Housing Affordability Near Metro's Light Rail – Expo Line & Gold Line

by

Michelle Ramirez

A Thesis Presented to the
FACULTY OF THE USC DORNSIFE
COLLEGE OF LETTERS, ARTS AND SCIENCES
University of Southern California
In Partial Fulfillment of
the Requirements for the Degree
Master of Science
(Geographic Information Science and Technology)

May 2020

Table of Contents

Acknowledgements

To professors, friends, coworkers, and family members that have helped me through this process thank you!

List of Figures and Tables

List of Abbreviations

ACS – American Community Survey

GIS – Geographic Information Systems

HRM – Hedonic Regression Model

Metro – Los Angeles County Metropolitan Transportation Authority

OLS – Ordinary least squares

TOD – Transit-Oriented Development

SFR – Single Family Residence

Abstract

The Los Angeles County Metropolitan Transportation Authority is the third largest public transportation agency in the country — established in 1993 due to a merger between Southern California Rapid Transit District and the Los Angeles County Transportation Commission. The mission of the company is to enhance the "quality of life for all who live, work and play within LA County" (Metro's website). In order to achieve this mission Metro has created services and has invested in bikes, buses, light rails, and rideshare programs. The most impactful investment was into the light rail systems. This rail system includes two rapid subways, four light rails, and ninety-three stations, which spans 105 miles throughout the county. This thesis examined how Metro's light rail system has influenced the value of single-family residential homes in the neighborhoods surrounding the Expo (South Los Angeles) and Gold (East Los Angeles) lines. This study examined how Metro has transformed the zoned area near these two Metro light rails. In order to analyze the impact, the Hedonic Regression Model was utilized to determine the correlation between these light rails and housing values. A series of maps were created to depict the demographic and physical change over a ten year period. The result of the study showed that the impact that Metro's light rails was minimal and did not show significant change in the social, economic, and structural landscapes of these communities.

Chapter 1 Introduction and Background

One of the biggest challenges that the City of Los Angeles has and is currently facing is a housing crisis, which has created ramifications in communities throughout Los Angeles. Like any major city of the United States the current landscape of housing prices, demographics of communities, and zoning of land can be traced back to the formal and informal practices of public and private institutions. Practices like redlining, discriminatory mortgage loans, and racially restrictive housing covenants were strategies used by the government and financial organizations to keep certain economic and ethnic groups restricted to live in specific areas of Los Angeles. Areas of the city such as South Central and the East Los Angeles were created by redlining which created neighborhoods that are historically associated with specific demographics (race and income). These restrictions practices also kept these communities from receiving financial services from banks and the government that would have provided them assistance to buy homes. The policies enacted in the past have had financial implications for the neighborhoods and community members, causing many of these individuals to become long time renters or have devalued housing prices.

In order to reverse some of these practices from the past, new strategies and programs were created to help economically disenfranchised communities. The city played an important role in creating new programs to help with the housing crisis – opportunity zone and first-time homebuyer programs. Another program that is currently having an enormous impact on the social, and physical layout of Los Angeles is the reinvestment of major projects like Metro's light rail system. Los Angeles Metro has received nearly two billion dollars from federal and state grants to finance this project. This funding has provided Metro the ability to expand and create new rail system on top of the 105 miles throughout the County, four rails, two subways,

and ninety-three stations. Los Angeles Metro has also acquired land such as parking structures and parking lots further expanding their economic influence on communities.

In order to determine whether Los Angeles Metro has had an impact on these communities this thesis examined South Central and East Los Angeles neighborhoods. These two communities were examined to determine the change in the past ten years and the type of changes that have occurred since these light rails opened to the public. The South Central and East Side areas of Los Angeles are historically minority communities and, to some extent, low-income communities as well. However, due to new city projects like the Metro's light rail reinvestment into these areas has grown. This change can be seen by stations in Boyle Heights and Crenshaw neighborhoods. This reinvestment into these areas can be seen through conversations, articles, and studies over the last couple of years have talked about gentrification, displacement, and redevelopment in these communities.

The scale of this project is enormous, and it will have economic and social ramifications for these communities. The financing, policy, and outreach needs to be holistically understood by all in order to be an inclusive project. These practices need to be examined in order to determine all the impacts it will have on people, so policy can be enacted to protect the most vulnerable community members. While also giving those communities the agency, they need to determine the best course for them financially. This thesis examined how Metro's light rails have transformed these areas regarding housing value, demographics (income and race), and the physical structures of single residential housing to other forms of housing (apartments & condominiums). Theorizing that these indicators have changed over the past ten years due to the development of the light rails at .5-mile proximity to single family residence housing.

1.1 Motivation

The motivation for this study examined whether Metro's light rails have increased or decreased the value of housing prices in the South Central and East Los Angeles communities. To understand the implications of what occurred (with regards to changing demographics and the physical zoning of the areas by the Metro's light rails). There needed to be an in-depth analysis between housing prices and the light rails to determine if there was a correlation between these two entities. By understanding if there was some type of correlation, then it becomes be easier to understand the changes that are occurring to the neighborhood's demographics and physical layout.

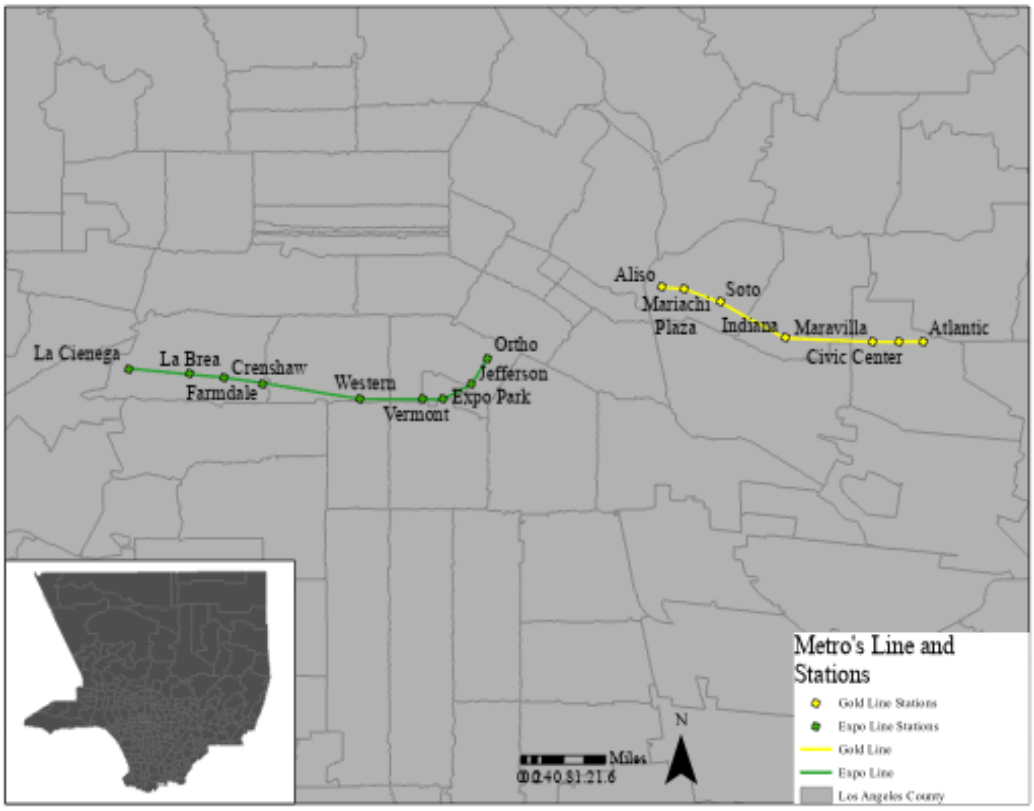


Figure 1: The study area the County of Los Angeles and Metro Lines (Gold & Expo).

1.1.1 Societal Impact

This study attempts to inform the general public, public leaders, planners, and city officials about the negative affect that there could be, so they can have a more holistic understanding of the implications about these light rails. This spatial phenomenon has had financial impacts on the homeowners and renters of these areas, by informing these individuals it better suits them to prepare financially. It also gives community members, that do not have property, the agency as stakeholders to have a voice in the process of planning, and policies that will affect their communities. It is important to involve the homeowners and renters in this process, because it is their community that will face the most displacement.

Like these stakeholders the government also plays a role as well and they too are stakeholders. By knowing the ramification of these lines, they can implement policies that can help these vulnerable communities from being displaced. As housing prices and rent continue to be unaffordable throughout the county, displacement will continue to increase. With a city that has increasing housing values, more and more people are spending more than 30% of income on rent. City of Los Angeles and Metro need to have a better understanding of their stakeholders needs by involving them in this process in order to avoid displacement. Even as these light rail lines were established and are currently being developed, conversations and policies can be created to stop the potential displacement.

1.1.2 Scientific Knowledge

Geographic information science is a very useful and powerful field that merges the social and scientific fields. There is an enormous amount of literature regarding GIS, transportation, and housing values that contribute to the conversation of GIS's role in studying social and scientific phenomenon. While past studies conducted focus heavily on statistical outputs (such as

regressions to determine their outcome), this study heavily focused on presenting a story through a regression model and cartographic maps. By presenting the data in this manner, it allowed the maps to depict the structural and social changes of a city with the statistical evidence to reinforce the story that the maps attempted to tell. The research conducted tested the working hypothesis with a regression model and depict the geographic data that was used to study this spatial phenomenon – demographics and structural changes of these neighborhoods.

1.2 Research Questions and Working Hypothesis

- Is there a correlation between these light rails and housing values?
- Has Metro's light rails made these areas unaffordable which has led to displacement?
- How have these lines changed the demographics and physical layout of these areas?

The working hypothesis was that Metro's investment into the Gold and Expo lines might have an influence on the value of housing which could attract new interest into these South and East Los Angeles areas. Potentially leading to the displacement of community members that are historically tied to those areas. In order to test this hypothesis and to answer these questions, this study is broken down into two parts: first part studies the single-family housing within a .5-mile proximity to these two Metro light rails from the years of 2008 to 2018. A regression model was conducted to determine if there was a correlation between these lines and the increase housing values. This regression model excludes mixed housing developments, condominiums, and plazas, due to the limitations of technical resources. The second part of this study focuses heavily on maps that illustrate the differences in the demographics and structural layout of these two neighborhoods.

