

Grouping properties in the city of Aracaju based on the characteristics of the neighborhood venues

Eduardo Fillipe da Silva Reis

Aracaju, SE - 2021

1. Introduction

Aracaju is the capital of the state of Sergipe, one of the smallest existing capitals in Brazil. Due to the last fact presented, it is not always easy to find information about properties for sale in the region. Such information can be found on specialized sales websites or through salespeople. However, the aforementioned options do not succinctly present information on trade in the neighborhood of the properties.

Thus, the work developed aims to:

1. Create an open DataSet with information on Latitude, Longitude and Price of properties in the Sergipe Capital and information on the trade around each property, containing the category, latitude and longitude of each trade
2. Group the properties of the DataSet in clusters according to the trade information around each property, so that it is possible, then, to make suggestions of properties from others, listing prices and other information.

The results of this work can be used by sellers and people interested in purchasing a new property in the region.

2. The DataSet

1. Properties Data

The data on real estate were extracted using Web Scraping techniques from the website <https://www.vivareal.com.br/> which works as a social network for the sale of houses and land.

2. Venues Data

Venues data were acquired using the Apis FourSquare suite, which gathers global geographic data. For each property, trade information was added within a radius of 600m from each property.

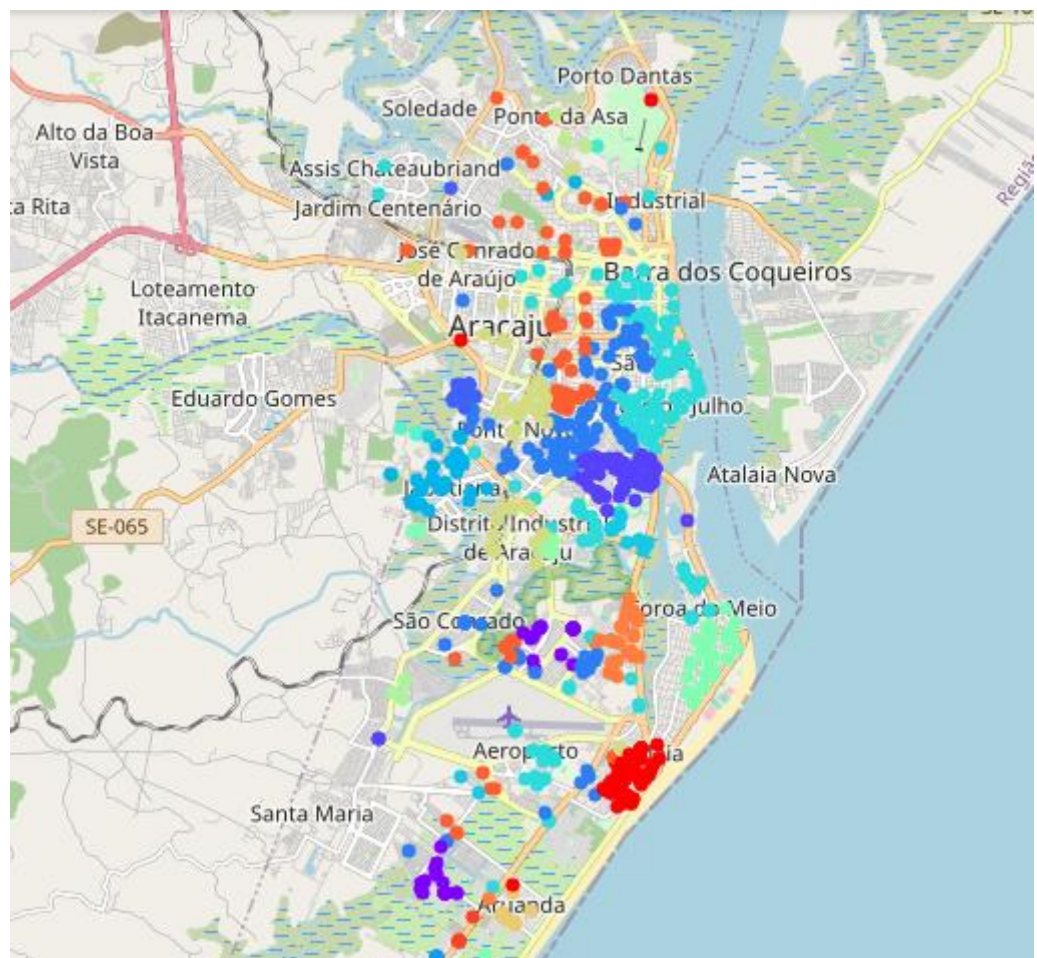
3. Results

1. The DataSets

The DataSet created in this work are one of the main results, as they gather in a clear, concise and clean way the information about real estate and commerce in the city of Aracaju and its relations.

2. Properties Sugestion

Another result of this work, is the application capable of suggesting the purchase of real estate based on his trade information. For this, the K-Means algorithm was used to group the data, obtaining the following result:



Thus, using any given property as an argument, the algorithm will return similar properties, using the neighborhoods, prices, distance and other information available in the DataSet as filters.

4. Trabalhos futuros

The realization of this work opens space for several others to be carried out and for several questions to be asked to the data, such as: “Do the characteristics of the trade influence the prices of the properties? How? Is there a direct relationship? Which?” In addition, it also allows applications to be created that use the algorithm created to improve existing search engines.