The Battle of Neighborhoods

Capstone Project

Daniela Moniz

1. Introduction

Over the last years, the Portugal's housing market has reached highs with a general price increase of about 9% on an annual basis, mainly in the capital city (Lisbon). Therefore, the main goal of this project is to help those who are looking to buy a house in Lisbon. In this context, an analysis of the variation of costs over the different neighborhoods is considered as well as a differentiation of the different places as residential or commercial areas based on the venues nearby.

2. Data

The information of the Portugal's housing market is based on a Kaggle data-set (Lisbon House Prices) that includes the house price, price m^2 for the different parishes in Lisbon. This data also includes the latitude and longitude that can be utilized for collecting the closes venues by the Foursquare application. After the data collection, the k-means clustering is utilized to divide the parish into residential and commercial areas. Moreover, the Folium package of Python will help to visualize the different characteristics of the clusters.