The Battle of the Neighborhoods – Report

1. 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.

2. Objectives

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.

3. Data

The data for this project has 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

	Postal Code	District	Area
0	15011	Ate	Ate_15011
1	15012	Ate	Ate_15012
2	15023	Ate	Ate_15023
3	15479	Ate	Ate_15479
4	15483	Ate	Ate_15483

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

	Postal Code	District	Area
0	7001	Callao	Callao_7001
1	7006	Carmen de la Legua Reynoso	Carmen de la Legua Reynoso_7006
2	7011	Bellavista	Bellavista_7011
3	7016	La Perla	La Perla_7016
4	7021	La Punta	La Punta_7021

The coordinates were obtained using google api.

	Postal Code	District	Area	Latitude	Longitude
0	15011	Ate	Ate_15011	-12.034139	-76.949903
1	15012	Ate	Ate_15012	-12.056509	-76.949903
2	15023	Ate	Ate_15023	-12.099424	-76.970067
3	15479	Ate	Ate_15479	-12.012905	-76.797125
4	15483	Ate	Ate_15483	-12.025183	-76.820198
5	15487	Ate	Ate_15487	-12.027929	-76.877861
6	15491	Ate	Ate_15491	-12.037547	-76.903801
7	15494	Ate	Ate_15494	-12.041299	-76.923973
8	15498	Ate	Ate_15498	-12.027729	-76.932617
9	15063	Barranco	Barranco_15063	-12.157589	-77.016144

	Postal Code	District	Area	Latitude	Longitude
0	7001	Callao	Callao_7001	37.926805	-122.056054
1	7006	Carmen de la Legua Reynoso	Carmen de la Legua Reynoso_7006	-12.046231	-77.088103
2	7011	Bellavista	Bellavista_7011	-12.064096	-77.111121
3	7016	La Perla	La Perla_7016	-12.070231	-77.122629
4	7021	La Punta	La Punta_7021	-12.059843	-77.139888
5	7026	Callao	Callao_7026	-12.043178	-77.112560
6	7031	Callao	Callao_7031	-12.005750	-77.119752
7	7036	Callao	Callao_7036	-12.013203	-77.099613
8	7041	Callao	Callao_7041	-12.025722	-77.098174
9	7051	Ventanilla	Ventanilla_7051	-11.882764	-77.119752

Foursquare api is being used for segmentation and clustering.

Two districts will be taken as reference: Santiago de Surco (Lima) and Callao (Callao).

4. Methodology

In this project, we will direct our efforts to detect similarities or differences between Lima and Callao, with respect to various areas of interest such as: sports, food, tourism and others.

In the first step, data has been collected from the zip code, district and area of Lima and Callao. The districts of Santiago de Surco (Lima) and Callao (Callao) will be taken as shown, and the areas will be reflated on maps.

The second step in our analysis will use the Foursquare API to obtain places in the surrounding area of Santiago de Surco and Callao.

In the third step, we will focus on finding each neighborhood along with the 5 most common places. We will present a map of all those locations, but we will also create groups (using k-means clustering) of those locations to identify general areas that differentiate or have similarities between both cities.

