hudportal/HUD?src=/program_offices/comm_planning/economicdevelopment/programs/rc (accessed une 17, 2014).

United States. Department of Housing and Urban Development. Office of the Secretary, Eunice Grier, and George. *Urban Displacement: A Reconnaissance.* Department of Housing and Urban Development, 1978.

Urban Dictionary. *Urban Dictionary.* April 23, 2014. http://www.urbandictionary.com/define.php?term=gentrification (accessed April 23, 2014).

Vandergrift, Janelle. "Gentrification and Displacement." *Urban Altruism.* Grand Rapids, MI: Calvin Center for Social Research, 2006. 1-10.

VOX Media. *Curbed Los Angeles Gentrification Watch.* April 23, 2014. http://la.curbed.com/tags/gentrification-watch (accessed April 23, 2014).

Weidemann, Chris. "International Journal of Geoinformatics." *Social Media Location Intelligence: The Next Privacy Battle - An ArcGIS andd-in and Analysis of Geospatial Data Collected from Twitter.com* 9, no. 2 (2013): 21-27.

Zook, Matthew, and Ate Poorthuis. "Offline Brews and Online Views: Exploring the Geography of Beer Tweets." *The Geography of Beer*, 2014: 201-209.

APPENDICES

Appendix A: Four Variables Indicating Demographic Change from 2000 to 2012

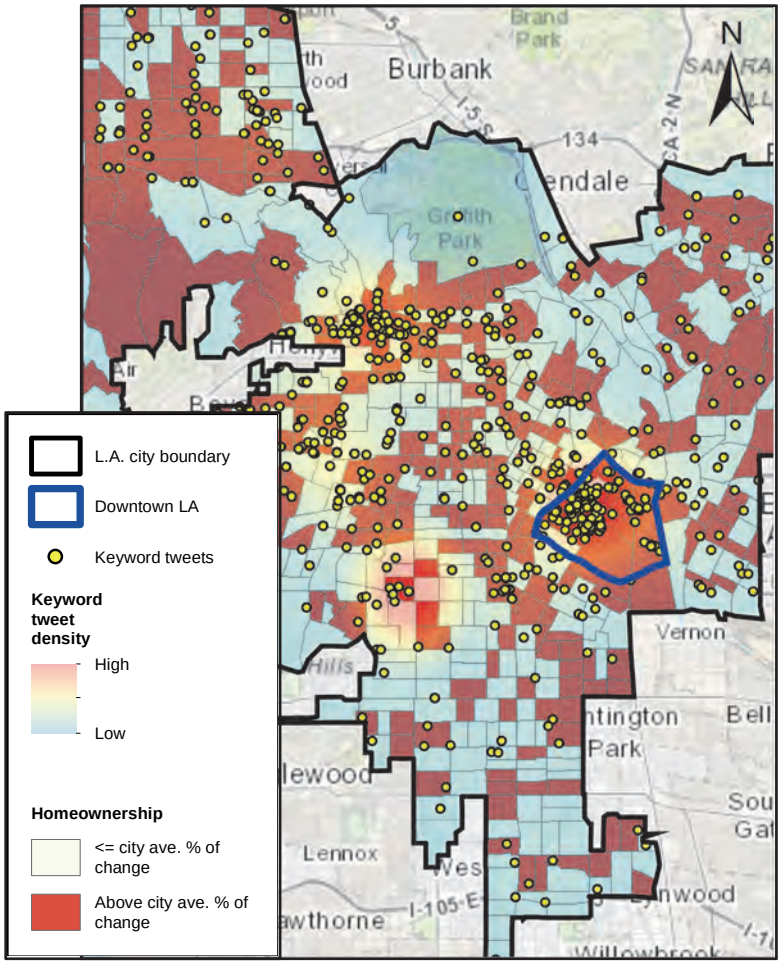


Figure 26: Demographic change factor of census tracts below, equal to, and above the citywide percentage in homeownership.