1.3 Project Overview

The project consisted of two aspects that illustrated the correlation between Metro's light rails and housing value in these neighborhoods. The first component was the Hedonic regression model that utilized parcel data, which consisted of data of each single-family housing from 2008 to 2018 of these two communities. This regression illustrated the influence that these lines could have on housing value. The dependent variable was the total value of housing with the independent variables being the distance to the metro light rail, size, and other external outputs. The second aspect of this study was the creation of a series of maps that illustrated the difference in demographics and physical layout of housing near these lines within 2008 and 2018. In this series of maps two indicators (income and race) were utilized to illustrate the change in these neighborhoods. Depending on the outcome, the regressions either reflect what the maps illustrate, or not.

To establish the research effectively defining displacement, affordable housing, and race were important to create the foundation of this study. It is crucial that these words are understood, because these words are subjective and can change depending on who used it. Displacement and affordable housing were defined based on other research articles related to this topic. Income and race were used to determine if displacement occurs and whether the unaffordability of these areas influenced by housing price which will led to this displacement. Race was defined because the manner in which that data is collected by the American Community Survey. The difficulty with race was based on ethnicity and race with the Latino population: Hispanic and Non-Hispanic. Race was defined and determined based on how the Census Bureau defined this word and how they categorized this data.

Currently, Los Angeles is going through a revitalizing transformation that is affecting every single Los Angelino. There was an influx of new investments from the public and private sectors that have sprung new projects throughout the city. New projects such as the expansion of the Metro light rails, housing complexes, and sport stadiums just to name a few projects are transforming the city. The main purpose of this research was to give an in-depth analysis to city planners, and government officials, so they can create policies to prevent displacement of community members that are historically associated with those neighborhoods. Housing affordability affects everyone in the city and if there is no policy implemented this problem will only continue to worsen.

1.4 Background

1.4.1 East Los Angeles Neighborhood

In 2003 East Los Angeles area was a very hegemonic community with the majority of the population having similar social, economic, and demographic characteristics. The total population was 124,283 with 96% of the population being Hispanic or Latino descent. The median income for a household was $28,544 annually with 30% of families living below poverty. 43% of the people living in this area had a 9th grade education, 23% had no high school diplomas, and 17% of the population had a high school diploma (ACS). In 2003 the Gold line was introduced to the public for ridership in an attempt to connect the city. During this time the majority of individuals that commuted to work utilized a car as a means of transportation.

Since the opening of the Gold line the community still remains very hegemonic with 96% of the residents of the community being Hispanic or Latino descent. However, changes in other demographics and social characteristics have changed. The median income has grown to $45,000 annually and the poverty levels dropping to 25%. Education levels have also increased with 32%

of the residents have 9th grade education, 16% of residents do not have a high school diploma, and 24% of residents that have higher than a high school diploma. Another change can be seen in the community which is public transportation as of 2017 public transportation has increased. Overall, education levels and income levels in these neighborhoods have risen.

Table 1: Characteristics of East Los Angeles from 2009 to 2017.

Characteristics	2009	2013	2017
Total Population	123,136	127,897	123,905
Latino/Hispanic	97.8%	97.6%	96.4%
Black/African American	0.2%	0.8%	0.2%
White	1.1%	2.4%	1.8%
Income	$37,128	$37,982	$42,544
High School Education	22.9%	23.3%	23.8%

1.4.2 South Los Angeles Neighborhood

In 2009 the Expo line was introduced to several neighborhoods in South Los Angeles including -- Baldwin Hills/ Crenshaw, West Adams, Leimert Park, Jefferson Park, and Exposition Park. The total population in 2000 was 117,861 with 96% of the population being black or African American descent. The median income for a household was $30,000 annually with 27% of families living below poverty. 46.3% of the people living in this area had a 9th grade education, 26% no high school diploma, and 18% of the population had a higher than a high school diploma. In 2009 the Expo line was introduced into to the public for ridership. During this time most individuals that commuted to work utilized a car as a means of transportation.

Since the opening of the Expo Line in 2009 the social and demographic indicators of these neighbors have changed. According to the ACS the education levels have risen from 74% of the population have a high school diploma to 77% with 28% of the population received a bachelor's degree. The median income has also risen to $46,000 annually and the amount of families living below the poverty line has dropped to 19%.

Table 2: Characteristics of South Los Angeles from 2009 to 2017.

Characteristics	2009	2013	2017
Total Population	248,666	126,496	171,703
Latino/Hispanic	48.5%	NA	50%
Black/African American	35%	NA	30%
White	15%	NA	20%
Income	$32,965	$49,138	$36,884
High School Education	66%	73.7%	76.4%

Overall, these neighborhoods in South and East Los Angeles social characteristics have improved. Levels of education have increased, and poverty levels have dropped. As these areas see an increase in education and income these areas remain some of the poorest communities. The ACS reported the County of Los Angeles median income in 2017 was around $70,000 which in comparison to South Los Angeles and East Los Angeles the median income was around $42,000. Conveying that these communities still are considered low income and poorer neighborhoods. These areas also remain very hedonic with one ethnicity live in these areas at a higher rate than any other ethnic group.

Chapter 2 Literature Review

This literature review examined that various amounts of work on research on housing patterns, displacement, the light rail transit, and transit-oriented development. This amount of information established a foundation to determine what approaches and limitations that should be considered. These literature reviews synthesized the connections (negative or positive) that these two entities have one each other. The second part of this thesis examined what methods were the best approach to determine what type of correlation could occur. Lastly, this section sought to understand the social ramifications behind the public transportation and what potential problems could occur.

2.1 Housing and Transit Development

Transit development in the County of Los Angeles was a practice that were used by companies like the Pacific Electric Railway Company to increase the value of land surrounding the rail lines. Till this day this practice such as transit oriented development, and zone policies are used to create ridership and increased land value. The condominium market in San Diego was used to analyze how affect transit-oriented development increased housing prices, how proximity to rail stations in Texas have influenced single family residence (SFR), and how different land types value fluctuates in proximity to a rail station.

In San Diego the transit-oriented development near condominiums has increased the price value of these homes. Michael Duncan researchers attempted to understand what variables were the factors that increased the value of these condominiums (Duncan 2010, 3). In the study the hedonic price model was utilized to estimate the statistical significance of transit-oriented development. According to the results a “condo in a good pedestrian environment and near a station (i.e. TOD) has a significantly higher value than a condo in a similar neighborhood not

near a station. Conversely, a condo in a less walkable residential neighborhood near a park-and-ride station (i.e. TAD) can have values that actually fell below" (Duncan 2010, 121). There was a strong connection between the dependent and independent variables illustrates that amenities added by transit orient development (TOD) increased the price of a condominium. By taking into consideration the walkability as variable the author argued that TOD did have the ability to alter to prices of housing. Suggesting that the further away the condominiums are from the station and limited walkable areas the lower the prices of the unit. The hedonic model was utilized in this study to show what factors influenced the price of housing. The next study takes a different approach and used single family residence (SFR) and not condominiums as this study did.

The METRORail in Texas is the 12 largest public light rail system and this study examined the influence it had on single-family residential home value. Multiple linear regression (MLR) and ordinary least squares (OLS) models were used to determine which statistical tool would be able to show to most statistical significance. Both regressions illustrated different results arguing that residential homes near stations have higher value than those further away. The conclusion states that "the OLS estimation reveals that the light rail line has positively significant effects on properties located between one-mile and three-mile distance from rail stops while the MLR model only shows insignificant effects of light rail on those properties" (Diao 2019, 166-167). There was a statistically significant reaffirming the prediction that a station near residential property was one major contributing factor to increasing resident house value. The correlation between distance and the property value of the single family residence (SFR) provided further evidence that proximity to a rail or station equals greater price value. Two different methods were used to analyze the correlation to understand the patterns that can occur

with spatial distribution. The study showed that OLS was more effective than MLR to determine which independent variables, in this case distance, to show the influence on housing values.

The next study examined the value of condominium by applying a different approach and examined the sale of properties utilizing parcel data. According to the findings of the study "among land uses, the largest share of condominium sales was within a half-mile of a transit stop – 16.9 percent; for other land uses, the shares of sales near rail stations or BRT stops were: commercial – 12.6 percent; multi-family housing – 12 percent; and single-family homes – 5.7 percent" (Cervero et al. 2002, 11). Robert Cervero utilized the hedonic model to analyze the parcel data and the different types of land use – commercial, single-family housing, multi-family housing, and condominium. This analysis showed that land types are different overall within the proximity to the light rails the values goes up almost 20 percent and the lowest land value 6 percent. The authors took into account different types of zoned land which showed that different type of housing has various values – a condominium closer to a station did have increased land values than single family residence. The analysis showed that the land zoned as condominiums did have more value than SFR suggesting that type of housing was as important as proximity. The type of housing can be an important indicator as much as a light rail, suggesting that there are many external factors other than that light rail that need to be considered.

These studies have shown that depending on the independent variable with a proximity to the light rail could play a greater role than other variables such as square feet or numbers of bedrooms. In the three studies above whether the HRM, OLS, or MLR are used it suggest that proximity between 1 to 3 miles to a rail station gave the greatest value to housing. The different regression showed that HRM or OLS have the capacity to show statistical significance while MLR was not as responsive. These two regressions illustrated the importance to the proximity,

while the last reading suggest that housing type plays more of a role than proximity, but it is still important factor. The hedonic regression in the last article conveyed that the type of housing can increase the value property when it sold. Suggestion that a combination of housing type and proximity have the greatest impact on housing prices than just proximity.

2.2 GIS Approach

In this section there are three different articles that were used to break down the role of GIS spatial analysis. These research paper inspected the effectiveness of GIS and the limitations that occur. Some of the readings are heavily based on understanding transportation with regressions and others take a more holistic approach to consider other forms of analysis with automated models or used other statistical tools. Overall, these articles are a mixture of different literature in different locations to illustrate how influential GIS was when it comes to studying urban planning, transportation, and housing.

Irregularities that affect housing near light rail in Portland were inspected to see the influences of light rails. These irregularities are negative impacts or nuisances that make the property values decrease with time and space. According to the authors, "the results of this study confirm our hypothesis that the light rail has both a positive effect (accessibility effect) and a negative effect (nuisance effect) on single-family home values. The positive effect dominates the negative effect, which implies a declining price gradient as one moves away from LRT stations for several hundred meters" (Chen et al. 1997, 11). The hedonic regression model was used to understand the negative and positive effects of the assets in the community and by adding a light rail the value goes up and adding a heavy rail the value goes down. The negative impacts or nuisances can also have an impact and lower the value of housing, this analysis showed that multiple influences can affect the value of single family residence.