5. Analysis

Santiago de Surco

Lima has 162 district and 286 areas. For the example, the district of Santiago de Surco will be taken, which has 4 areas according to zip code.
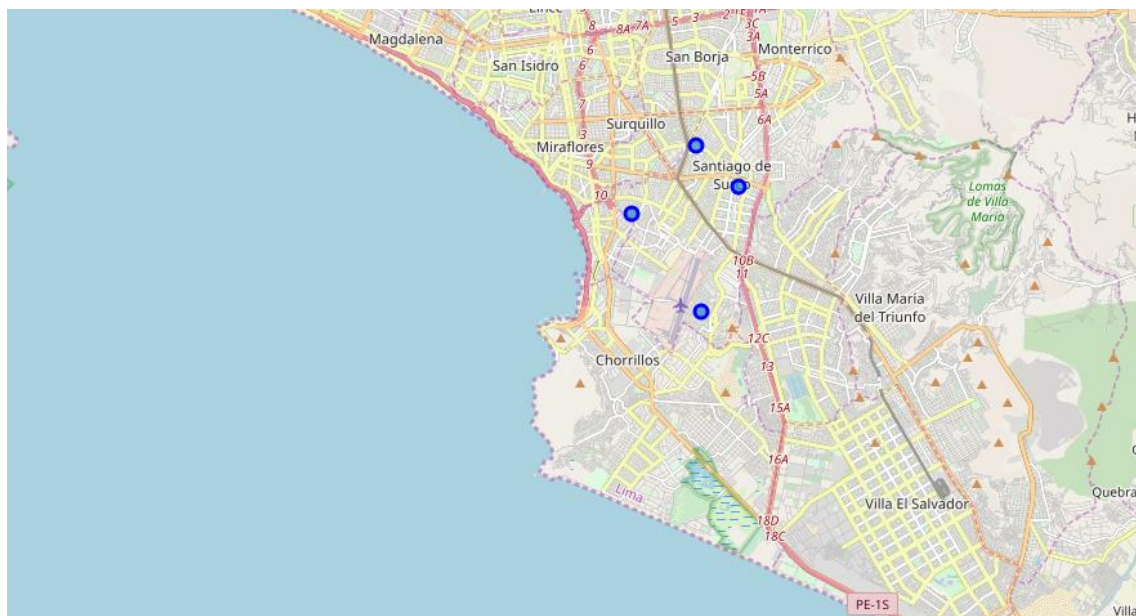
Using Foursquare API to get venues at surrounding area of Santiago de Surco.

25 venues were returned by Foursquare for Santiago de Surco, Lima.

	name	categories	lat	lng
0	Embarcadero 41 Fusión	Seafood Restaurant	-12.120941	-76.996540
1	El Piombino	Snack Place	-12.120579	-76.995330
2	La Casa del Ceviche	Seafood Restaurant	-12.120855	-76.996136
3	Wing Factory	Wings Joint	-12.121674	-76.992644
4	Pisco y Pesca	Seafood Restaurant	-12.118546	-76.997427

See each neighborhood along with the 5 most common places of Santiago de Surco.

	Area	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	Santiago de Surco_15038	Bakery	Seafood Restaurant	Peruvian Restaurant	Pool	Snack Place	Wings Joint	Sandwich Place	Italian Restaurant
1	Santiago de Surco_15039	Pizza Place	Sandwich Place	Coffee Shop	Fast Food Restaurant	Park	Chinese Restaurant	Bakery	Peruvian Restaurant
2	Santiago de Surco_15049	Nightclub	Athletics & Sports	Pub	Soccer Stadium	Trail	Donut Shop	Bakery	Music Venue
3	Santiago de Surco_15054	Asian Restaurant	Fried Chicken Joint	BBQ Joint	Gym / Fitness Center	Burger Joint	Park	Peruvian Restaurant	Dog Run

Using K-mean Cluster

	Postal Code	District	Area	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	15038	Santiago de Surco	Santiago de Surco_15038	-12.120241	-76.995987	0	Bakery	Seafood Restaurant	Peruvian Restaurant	Pool	Snack Place	Wings Joint	Sandwich Place	Italian Restaurant
1	15039	Santiago de Surco	Santiago de Surco_15039	-12.130890	-76.984468	0	Pizza Place	Sandwich Place	Coffee Shop	Fast Food Restaurant	Park	Chinese Restaurant	Bakery	Peruvian Restaurant
2	15049	Santiago de Surco	Santiago de Surco_15049	-12.137855	-77.013264	1	Nightclub	Athletics & Sports	Pub	Soccer Stadium	Trail	Donut Shop	Bakery	Music Venue
3	15054	Santiago de Surco	Santiago de Surco_15054	-12.163562	-76.994547	2	Asian Restaurant	Fried Chicken Joint	BBQ Joint	Gym / Fitness Center	Burger Joint	Park	Peruvian Restaurant	Dog Run
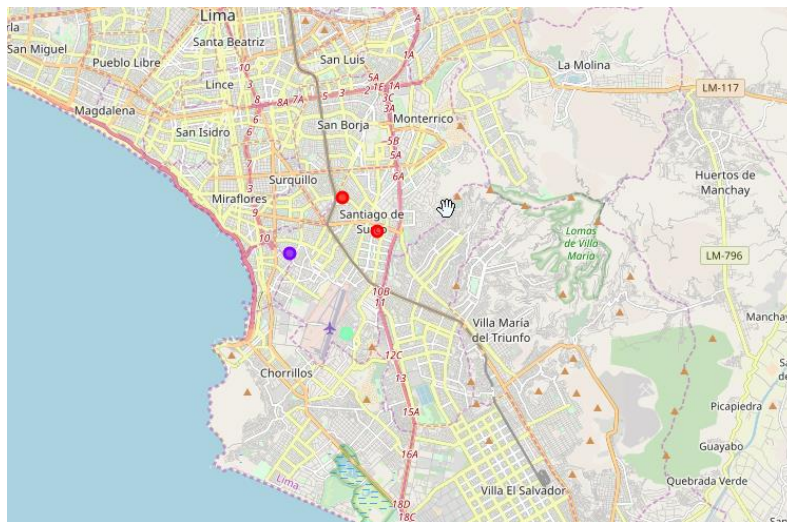
Callao