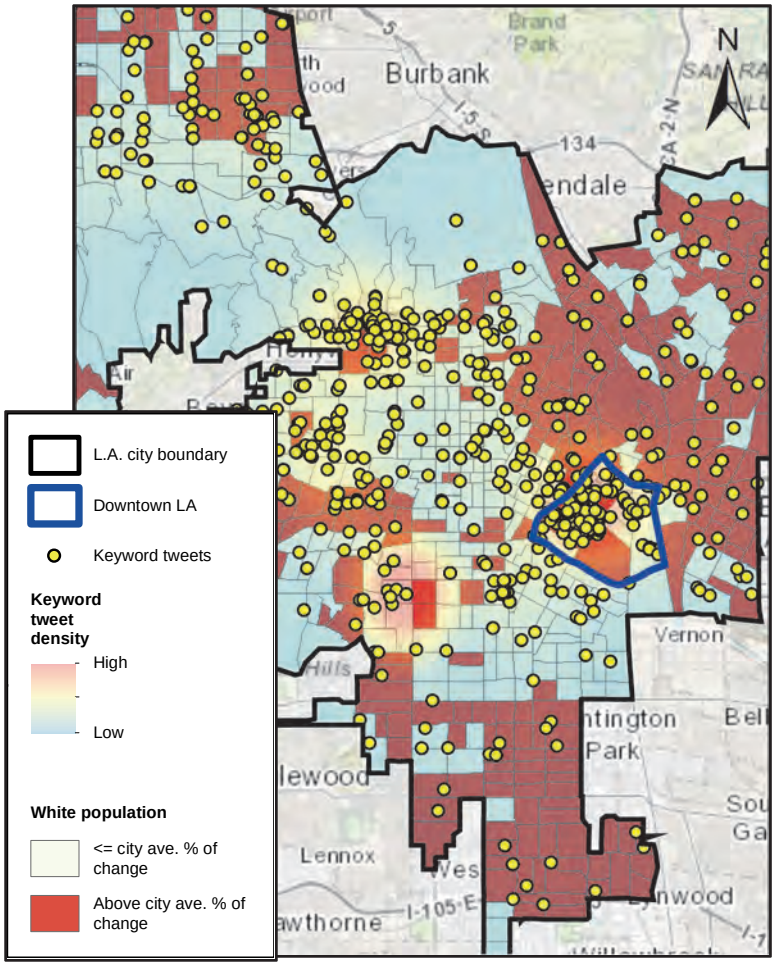


Figure 27: Demographic change factor of census tracts below, equal to, and above the citywide percentage the white population.

