

# **IBM Applied Data Science Capstone**

## **Using Python for researching a neighborhood before buying a house in the Southwest Houston area**

By: Dante A. Zorrilla

June 2020



# **1. Introduction**

Houston is one of the fastest growing and most diverse places in the United States. Wealth of job opportunities is bringing people to move in Houston, and thanks to the influx of new residents and median income growth rate that exceeds the national rate, newer and cheaper houses are being built.

## ***1.1. Description of the Problem***

Buying a house in Houston seems like a no brainer. However, for most people the whole process of buying a house is a hassle; it doesn't matter if the house will be for personal use or rental purposes, every person should conduct an extensive research before making such an important decision. Researching the market as per preference and budget along with looking into the neighborhoods is time consuming, therefore with this project I am looking to ease the process by applying data science methodology.

## ***1.2. Target Audience***

This project is useful for people looking to buy a house in Southwest Houston area. I will use Python language to gather real estate information, crime index per neighborhood and venues from different sources and present them to you in order to help you buying a new house in Houston.

## 2. Data

For this project we will use the following data:

- List of real estate properties in Southwest Houston area extracted from [zillow.com](https://www.zillow.com).
- Crime index for each neighborhood from [neighborhoodscout.com](https://neighborhoodscout.com)
- Geojson file containing the map of the neighborhoods in Southwest Houston created in [geojson.io](https://geojson.io).
- List of venues for different neighborhoods extracted using Foursquare API.