Geo-statistical analysis was used as a tool to understand transportation cost. A cost analysis tool was created to have a quick and accessible manner to process data to determine cost and accessibility by the light rail. In order to have an effective tool the authors suggested that an automotive model through ModelBuilder should be "generic and transferable, able to accept standard GIS input in the form of spatial data, at any given scale, and perform the generalized cost computation to produce a set of accessibility measures, an add-in" (Ford et al. 2015, 128). Once this tool was created it showed that accessibility was imported to commuter traveling on public transportation. The automated model illustrated that light rail does have an influence on accessibility and shows that light rail allowed a greater connective network. Automating the tool to perform an analysis gave the user the ability to study a variety of scenarios. The ModelBuilder showed that walkability and access to the light rails are important for users. To understand the full impact of accessibility it would be important to consider the condominiums or apartments.

The article examined ridership of Los Angeles Metro's light rails by analyzing aggregated data such as the census to understand these spatial patterns. The author concluded that Metro lost ridership between 2009 and 2016 "ridership of about 26% along the extension corridor of Gold Line over the seven years. System-wide, overall Metro ridership in LA County decreased 10.5% and bus ridership decreased 18.6% for the same time period" (Lee et al. 2013, 13). The data provided illustrated that overall there was a decrease in ridership, which can question if these light rails are effective or a service to the communities. There was a significant decrease of ridership in specific communities, even though this study does not discuss the values of housing it does present important questions. If ridership is falling who is uses the light rail? Is it worth investing in? And why has light rail transit falling in certain neighborhoods in Los Angeles?

2.3 Displacement

This section defined what displacement actually means and what was considered affordable. Being able to define these two words clearly were important to determine the standards. To understand what displacement means and what occurred when displacement occurs these following studies inspected the term and specific locations. Displacement was defined as the force movement of people by external forces. While, gentrification was defined as improving of houses in the communities to appeal to outside communities. Gentrification in this thesis was not used because it does not consider housing improvements as an influence, but only external factors.

To analyze data, it was important to understand what approach best suits the needs of that data in, the author gave an in-depth insight on what approach best fits for this data. This research was conducted by running several regression models to understand the patterns of gentrification near light rail. In the paper the authors state that over 50 years the term gentrification has be utilized and defined differently due to the complexity of the word it has evolved to gentrification being defined as "neighborhood-level changes in the income distribution of residents" (Boarnet et al. 2018, 3). For that reason, gentrification was excluded because term only focused on one factor and that was income. This report allowed the reader to understand what the data represented. They break down 18 years of data to analyze the patterns of gentrification and what approaches were used in this analysis such as a statistical regression.

In another article Marlon Boarnet established what criteria was needed to determine displacement stating the indicators such as affordability, hazards, or environmental influences can cause displacement. The authors state that displacement was defined as "a household is forced to move from its residence because of conditions outside of its control, which occur

despite meeting pre-imposed conditions of occupancy, and which make occupancy "impossible, hazardous, or unaffordable" (Boarnet et al. 2017, 3). This definition allowed the authors to create an understanding of what displacement actually means. The authors guided the reader to understand how historic neighborhoods were created and what it means when the community members are displaced. Unaffordability was one factor considered to determine if East and South Los Angeles are currently going through the process of displaced.

The last article of this section defined affordable housing and the concept of how much income should be spent on rent or housing. The Mary Schwartz begins by stating that about 30% of one's income should go to rent so it can be affordable for the individual to sustain a life style. According to the article "the conventional public policy indicator of housing affordability in the United States is the percent of income spent on housing. Housing expenditures that exceed 30 percent of household income have historically been viewed as an indicator of a housing affordability problem" (Schwartz 2006, 1). This means that when individuals are spending more than 30% of their income on housing it becomes a financial burden and therefore, becomes unaffordable. The authors' utilization of census data was another aspect that contributed to the understanding on how to determine unaffordability. This clearly help understand how much income should be spent on housing. When housing is unaffordable it makes it harder for individuals to go from renter to homeowners and that external influence can cause displacement of these communities.

The articles present here vary greatly where some of the readings emphasize mostly on housing and transportation, and others on GIS, defining displacement and affordability. The variation of these articles depict the complexity of this topic. By understanding the interconnection between these fields it guided the research to be holistically in its approach. The

most common research tool utilized in these articles was the hedonic regression model. Which was the approach that was taken to determine whether there was a correlation between Metro's light rails and housing prices.

Chapter 3: Methodology

The objective of this section was to describe the data, present the scope of the project, and to present how this data was analyzed. The Hedonic Regression Model explained and examined what variables were used to understand if there was a correlation or some type of relationship between these two entities. The second part of this section layout how demographic data was visualized for the five year period of demographics, income, and physical landscape of the East Los Angeles and South Los Angeles area.

3.1 Data Sources

Most of the data that was used comes from local and government sources - parcel county and census data. The Assessor Parcel Data was an assessment that the local government conducted of land data to determine the value of real land and property. The houses are represented by parcel, which are the boundaries of the land that was owned by residents. The purpose of this data was for local government to assess the value of the properties within the boundary of the County of Los Angeles — including all types of land usage commercial, residential, or vacant. The local government used this information to determine future property taxes dependent on the value of the property. The Assessor Parcels Data - 2006 thru 2019 was the largest dataset used having 30 million rows and 51 columns. Located on the County of Los Angeles Open Data website, which the assessor parcel data was collected 2006 and was updated every year currently up to 2019. The metadata was updated a couple times a year – the website showed it was updated in July and October of this year.

The next datasets consisted of demographic data which focused on ethnicity, and income data that span from 2009 to 2017 a five year period. The first dataset that came from the American Community Survey found on the Census Bureau website. The datasets were used for

this study came from Census Bureau website which is an ongoing survey that informs local government of the changes that go on like demographics. The American Community Survey provided more information regarding jobs, education, whether homes are owned or rented. The ACS also determined how billions of federal and state funds are spent each year. This data looked at the block group and provided detail information that gave a greater understanding of each neighborhood and not just the overall scope that a national dataset would give.

The last dataset came from Los Angeles Metro's website. This source kept track of all the rails and bus locations in the County of Los Angeles. These buses and light rail lines are updated twice a year in June and December. The GIS data consisted of the location of the stations and the lines that the rails go through, estimating the distances between the locations of each station and the closest SFR. The data that was collected and analyzed all came from public institutions that provided the information on their website. Having the accessibility to this information showed that it is possible to for these public intuitions to be involved in the process of studying this spatial phenomenon. There is the capacity and ability to have easy access to research this topic.

Table 3: Table of the needed data to conduct the study.

Dataset	Temporal Scale	Spatial Scale	Data Format
Metro's Light Rail Line	Updated December 2018 - Dataset updated every 6 months	Los Angeles County	Shapefile - Origin of dataset or source: Metro Developer GIS data (Free)
Assessor Parcels Data - 2006 thru 2019	Updated October 2019 - 2006 to 2019	Los Angeles County	CSV File - County of Los Angeles Open Data website (Free)
American Community Survey	Updated November 2017 – 2009 to 2017	East and South Los Angeles Neighborhoods	CSV File - Census Bureau, City or County website (Free)
Zoned Land	Updated October 2019 - 2006 to 2019	Los Angeles County	CSV File - County of Los Angeles Open Data website (Free)

3.2 Data Description

This chapter described the data, cleaning methods, research design, and the methods used to analyze the data. In order to achieve the best outcome all, the data was altered in order to fit the needs of the spatial and temporal parameters. The most complicated dataset was the Assessor Parcels Data - 2006 thru 2019 that contained all the housing and zoning data. The ACS data consisted of all the demographic data, which was altered to normalize the data. Lastly, the Metro data was shortened to the South Los Angeles and East Los Angeles neighborhoods.

Once the data was cleaned and sorted the methods that were used are cartography and spatial analysis techniques. The demographic data was represented through a series of maps to illustrates if there was a change between a nine year span from 2009 to 2017. The next maps illustrate the zoning data, which came from the assessor parcel dataset and these maps span from 2009 to 2017 as well. The dataset was analyzed through a regression, which attempted to determine a correlation between housing prices and the Metro Expo and Gold line in these neighborhoods.

3.2.1 Assessor Parcels Data - 2006 thru 2019

This dataset contained an enormous amount of data regarding housing information for the County of Los Angeles. The information that was displaced includes zip code, assessor's identification number, year built, square feet which was just a few to name – a dataset that has 45 columns and 33,348,628 rows to be exact. The issue with this amount of data was that Excel had the inability to hold all the data within one spreadsheet. To have all the information needed for the regression the data was downloaded in several spreadsheets sorted by year from 2008 to 2018. Each of these spreadsheets hold the same information but with different years, location, and the difference between the data from year to year. Once these spreadsheets were created a

database which was centralized in one location and for the purposes of running the regression. This made it easier to access the information in the most effective manner to run the regression.

The data was filtered so it only focused on the geographic areas of South Los Angeles and East Los Angeles excluding housing outside those boundaries. This dataset did not fully download as one CSV file, so it was broken down by year (2008 to 2018) and each neighborhood to have individual dataset. The Assessor Parcels Data - 2006 thru 2019 consisted of land value, land improvement, year built, year of data recorded, square feet, and geographic location (etc.). When the network analysis was created it was filtered to five thousand SFR locations in order to be most effective to run the network analysis.

3.2.2 Block Group Data

The ACS provided two datasets the main component that was used to analyze the income and ethnicity of community members of South Los Angeles and East Los Angeles. The data consisted of the six different neighborhoods that the Expo line passes through, while the Gold line only pass through two neighborhoods in East Los Angeles. The ACS also provided information like the occupancy, education level, unemployment levels, and even transportation pattern of the block group. In order to have the most accurate representation of the data it was altered by normalizing the information were the total population was divided by the population of the ethnic group for it to be statistically relevant. The normalization provided the most accurate representation of the geographic areas that were studied, but unlike ethnicity the income was already normalized because the median income was used and not just income.