Callao has 7 district and 15 areas. For the example, the district of Callao will be taken, which has 4 areas according to zip code.

Using Foursquare API to get venues at surrounding area of Callao.
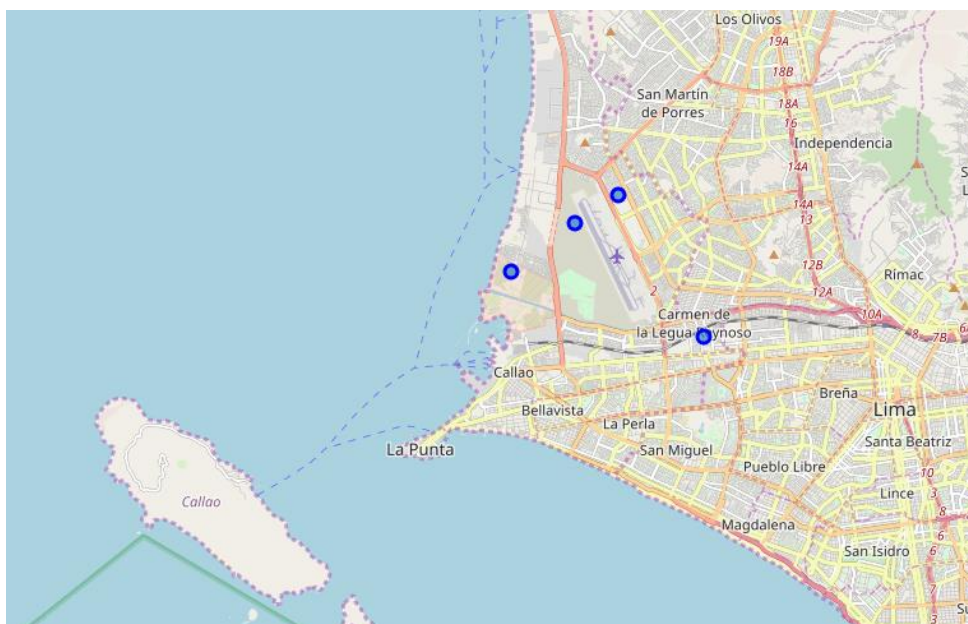
```
22 venues were returned by Foursquare for Callao, Callao.
```

	name	categories	lat	lng
0	iLoveKickboxing	Boxing Gym	37.926769	-122.056626
1	Renaissance ClubSport Walnut Creek Hotel	Hotel	37.925007	-122.056368
2	Bay Club Walnut Creek	Gym / Fitness Center	37.925261	-122.056904
3	Parada	Peruvian Restaurant	37.926811	-122.055849
4	Taheri's Mediterranean	Mediterranean Restaurant	37.928452	-122.057590

See each neighborhood along with the 5 most common places of Callao.

	Area	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	Callao_7001	Hotel	Coffee Shop	Mediterranean Restaurant	Bakery	Boxing Gym	Breakfast Spot	Café	Chinese Restaurant
1	Callao_7026	BBQ Joint	Big Box Store	Electronics Store	Market	Fish Market	Train Station	Gym / Fitness Center	Bakery
2	Callao_7031	Airport	Breakfast Spot	Trail	BBQ Joint	Bakery	Big Box Store	Boxing Gym	Burger Joint
3	Callao_7036	Plaza	Burger Joint	Sushi Restaurant	Bakery	Restaurant	Pizza Place	Chinese Restaurant	Peruvian Restaurant
4	Callao_7041	Bakery	Train Station	Trail	BBQ Joint	Big Box Store	Boxing Gym	Breakfast Spot	Burger Joint

