

# Houston, TX Population`

---

## 1 Introduction

### 1.1 Background

Houston is the 4<sup>th</sup> most populous city in the US with over 2 million within the city limits and an additional 3 million in the suburbs surrounding Houston.

### 1.2 Purpose

The purpose of this project was to evaluate the population data within Houston city limits. The goal of the analysis was to identify the most populous areas and the areas that are growing the fastest.

## 2 Data Acquisition and Cleaning

### 2.1 Data Sources

The population data was collected and cleaned from the US census website broken down by the zip code of residence. The zip code locations were then identified with coordinate data. The “json” file was imported from an open source website

### 2.2 Data Cleaning

Fortunately the data was pretty clean from most of the sources. The one exception was the location data had all position numbers while the folium library needed “negative” data to be able to reference Houston.

## 3 Data Analysis

### 3.1 Visualization

The first step I completed once the data was cleaned was a means to visualize the data to see if there was anything that stood out. The first plot is the current population data, the second plot is the population data from 10 years ago, and the third plot is the population change.

The data visualization found the greatest population was in West Houston area which is also experiencing the greatest population increase. While the Eastern neighborhoods are far less populated and have slower population growth.

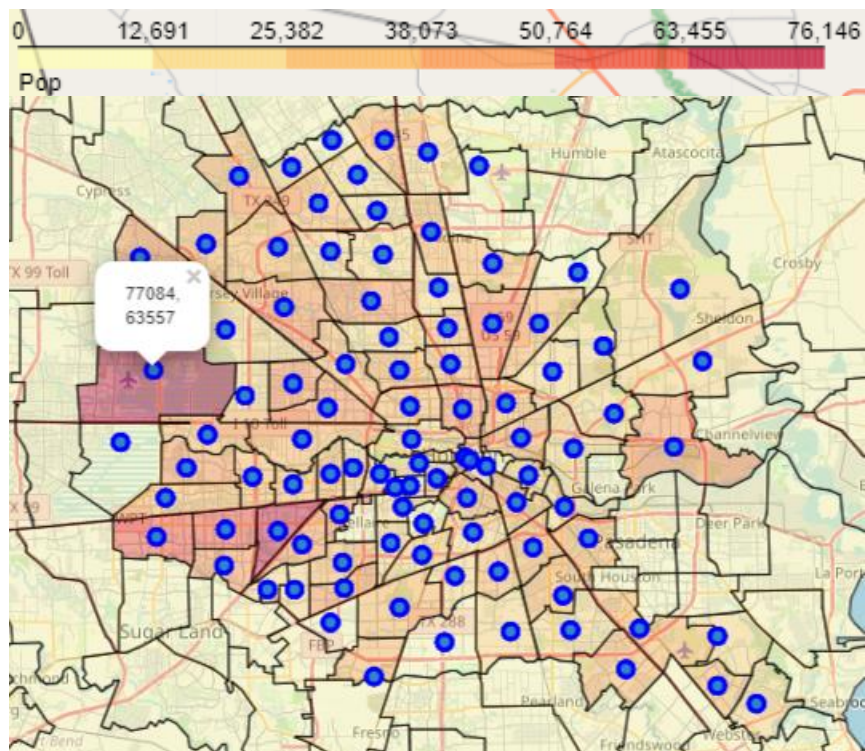


Figure 1: Current Population

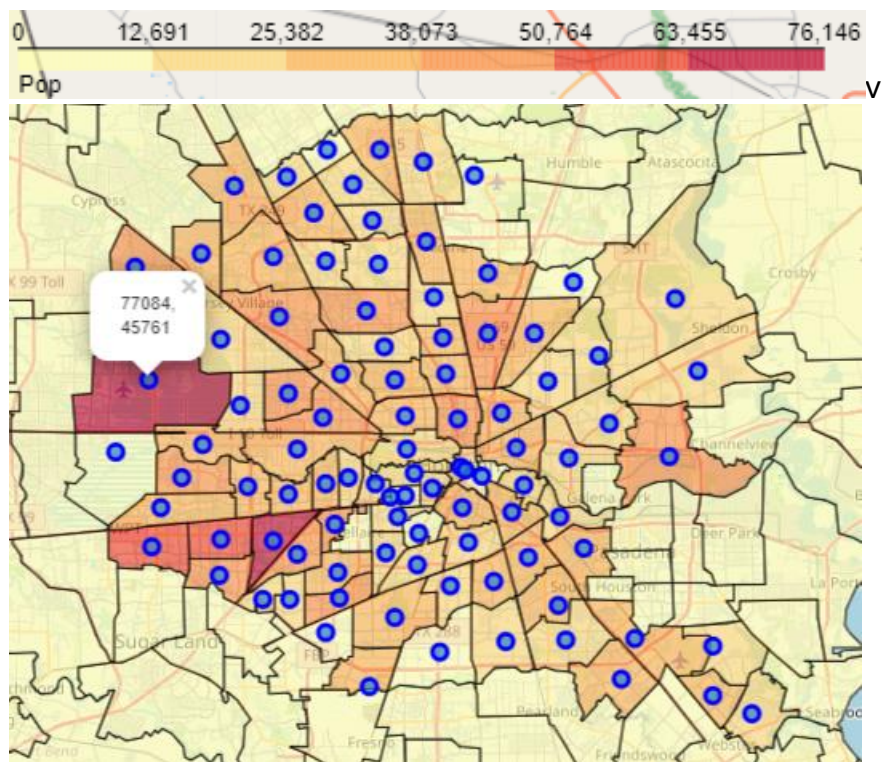


Figure 2: Old Population



## 5 Next Steps

The next steps for this analysis will be to determine why there is such a large population in certain areas while very small in others. This analysis will include an evaluation of the major industry in each of the zip codes to look a for a correlation.