Coursera Capstone Project- The Battle of Neighborhoods (Week 1)

Introduction

Greater Seattle is home to Amazon and Microsoft, two of the biggest technology companies in the world. There are more than a dozen companies that have headquarters in Seattle including Aerospace manufacturer Boeing. This makes Seattle real estate one of the hottest in the nation. According to Redfin.com, Seattle is rated 91 out of 100 Most Competitive where most homes get multiple offers often with waved contingencies. Homes sell for about 1% to 4% above list price and go pending in about 3 to 7 days. Some homes sell for as high as 22% above list price in that same site.

Audience

This makes it a perfect place for property developers and stakeholders to build and invest in this hot real estate market!

Business Problem

In this project we are going to analyze where is best location to build a new home. To do so we are going to use Seattle Neighborhoods and location data.

Data (Types of data needed)

-First, we need to get the list of neighborhoods data in order to explore city

-We then get latitude and longitude of the neighborhoods using python package geocoder.

-We also need to utilize the Foursquare Venues data and machine learning tool k-means clustering to group the neighborhoods into clusters.

Where we get the data

We get our data from https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Seattle which has all the 127 neighborhoods in Seattle.

Techniques used

Since our data is not readily available, we will have to use web scraping techniques such as beautiful soup or simply pandas to extract our table from the web page.

We will use geocoders package in python to get the geographical coordinates latitude and longitude of the neighborhoods.

We then use python request package to request venues data from the Foursquare location data and k-means clustering to cluster neighborhoods. The important step of visualizing the data in a map we use the python Folium package.

There are several steps used and that would be in the next section!

THANK YOU!!!