3.3 Research Design

This section layout the steps and processes that are taken to determine if the working hypothesis was accurate: to determine whether Metro's light rail proximity has a relationship with increasing housing prices which changed the demographics, and structural layout within .5 miles of these light rails. The flowchart below illustrates the steps taken throughout the analysis to create outputs depicting the outcome of the data. As mentioned in the beginning of this thesis there were two outputs at the conclusion of this analysis – a series of maps and the statistical analysis of the data.

The first step for the analysis sets the parameters of the data this includes uploading all the Metro light rail lines to create one file with all the lines located together. A .5 mile buffer was added which was used to analyze the data within that buffer only housing within that buffer was considered. Then the parcel data as mentioned before was downloaded and filtered with Excel and Access to create one database that contains all those files together. By creating this database, it allowed the software program ArcGIS pro, Microsoft Excel, and R Studio to run the analysis more efficiently without any technical issues- crashing software and loss of data.

Then this database was uploaded onto ArcGIS Pro where the coordinates were used to create an XY table that was then geocode the locations of each single family residence as points. These points were then used to create a series of maps and run statistical analysis along with the ACS data to illustrate change in demographics and structural change in zoning practices as well.

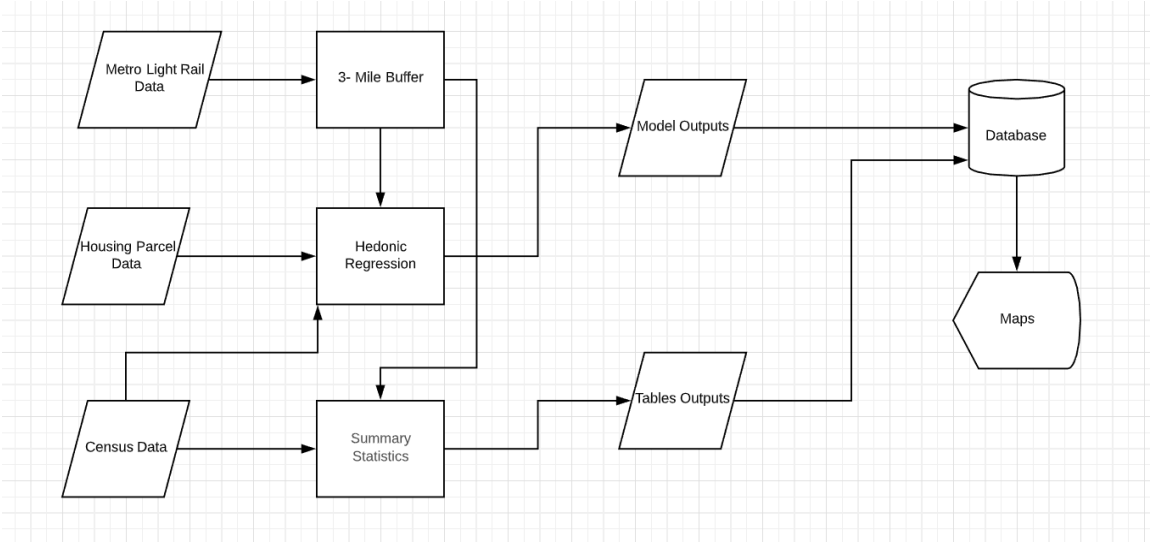


Figure 2: Flowchart explaining the process to conduct the analysis and outcomes.

3.4 Analysis Techniques and Tools

3.4.1 Hedonic Regression Model

The Hedonic Regression Model is an economic theory and statistical tool utilized to determine the impact that independent variables have on dependent variables. In this study the HRM was used to determine if there was a correlation between housing prices and the metro light rails stations near these homes. The regression was used to examine how much of a potential relationship there was between the dependent variable (housing prices) and independent variable (Metro's light rail and housing characteristics). Based on these two entities the HRM determined how much of an influence the independent variables had on the dependent variable.

The formula below examined how the attributes, parameters, and random variables affect housing prices and the correlation between the dependent and independent variables. The assets attribute (**X***i*) was the housing prices which was represented by total value in the dataset – the housing structure and land value. The model parameters (**α***i*) that were utilized for this regression were year built, square feet, bedrooms, bathrooms, land value, and miles to the light rail. Random variables (**ε***i*) were not taken into consideration like noise, parks, or air pollution.

$$\mathbf{P}i = f\,(\mathbf{X}i, \alpha i, \varepsilon i)$$

Where:

X*i*: asset attributes

α*i*: model parameters

ε*i*: random variables

However, there are many manners in which this regression was manipulated to give a different outcome depend on that types of influences that were used. But like with any theory this regression model has limitations and in order to combat this limitation another step needed was taken so there were no over estimation in the model or bias. The Hedonic index suggest that to determine what fits or parameters were adjusted some studies used OLS regression with HRM to reduce the bias in the model. These models were adjusted for the HRM to reduce the bias to limit amounts of independent variables that were used, since random variable were not utilized there was not a concern to reduce the amount of basis for both models.

3.4.2 Network Analysis - Closest Facility Analysis

The closest facility analysis is one tool in the network analysis toolset which analysis network datasets — lines representing the flow of routes. The Assessor Parcel data contained a large amount of data regarding housing, but it did not have data regarding housing distances to each station. To provide that data a network analysis was created in order to join the original data with the distance that was created from each station to SFR. The network analysis that was used to add distance was the closest facility. The assessor parcel data had the coordinates of each house, which was geocoded and uploaded into the network analysis. Metro's light rail stations also contained the coordinates of the location of each stop and these stops were uploaded to the network analysis. Once that was completed the light rail stations were used as the facilities and the housing location were used as incidents. This tool added the distance by settings the parameters to miles and miles walked which tracked the distance from these two locations. Due to the limitation of this tool only housing at .5 miles of each station were considered which was about five thousand SFR housing locations. The network analysis had a limitation to the amount of information allowed to be analyzed.

3.4.3 R Studio

To process the thousands of rows between the light rail and housing, it was processed through R Studio which ran a regression to determine if there was a correlation between these two entities. The regression was analyzed with lm (linear model) functions to determine if the points were random or showed statistical significance. The function used the dependent variable as the response and the independent variables as the terms which allows the function to run the linear predictor for the outcome. The outcome of this function determined which independent variable has a stronger influence and if the model had any statistical significance. The lm function allows for an argument that was made to determine if the parameters had an influence over each other.

3.5 Map Creation

3.5.1 Demographics Maps

The maps created represent nine years of block group data from the ACS, which illustrate income and ethnicity for each community. For income and ethnicity, a series of maps (total of 10 maps) showed any changes from 2009 to 2017 time period. Temporal time frame on these maps were established between 2009 and 2017, because the ACS only provided a nine year period worth of data. The maps were represented by a choropleth map to illustrate the findings and outcomes of the quantitative data.

3.4.2 Zoning Maps

The zoning data came from the Assessor Parcel dataset, which contained how each parcel were zoned. For these maps a time span of nine years were used from 2009 to 2017 to illustrate whether there was a change in the type of zoning in the areas around these Metro lines. In order to depict that image choropleth maps were created to illustrate the zoned area during this time

frame. These maps are straight forward but are utilized to depict a story that useful to understand changing practices and demographics.

Chapter 4 Results

This section examined the results of the data to determine if the outcomes supported the hypothesis. The HMR determined the correlation between the dependent variable and independent variables were weak and did not find a strong relationship. The dependent variable and independent variables that were utilized to analyze both South and East Los Angeles housing prices resulted with a stronger relation between the Gold line and Single family residence. The model run for the Gold and Expo line showed a greater significance. The maps for income, ethnicity, and zone types utilized determined that there was little to no change within the 0.5-mile proximity of each light rail station. These approaches showed that Metro's Expo and Gold lines have had a small impact on these communities that cannot necessary show a relationship that caused displacement of members from these communities

4.1 Hedonic Regression Model

In chapter 3, it was established that HRM determined how much of an influence the independent variables had on the dependent variable. When each of the independent variables were examined before running the linear regression much of the data showed that they tended to skew right when histograms were created. These right skewed distributions suggested that the independent variables are around the median of the data with some variables that were above the mean and no variables below the mean. Which suggests that the housing characteristics do not vary dramatically. The right skewness appeared for both East and South Los Angeles housing characteristics show that the data does not depart from the median value. The variable "Total Value" that was created showed that most of the housing land price were the same suggesting that most of the data did not deviate too much to create outliers.

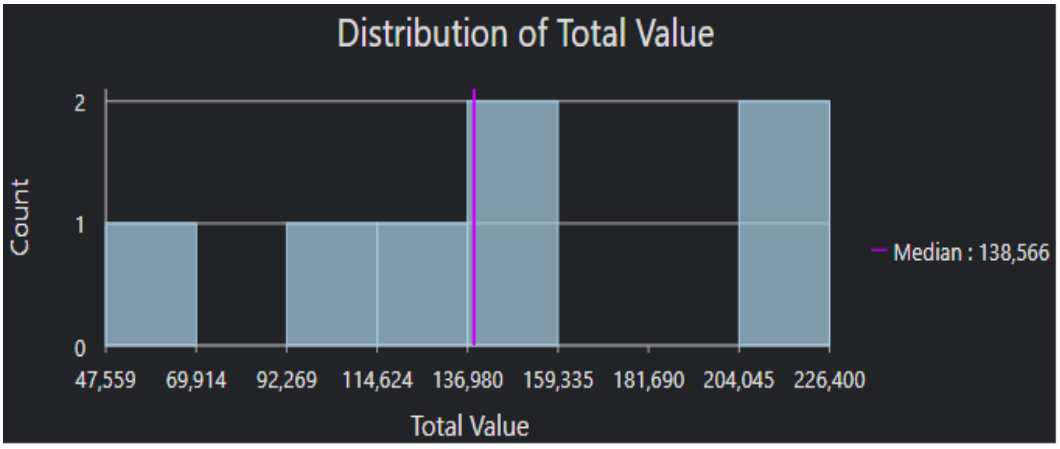


Figure 3: Housing value near the Gold line in 2009.

