The Battle of Neighborhoods: Lima and Callao

Carlos Alberto Palomino Chalco

Introduction

The analysis that will be carried out in this project is the comparison of two important cities of Peru: Lima and Callao.

Lima is the capital and most populous city of the Republic of Peru.6 It is located on the central coast of the country, on the shores of the Pacific Ocean, forming an extensive and populous urban area known as Metropolitan Lima, flanked by the coastal desert and extended over the valleys of the Chillón, Rímac and Lurín rivers. According to the Peruvian census of 2017, Lima has more than 8.5 million inhabitants; note 1 while its urban agglomeration has more than 11 million inhabitants, 8 30% of the Peruvian population, figures that make it in the most populous city in the country.

El Callao is a port city located in the constitutional province of Callao, which is located in the center-west of Peru and in turn on the central coast of the Peruvian coast and in the central western area of South America. Being on the shores of the Pacific Ocean, it develops to the west of the province of Lima and 15 kilometers from the historic center of Lima, a city with which it is conurbed.

Objective

Throughout this project, the study will be conducted with a focus on the classification of the area using Foursquare data, segmentation and clustering.

The main objective of this data analysis project is to segment areas of Lima and Callao based on most common places captured from Foursquare.

In addition, through the use of segmentation and clustering, we will determine the following:

- If there is similarity or difference of both cities.

- The classification of the area located within the city, whether residential, tourist, sport, art or others.

Data

The data for this project have been acquired from the wikipedia pages and have also been restructured to a csv file for a more understandable reading and manipulation.

The links to the files are:

- https: //github.com/cpalominoch/Coursera_Capstone/blob/master/Lima_df.csv

- https: //github.com/cpalominoch/Coursera_Capstone/blob/master/Callao_df.csv

The coordinates were obtained using google api.

Foursquare api is being used for segementation and clustering.