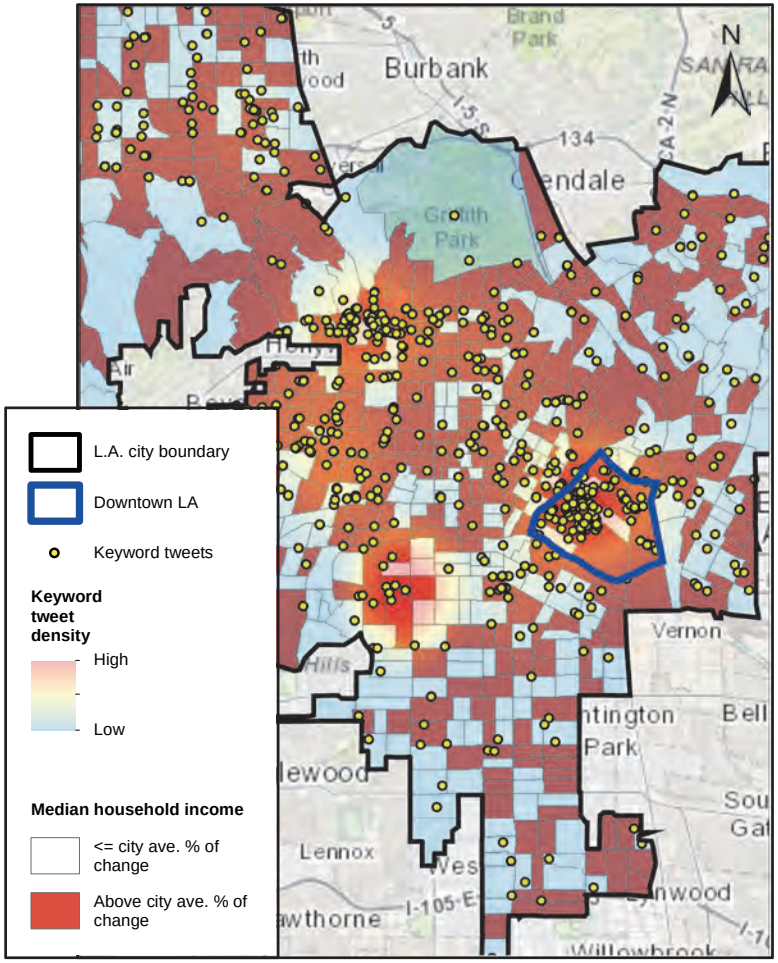


Figure 28: Demographic change factor of census tracts below, equal to, and above the citywide percentage of change in median household income.

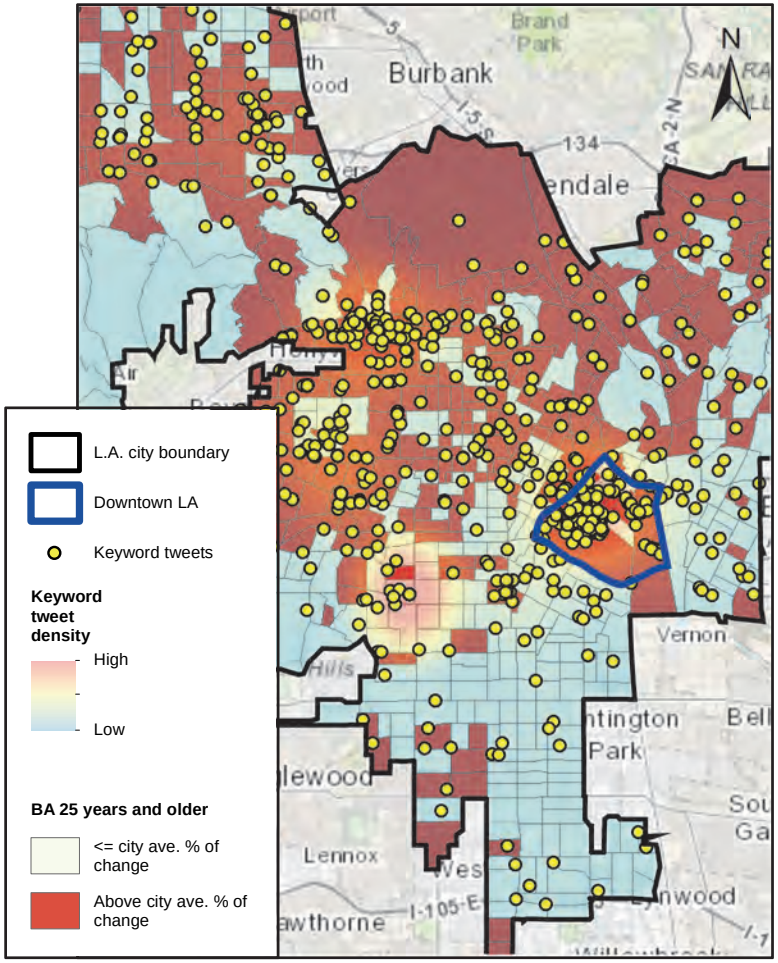


Figure 29: Demographic change factor of census tracts below, equal to, and above the citywide percentage of change of 25 years and older population with a bachelor's degree.