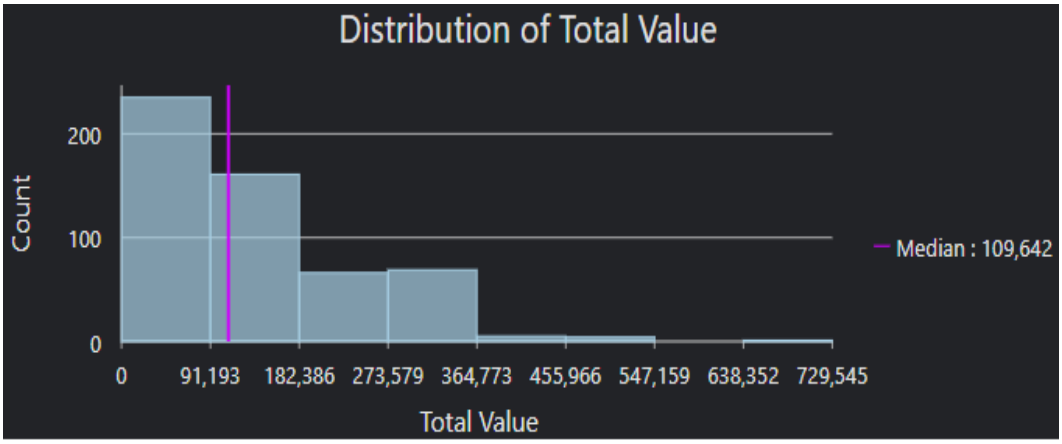


Figure 4: Housing value near the Expo line in 2009.

4.1.1 South Los Angeles

To analyze housing values in South Los Angeles total value of single-family residential property was utilized as the dependent variable. While, year built, square feet, bedrooms, bathrooms, land value, and miles to the light rail stations were utilized as the independent variables. Once the HRM was run the results were quite confounding were the Adjusted R-squared resulted at 0.8752 and the Multiple R-squared as 0.8754, suggesting that there was some type of correlation between these entities. However, when looking at the significance level all the independent variables with the exception of miles' distance to the station, suggested the null hypothesis was rejected. The model suggested that the independent variable with the most significance was miles to the stations with the p-value was 0.731875, this variable was significant compared to the other variables which showed little correlation to the values of SFR.

Table 4: Model 1 of South Los Angeles housing and Expo line proximity.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-6.26	1.10	-5.694	1.32 ***
YearBuilt	-1.88	5.03	-3.727	0.00 ***
SQFTmain	3.70	9.03	40.927	< 2.00 ***
Bedrooms	1.19	7.80	15.22	< 2.00 ***
Bathrooms	1.09	8.73	12.515	< 2.00 ***
LandValue	1.22	1.11	110.148	< 2.00 ***
Miles	-5.08	1.48	-0.343	0.73

4.1.2 East Los Angeles

The HRM that was run for the data regarding SFR price in East Los Angeles parameters were kept the same as the parameters for the data of South Los Angeles. However, the findings for this model were more surprising than those of South Los Angeles. The results showed that the model had a stronger Adjusted R-squared at 0.8791 and the Multiple R-squared at 0.8794. This model was stronger and showed that there was more significance with the data. The high p-value suggest that these data is not random. This model showed that the null value should not rejected for the year built, square feet, and bathrooms. While, the suggested that the land value, bedrooms, and miles to the stations should rejected the null because there was a stronger relationship.

Overall, the HRM models for SFR housing in neighborhoods in South and East Los Angeles showed that light rails did have some influences when it comes to housing prices, but overall the significance was more that relevant with certain variables like distance to the light rail, size of housing, and number of rooms. These models debunk the hypothesis, which suggests that the light rail had the ability to change the prices of housing at a significant level that could cause a spike in housing price – leading to displacement. In accordance with the two models the light rails do not have a great impact on the price of housing. However, there is more of a significance and a stronger relationship with the Gold line than with the Expo line. Since the opening of the Expo and Gold lines there was little change in the social and physical landscapes of East and South Los Angeles as shown with the regression, which the series of maps also support those findings as well.

Table 5: Model 2 of East Los Angeles housing and Gold line proximity.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-9.06	1.15	-0.785	0.43228
YearBuilt	-1.21	4.55	-2.668	0.00770 **
SQFTmain	2.57	2.30	111.6	< 2.00 ***
Bedrooms	3.47	1.11	3.124	0.00181 **
Bathrooms	-1.11	1.19	-9.312	< 2.00 ***
LandValue	-1.39	1.61	-8.639	< 2.00 ***
Miles	-4.62	2.27	-0.204	0.8384

4.2 Demographics in South and East Los Angeles

In order to see if there was a significant change in these communities there were two indicators that were used to cartographically illustrate whether there was a change. In this section thematic maps were created and visualized whether was a change in the communities in from 2009 to 2017. In chapter 1 there was an attempt to display the characteristics of these communities with levels of education, ethnicity, and income levels. For the last couple of years these communities were statics with little change in income and ethnicity. Each of these communities have large population of minority groups that live in these areas. Where South Los Angeles has more than 80% of the population was black or African American and East Los Angeles has more than 90% of the population was Hispanic or Latino. The second largest populations that live in East Los Angeles were white individuals and Hispanic or Latinos that live in South Los Angeles.

4.2.1 Ethnicity

Figure 4 illustrate the ethnicity of the black and white community members of South Los Angeles. The black members of these communities were used because they are the largest population in this area and white community members were used because they are the largest growing demographics in these specific neighborhoods of South Los Angeles. 2009, 2013, and 2017 were the years used to determine if there was a significant change over the five years time period in these areas. As seen in 2009 83% of the population was black with the highest concentration were in the Leimert, Baldwin Hills/Crenshaw, and the further east and closer to downtown areas less than 10% had a concentrated of black people. In 2013 and 2017 the maps illustrated a small significance of change that occurred in 2017 with more black families or individuals moving more East.

However, change was seen with the white community member in 2009 only 6% or less of the population lived in the Leimert, Baldwin Hills/Crenshaw area of South Los Angeles. Where 66% of the population lived in the downtown and eastern communities. In 2017 an increase of the overall population of white individuals living in this area did not increase but showed slight decrease with higher concentration moving to the east closer to Downtown area.

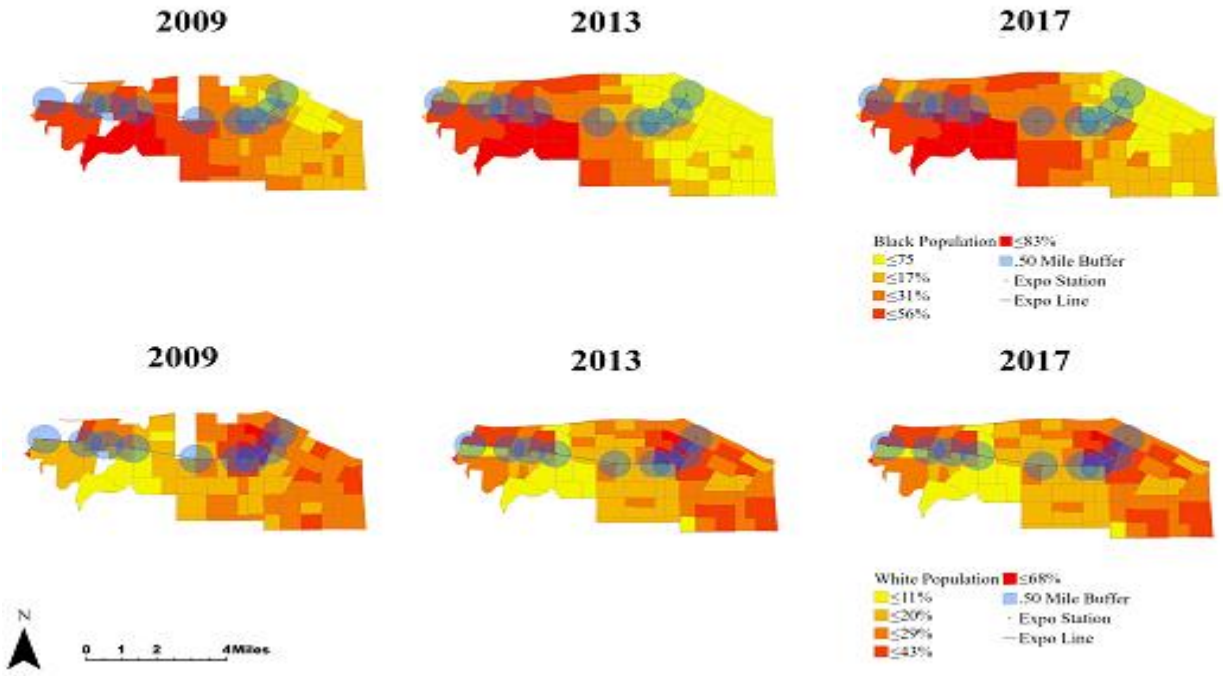


Figure 5: Demographic changes of South Los Angeles from 2009 to 2017.

East Los Angeles largest population was the Hispanic/ Latino community with the population was 96% having one of the highest concentrations of Hispanic/ Latino in Los Angeles. Like South Los Angeles the same phenomenon was seen in East Los Angeles with highest concentration of Hispanic/ Latino living in Boyle Heights and East Los Angeles neighborhoods. The further west and closer to downtown Los Angeles between 28% and 55% of the population was Hispanic/ Latino. From 2010 to 2017 the population of Hispanic/ Latino stays consistent and do not show significant change with the majority of the population still heavily concentrated area of the same ethnic group and income levels.

However, the white population has grown slightly but not to a significant level. In 2010, about 2% or less of the white individuals lived in the Boyle Heights and East Los Angeles neighborhoods this has grown. This change was seen in the 2017 map where pockets of neighborhoods do have a slightly higher population than in 2010. However, 2013 did show the biggest increase with higher numbers of the white population were seen on the map. The same trend as in the maps for South Los Angeles was seen as well where higher populations of white individuals live closer to the Downtown area with 56 % of the population being white in 2010 and in 2017 grown to 65%. Overall, there was a slight increase in the white population East Los Angeles, but not enormous increase compared to the Hispanic/ Latino.

