1. Introduction

1.1 Background

As of the year 2020, the United States housing market is worth an estimated 36.2 trillion dollars. The underlying factors that go into this astronomical amount can be debated, but when it comes to the purchasing of a home, it is the individual that must determine a house’s value and decide what metrics that value is based on. As the old allegory ‘location, location, location’ would indicate, neighborhood characteristics play a major role in the property values of the homes within them. Landing a dream house in the middle of the desert does not leave many options for dinner, regardless of the price paid. Fundamentally, a house is an asset, and to properly value it you cannot overlook its’ location. Thus understanding the characteristics of the area that surrounds it is of the utmost importance.

1.2 Problem

Chicago Illinois is one of the most diverse and economically developed cities in the US. The problem this project will focus on is determining the locations in Chicago that a new home buyer will have the highest likelihood of retaining property value based on the characteristics of the surrounding neighborhood.

1.3 Interest

This information would be of interest to new Chicago home buyers, Chicago real estate investment firms, Chicago municipal departments focused in community development and small business owners looking to locate themselves near higher valued neighborhoods.

2. Data

2.1 Data Sources

Chicago is divided into 77 Community Areas. To understand the population characteristics, general information can be pulled from Wikipedia using python’s beautifulsoup library and looping through the individual Community Area pages. I will also be pulling data from HousingStudies.org, which utilized the census to build easily accessible tables which can be grouped by Community area. I will also use a complex dataset sourced from the Environmental Data Initiative, which contains a study that analyzed tree canopy in 37 major US cities(including Chicago). Obviously, to understand the surrounding venue characteristics and to complete the picture of the neighborhood, I will be using Foursquare.