Appendix A: Additional Geotagged Tweet Maps

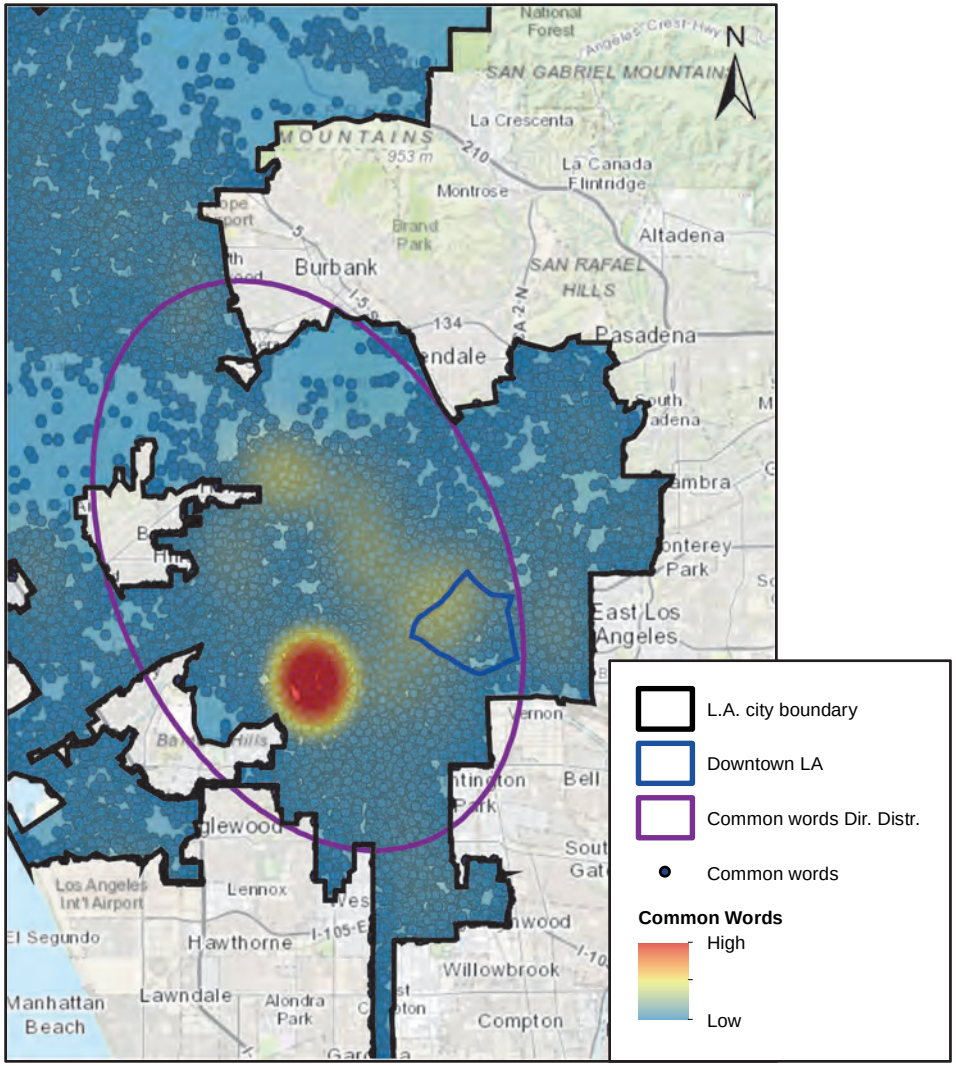


Figure 30: Density map of all geotagged tweets within the City of Los Angeles that contain the ten most common used words excluding pronouns.