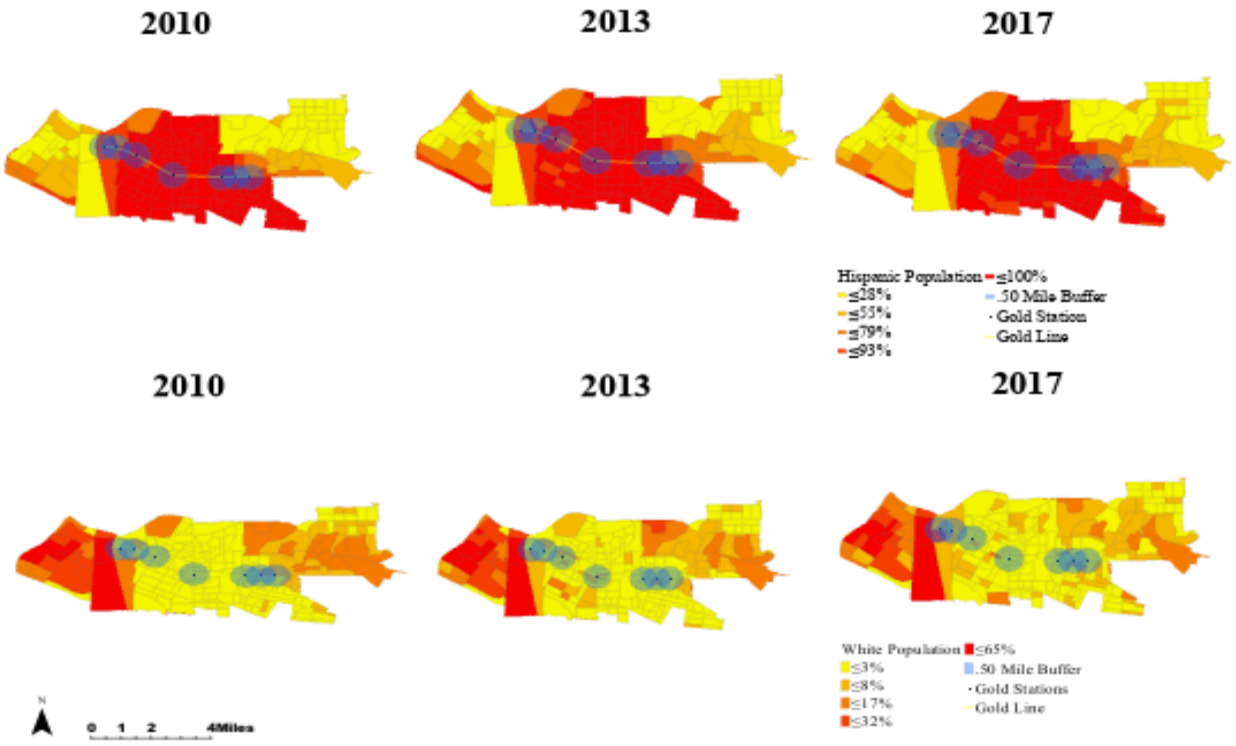


Figure 6: Demographic changes of East Los Angeles from 2010 to 2017.

4.2.2 Median Income

The median income for both East and South Los Angeles varies greatly with the median income on the lowest rage was around $22,000 and on the highest range was around $100,000. In South Los Angeles the median income in 2009 was higher with less of the income was around $22,000. However, a change was seen in 2013 in neighborhoods more to the west and away from the downtown area higher income levels with couple of communities having household that earned up anywhere from $56,000 to $100,000 annual income. In 2017 a change in median income was seen where household began to drop the median income that ranged from $22,000 to $46,000 which dominated. Since 2009 a decrease in median income showing that households are not making the same income as they were in 2009.

The same argument was made from median household income in East Los Angeles. In 2009 the median income ranged from $22, 000 to $100,000. The highest levels of income were located in the downtown area and Monterey Park (the city was included because a station was located in proximity to East Los Angeles). In Boyle Heights and East Los Angeles, the average income was between $40,000 to $52,000 and the lowest median income of $22,000 were concentrated in the Downtown Los Angeles area in 2009. Unlike the phenomenon in South Los Angeles in East Los Angeles incomes increased from 2009 to 2017 were incomes increased in the downtown area but decrease in certain parts of Boyle Heights. Incomes did not change too dramatically in that five-year period.

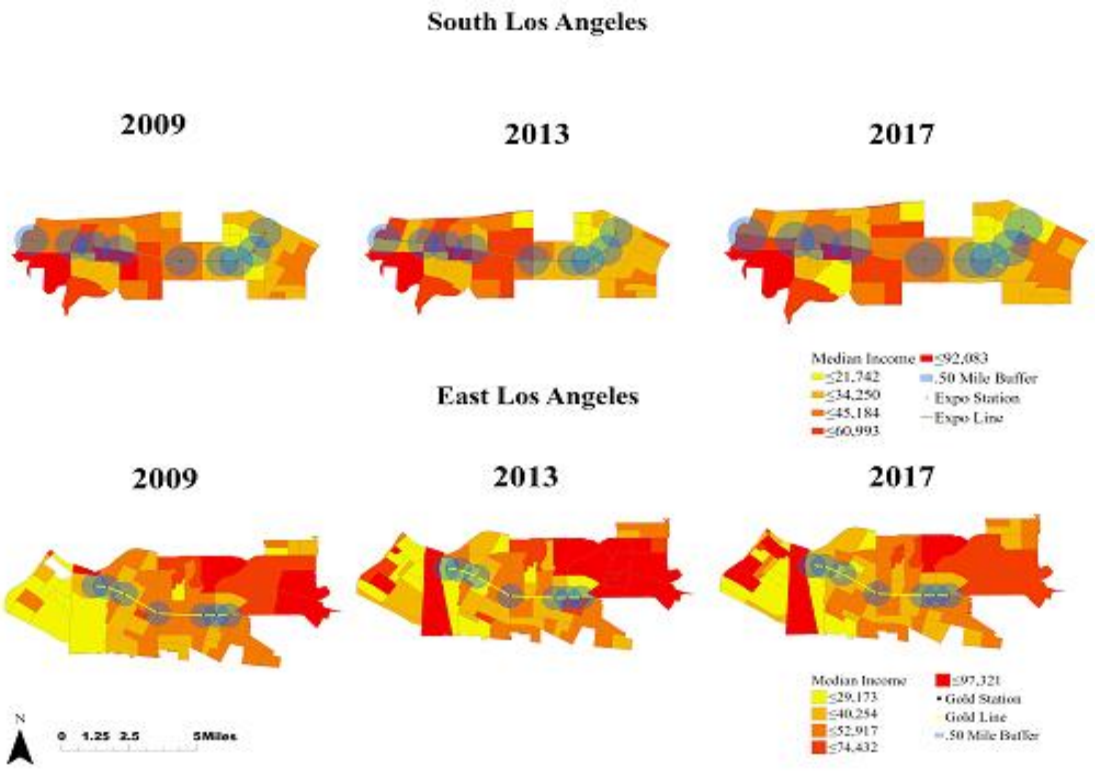


Figure 7: Income changes of South and East Los Angeles from 2009 to 2017.

4.3 Zoning of South and East Los Angeles

The last indicator that was used to determine the influence of the Expo line was zoning data. This section looked at the number of apartments, condominiums, or non-single homes residency created from 2009 to 2017. In 2009 there were a total of 998 buildings that were zoned as apartments, condominiums, or non-single homes residencies in Exposition Park. The neighborhoods of Leimert Park and Baldwin Hills/Crenshaw had a total of 619 buildings zoned as apartments, condominiums, or non-single homes residencies. In 2017 the figures below illustrate that there was some change but not a large amount to different zone types.

This same pattern was seen in the East Los Angeles as well. It showed that there was a slight increase in zoned apartments, condominiums, or non-single homes, but no to an extent where large amount of properties were zoned differently from 2009 to 2017. The maps suggest that has been not an increase number of zoned housing to non single family residence make the argument invalid that the Expo and Gold line have changed the physical structure within that .5-mile proximity. overall, that data does not support that the hypothesis that the two light rails have changed this area dramatically, slight numbers of non single family residence were zoned.

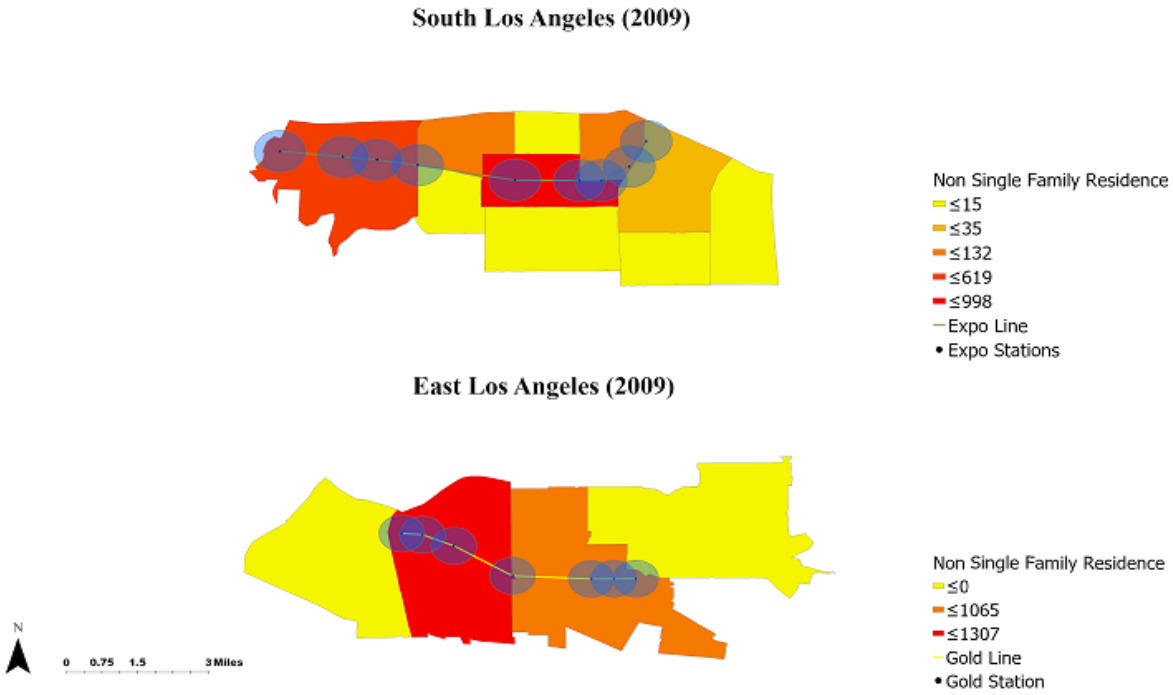


Figure 8: Zone types from 2009 depicts the number of non-single family homes within proximity to the Expo (South Los Angeles) and Gold Lines (East Los Angeles).

