The BATTLE OF NEIGHBORHOODS

Selecting Neighborhood of Houston

1 Introduction/Business Problem

1.1 Background

Houston is the largest city in Texas, the fourth-largest in America, and brimming with opportunity, excitement, culture and history. A diverse city where more than 145 languages are spoken results in a food, art and community fusion unmatched in other destinations.

Unparalleled performing arts and nightlife activity keep this city running 24/7. The city's relatively compact Downtown includes the Theater District, home to the renowned Houston Grand Opera, and the Historic District, with 19th-century architecture and upscale restaurants. Tucked between the bustling Midtown and Medical Center areas, rests Houston's leafy, park-like Museum District. Nineteen world-class institutions, including the Menil Collection and Museum of Fine Arts, Houston are clustered in this area, drawing seven million visitors to the district each year.

1.2 Business Problem

As a visitor who has never been to Houston before, it is always difficult to find out best places to visit and find something which really suits their liking in the restricted time period. The visitor may enjoy Houston's outstanding performing and visual arts venues. Try one of the countless restaurants available, offering cuisine in everything from Tex Mex and South American to Middle Eastern and Vietnamese. For sports fans we have local teams representing all major sports. Do some shopping; Houston offers something to fit every budget - from the exclusive shops in Houston's Uptown area to the outlet malls just outside the City.

Every person who visits Houston faces this problem to find suitable top-rated locations of their liking. Therefore, as part of this Capstone project I tried solving this problem by segregate different neighborhoods to find different types of locations based on the surrounding.

1.3 Interest

We will use data science to generate and clustering the neighborhood in Houston. The advantages and feature of venue categories in each Houston neighborhood will then be clearly expressed so that the best neighborhood to visit can be chose by our target tourist.

2 Data

2.1 Data Sources

Following data sources are used for the analysis of neighborhoods of Houston:

1. The name, location and boundaries for all the Houston neighborhoods can be retrieved from the following online open data : https://opendata.arcgis.com/datasets/deae15e7f3db4b198946c17f0d169c39_3.geojson
2. Get the latitude and longitude values of Houston using geopy library.
3. Explore the venues and neighborhoods using Foursquare API.

2.2 Data Cleaning

The data of Houston neighborhoods is imported and transformed to pandas dataframe.

The latitude and longitude of each Houston neighborhood can be calculated by taking the average center of all of its shape boundary points. If a neighborhood is defined by two or two more shape areas, only the first shape area will be used for the calculation of the latitude and longitude.