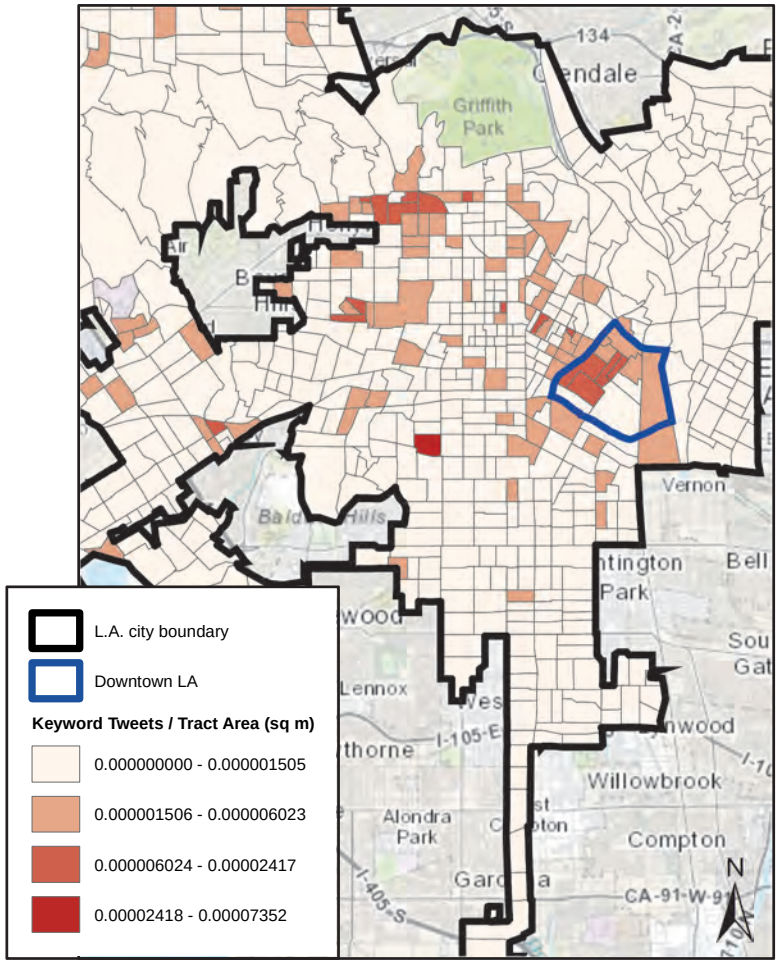


Figure 31: Geotagged tweets containing keywords for displacement per census tract area from August 2013 to January 2014.

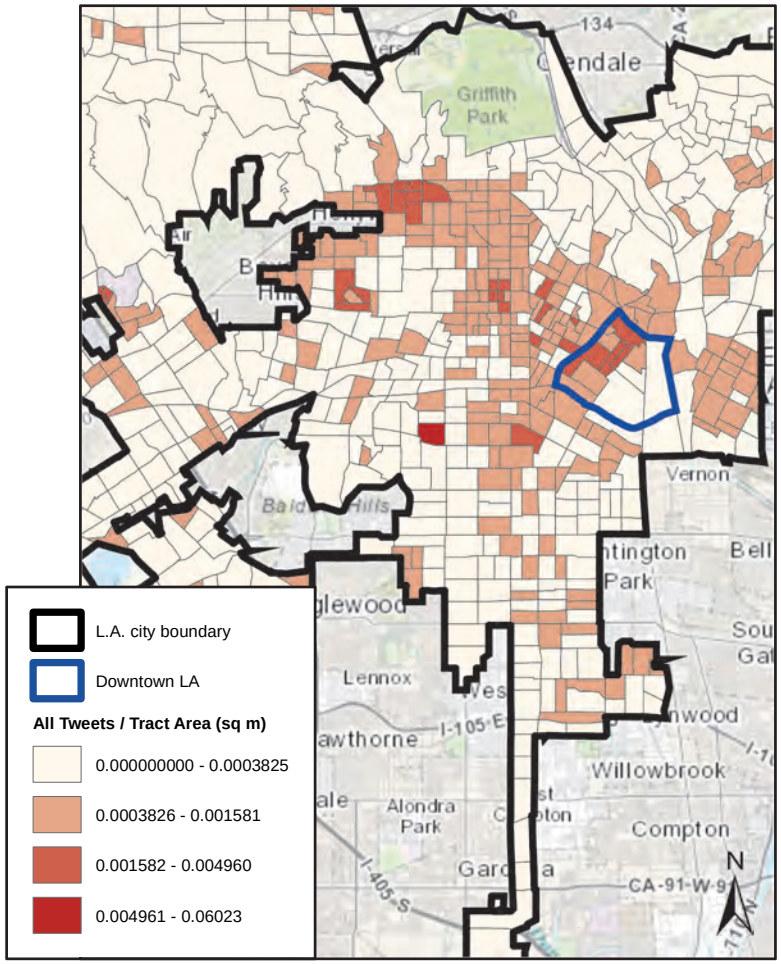


Figure 32: All geotagged tweets supplied by Floating Sheep's DOLLY project per census tract area from August 2013 to January 2014.