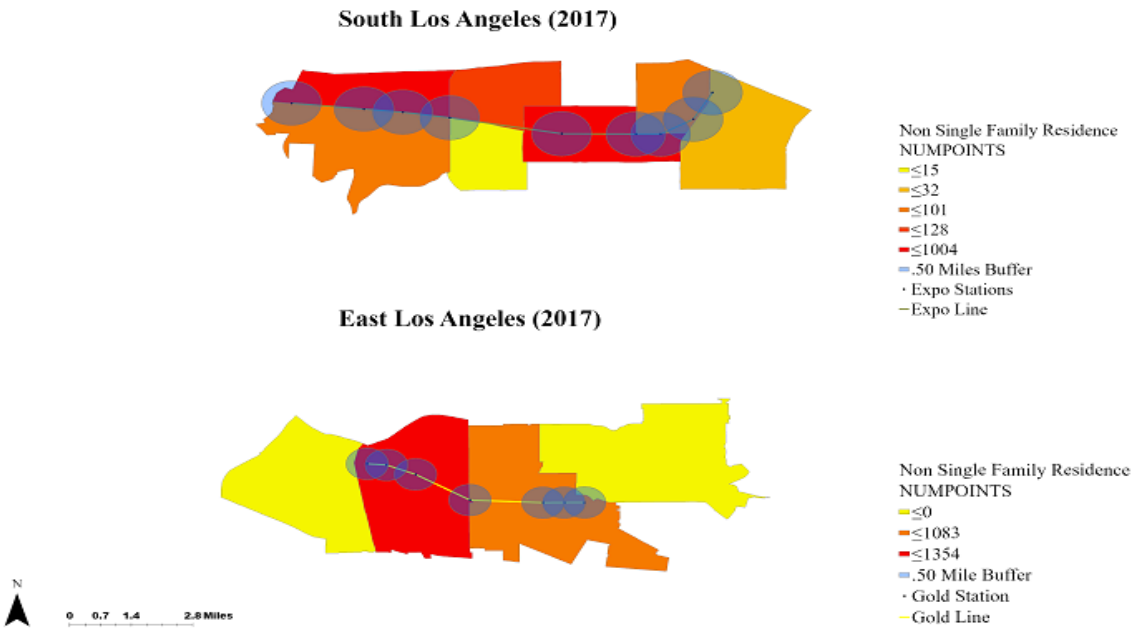


Figure 9: Zone types from 2017 depicts the number of non-single family homes within proximity to the Expo (South Los Angeles) and Gold Lines (East Los Angeles).

Chapter 5 Conclusion and Recommendations

The regressions and maps provided housing, demographic, and physical patterns, even though the results of the regression did not find a strong correlation between the distances of the stations to single family residence. The predictors that were used to run the model for South Los Angeles data yielded a strong relation between the total value of single-family home and the other variables such as room number of room, square feet, and number of bathrooms. However, the second model that was run for East Los Angeles did show that there was a weaker relationship between years built, bedrooms, and a higher relationship mile to the stations and rejected the rest of the predictors. The stronger correlation between the East Los Angeles stations and housing prices can be due to the time period that the stations opened. Compared to the Gold line, opened in 2003, the Expo line has only been open since 2009 which might be a factor as to why there was a stronger correlation with the Gold line and not the Expo line. These models failed to reject the hypothesis but even then, there was some relationship with increased housing prices and the Metro stations for the Gold line. There was not enough evidence to prove that the hypothesis was not corrected.

The second aspect to the hypothesis was to determine whether there were changes in income, ethnicity, or zoning practices near the Expo and Gold lines. These maps showed that within a five year period from 2009 and 2012 there was little to no change in these indicators. There was no substantial change in demographics in the South or East Los Angeles neighborhoods. Most neighborhoods still had extreme concentrations of minority families and individuals living in these areas – South Los Angeles (African Americans) and East Los Angeles (Latinos). The second series of maps did illustrate a change in income, as shown from 2009 to 2017 there was a decrease in median income, but 2013 the maps showed an increase income and

then a drop in 2017. Lastly, zoning patterns did illustrate an increase in apartments, condominiums, or non-single homes residencies. These maps cartographically visualized that there was no change in the 5 year period and that the metro did not have much of an impact that could have changed these indicators.

5.1 Resources Limitation

As mentioned in the earlier chapters the Assessor Data parcel was an enormous amount of data that needed to be processed in order to create meaningful results. Even with the amount of information that was obtained there were challenges regarding the network analysis and how to process this data. ArcGIS Pro limited the ability to study more housing points than just the five thousand housing points that were allowed by the network analysis tool. This limitation did not allow for other forms of housing that could have been analyzed condos, vacant lots, or apartment buildings, but were filtered out in order to meet the constraints. Having the ability to analyze those forms of housing would have made this thesis more holistic and could have changed the outcomes.

Another aspect that hinder that outcomes and provided limitations was the Block Group data (ACS), because the limited five-year period. It was impossible to go into depth because it did not provide a larger temporal span that a ten year period could have provided. The Block Group data on the ACS did not allow data to attach geographically that did not fall into this time period. Due to data prior to 2009 and after 2017 that was not available made it was a challenge when attempting to have continuity with the rest of the data that was obtained.

5.2 Future Work

The work presented in this thesis was just scratching the surface of what was done with this data. Whether it is finding more data that can fit the current time frame or finding other platforms that provide larger amount of data processed, there is still a lot that could be done. This thesis just looked at one variable that could have altered the value of single-family residential housing and it did not take into consideration walkability, TOD, shopping centers etc. There is so much that could be done to improve on this subject and provide more information on housing prices and displacement that could be utilized to understand housing patterns. Which will inform local officials and companies how to reach out the disenfranchised communities with solution that can improve house affordability.

Studying housing patterns is important, because it allows individuals, companies, or public institution to understand the impact that major infrastructural projects can have on communities. Especially in Los Angeles were so many people struggle to find housing that is affordable and with population grow if not understood this problem will only get worse. Studying this phenomenon also allows individuals, and communities to know what type of impact a project like a transit station can have on their community and how they can prepare themselves financially, socially, or environmentally. The more knowledge there is housing patterns, the better it is, because it allows all stakeholders to understand the ramifications that can occur if vulnerable communities are not informed or protected.

References

"About Metro." LA Metro Home. Accessed July 15, 2019. https://www.metro.net/about/about-metro/.

"Assessor Parcels Data - 2006 thru 2019: LAC Open Data." data.lacounty.gov, July 16, 2019. https://data.lacounty.gov/Parcel-/Assessor-Parcels-Data-2006-thru-2019/9trm-uz8i.

Boarnet, Marlon G., Raphael W. Bostic, Evgeniy Burinskiy, Seva Rodnyansky, and Allen Prohofsky. "Gentrification Near Rail Transit Areas: A Micro-Data Analysis of Moves into Los Angeles Metro Rail Station Areas." (2018).

Boarnet, Marlon G., Raphael W. Bostic, Seva Rodnyansky, Raúl Santiago-Bartolomei, Danielle Williams, and Allen Prohofsky. "Sustainability and Displacement: Assessing the Spatial Pattern of Residential Moves Near Rail Transit." (2017).

Cervero, Robert, and Michael Duncan. "Land value impacts of rail transit services in Los Angeles County." *Report prepared for National Association of Realtors Urban Land Institute* (2002).

Chapple, Karen, Anastasia Loukaitou-Sideris, and University of California, Berkeley. *Developing a new methodology for analyzing potential displacement*. California Air Resources Board, Research Division, 2017.

Chen, Hong, Anthony M. Rufolo, and Kenneth Dueker. "Measuring the impact of light rail systems on single family home values: A hedonic approach with GIS application." (1997).

Duncan, Michael. "The Impact of Transit-oriented Development on Housing Prices in San Diego, CA." *Urban Studies* 48, no. 1 (2010): 101-27. doi:10.1177/0042098009359958.

Ford, Alistair, Stuart Barr, Richard Dawson, and Philip James. "Transport Accessibility Analysis Using GIS: Assessing Sustainable Transport in London." *Urban Development for the 21st Century*, 2015, 201-33. doi:10.1201/b18765-11.

"GIS Data." Metro LA Developer, September 26, 2019. https://developer.metro.net/docs/gis-data/.

Glaeser, Edward L., and Joseph Gyourko. *The impact of zoning on housing affordability*. No. w8835. National Bureau of Economic Research, 2002.

Greene, Richard P., and James B. Pick. *Exploring the urban community: A GIS approach*. Prentice Hall, 2012.

Hawthorne, Timothy, John Krygier, and Mei-Po Kwan. "Mapping ambivalence: Exploring the geographies of community change and rails-to-trails development using photo-based Q method and PPGIS." *Geoforum* 39, no. 2 (2008): 1058-1078.

Heravi, Saeed, Alan Heston, and Mick Silver. "Using scanner data to estimate country price parities: A hedonic regression approach." *Review of Income and Wealth* 49, no. 1 (2003): 1-21.

Johnson, Geoffrey. "Developing a Methodology for Assessing the Energy Potential of Residential Micro-Cogeneration Systems." University of California, Berkeley and the University of California, Los Angeles.

Lee, Jeongwoo, Marlon Boarnet, Douglas Houston, Hilary Nixon, and Steven Spears. "Changes in Service and Associated Ridership Impacts near a New Light Rail Transit Line." *Sustainability*9, no. 10 (2017): 1827. doi:10.3390/su9101827.

Los Angeles Times. Los Angeles Times. Accessed October 9, 2019. http://maps.latimes.com/neighborhoods/.

Malczewski, Jacek, and Claus Rinner. "Exploring multicriteria decision strategies in GIS with linguistic quantifiers: A case study of residential quality evaluation." *Journal of Geographical Systems* 7, no. 2 (2005): 249-268.

McConnell, Eileen Diaz. "Who has housing affordability problems? Disparities in housing cost burden by race, nativity, and legal status in Los Angeles." *Race and social problems* 5, no. 3 (2013): 173-190.

Melser, Daniel. “The Hedonic Regression Time-Dummy Method and the Monotonicity Axioms.” *Journal of Business & Economic Statistics* 23, no. 4 (2005): 485–92.

Pagliara, Francesca, and Enrica Papa. "Urban rail systems investments: an analysis of the impacts on property values and residents' location." *Journal of Transport Geography* 19, no. 2 (2011): 200-211.

Pan, Qisheng. "The impacts of light rail on residential property values in a non-zoning city: A new test on the Houston METRORail transit line." Journal of Transport and Land Use 12, no. 1 (2019).

Oladunni, Timothy, Sharad Sharma, and Raymond Tiwang. “A Spatio - Temporal Hedonic House Regression Model.” *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2017.

Rapach, David E., and Jack K. Strauss. "Differences in housing price forecastability across US states." *International Journal of Forecasting* 25, no. 2 (2009): 351-372.

Richmond, Jonathan. *Transport of delight: the mythical conception of rail transit in Los Angeles*. The University of Akron Press, 2005.

Riggs, William, and Forrest Chamberlain. "The TOD and smart growth implications of the LA adaptive reuse ordinance." *Sustainable cities and society* 38 (2018): 594-606.

Schwartz, Mary, and Ellen Wilson. "Who can afford to live in a home?: A look at data from the 2006 American Community Survey." *US Census Bureau* (2008).

Sung, Seyoung. "Transit oriented development and neighborhood change along the light rail system: the social equity impact of the Metro Blue line in Los Angeles." PhD diss., 2014.

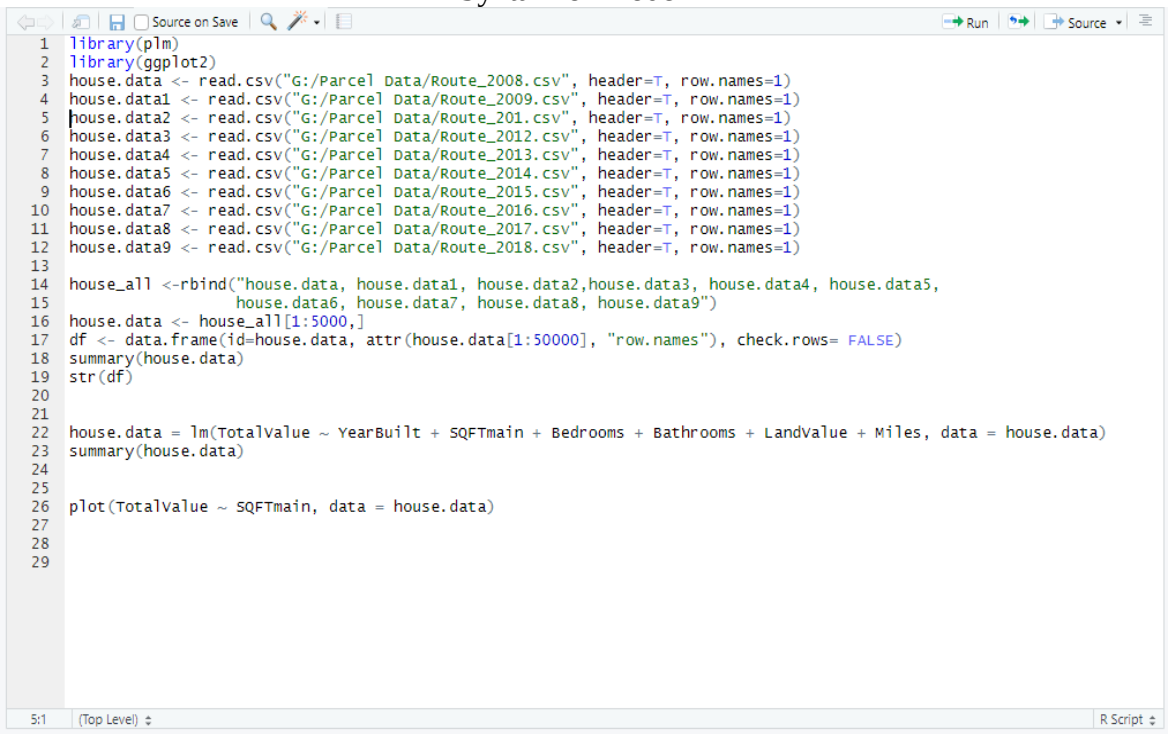
US Census Bureau. “American Community Survey (ACS).” The United States Census Bureau, November 5, 2019. https://www.census.gov/programs-surveys/acs/.

Wei, Yehua Dennis, and Reid Ewing. "Urban expansion, sprawl and inequality." (2018).

Zuk, Miriam, Ariel H. Bierbaum, Karen Chapple, Karolina Gorska, and Anastasia Loukaitou-Sideris. “Gentrification, Displacement, and the Role of Public Investment.” *Journal of Planning Literature* 33, no. 1 (February 2018): 31– 44.

Appendix A

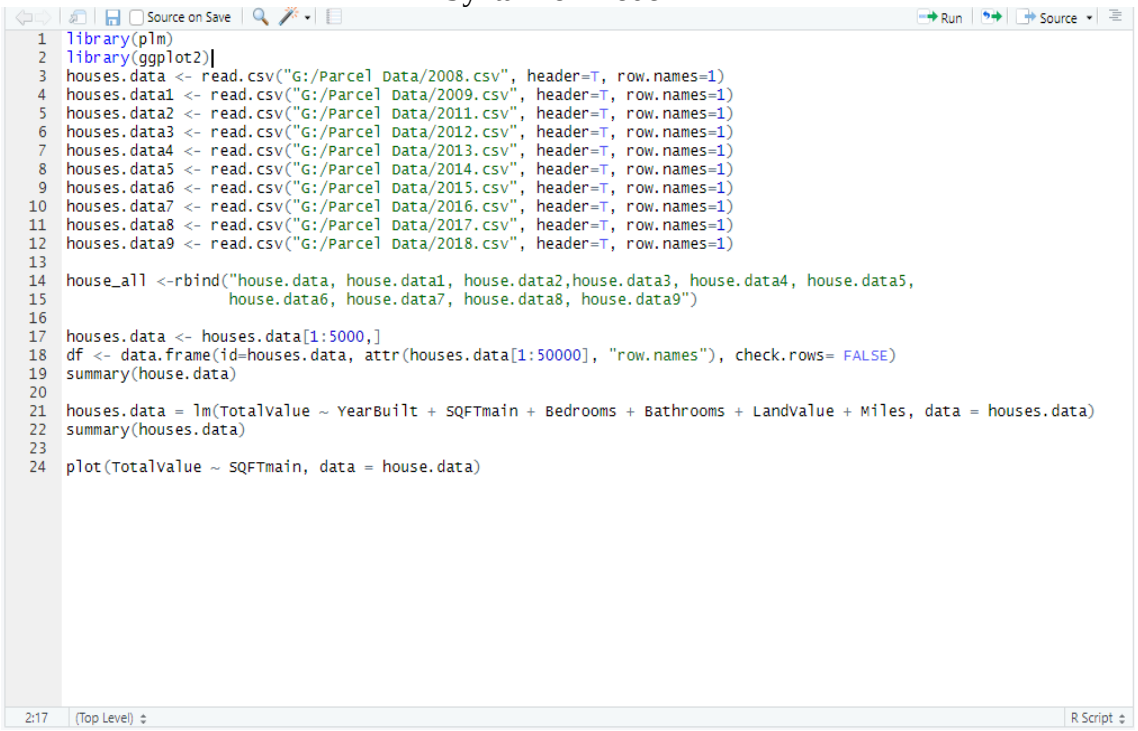
Syntax for Model 1



```
library(plm)
library(ggplot2)
house.data <- read.csv("G:/Parcel Data/Route_2008.csv", header=T, row.names=1)
house.data1 <- read.csv("G:/Parcel Data/Route_2009.csv", header=T, row.names=1)
house.data2 <- read.csv("G:/Parcel Data/Route_201.csv", header=T, row.names=1)
house.data3 <- read.csv("G:/Parcel Data/Route_2012.csv", header=T, row.names=1)
house.data4 <- read.csv("G:/Parcel Data/Route_2013.csv", header=T, row.names=1)
house.data5 <- read.csv("G:/Parcel Data/Route_2014.csv", header=T, row.names=1)
house.data6 <- read.csv("G:/Parcel Data/Route_2015.csv", header=T, row.names=1)
house.data7 <- read.csv("G:/Parcel Data/Route_2016.csv", header=T, row.names=1)
house.data8 <- read.csv("G:/Parcel Data/Route_2017.csv", header=T, row.names=1)
house.data9 <- read.csv("G:/Parcel Data/Route_2018.csv", header=T, row.names=1)

house_all <-rbind("house.data, house.data1, house.data2,house.data3, house.data4, house.data5,
                  house.data6, house.data7, house.data8, house.data9")
house.data <- house_all[1:5000,]
df <- data.frame(id=house.data, attr(house.data[1:50000], "row.names"), check.rows= FALSE)
summary(house.data)
str(df)


house.data = lm(TotalValue ~ YearBuilt + SQFTmain + Bedrooms + Bathrooms + LandValue + Miles, data = house.data)
summary(house.data)


plot(TotalValue ~ SQFTmain, data = house.data)
```

Appendix B

Syntax for Model 2



```
library(plm)
library(ggplot2)
houses.data <- read.csv("G:/Parcel Data/2008.csv", header=T, row.names=1)
houses.data1 <- read.csv("G:/Parcel Data/2009.csv", header=T, row.names=1)
houses.data2 <- read.csv("G:/Parcel Data/2011.csv", header=T, row.names=1)
houses.data3 <- read.csv("G:/Parcel Data/2012.csv", header=T, row.names=1)
houses.data4 <- read.csv("G:/Parcel Data/2013.csv", header=T, row.names=1)
houses.data5 <- read.csv("G:/Parcel Data/2014.csv", header=T, row.names=1)
houses.data6 <- read.csv("G:/Parcel Data/2015.csv", header=T, row.names=1)
houses.data7 <- read.csv("G:/Parcel Data/2016.csv", header=T, row.names=1)
houses.data8 <- read.csv("G:/Parcel Data/2017.csv", header=T, row.names=1)
houses.data9 <- read.csv("G:/Parcel Data/2018.csv", header=T, row.names=1)

house_all <-rbind("house.data, house.data1, house.data2,house.data3, house.data4, house.data5,
                  house.data6, house.data7, house.data8, house.data9")

houses.data <- houses.data[1:5000,]
df <- data.frame(id=houses.data, attr(houses.data[1:50000], "row.names"), check.rows= FALSE)
summary(house.data)

houses.data = lm(TotalValue ~ YearBuilt + SQFTmain + Bedrooms + Bathrooms + LandValue + Miles, data = houses.data)
summary(houses.data)

plot(TotalValue ~ SQFTmain, data = house.data)
```