Using K-mean Cluster

	Postal Code	District	Area	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	7001	Callao	Callao_7001	37.926805	-122.056054	1	Hotel	Coffee Shop	Mediterranean Restaurant	Bakery	Boxing Gym	Breakfast Spot	Café	Chinese Restaurant
1	7026	Callao	Callao_7026	-12.043178	-77.112560	1	BBQ Joint	Big Box Store	Electronics Store	Market	Fish Market	Train Station	Gym / Fitness Center	Bakery
2	7031	Callao	Callao_7031	-12.005750	-77.119752	2	Airport	Breakfast Spot	Trail	BBQ Joint	Bakery	Big Box Store	Boxing Gym	Burger Joint
3	7036	Callao	Callao_7036	-12.013203	-77.099613	1	Plaza	Burger Joint	Sushi Restaurant	Bakery	Restaurant	Pizza Place	Chinese Restaurant	Peruvian Restaurant
4	7041	Callao	Callao_7041	-12.025722	-77.098174	0	Bakery	Train Station	Trail	BBQ Joint	Big Box Store	Boxing Gym	Breakfast Spot	Burger Joint
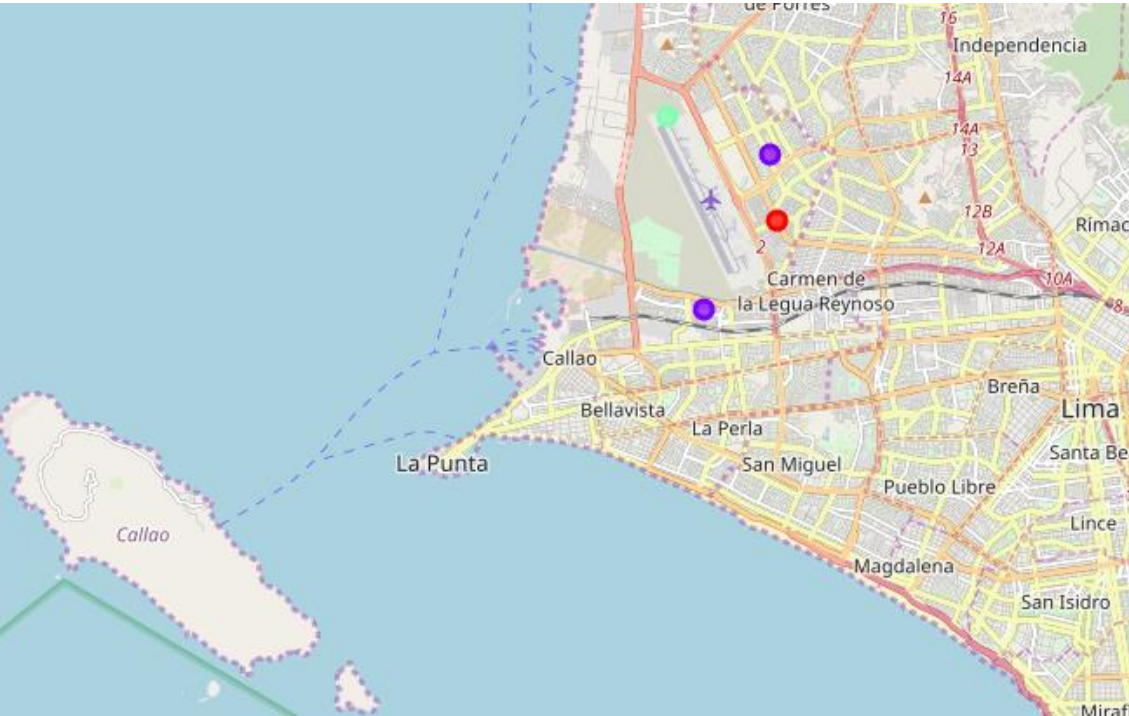
6. Results and Discussion

After performing the analysis of the district of Santiago de Surco (Lima), it can be seen that there are categories associated with: restaurant, sport, pub, park, fast food.

	Area	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	Santiago de Surco_15038	0	Bakery	Seafood Restaurant	Peruvian Restaurant	Pool	Snack Place	Wings Joint	Sandwich Place	Italian Restaurant
1	Santiago de Surco_15039	0	Pizza Place	Sandwich Place	Coffee Shop	Fast Food Restaurant	Park	Chinese Restaurant	Bakery	Peruvian Restaurant

	Area	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
2	Santiago de Surco_15049	1	Nightclub	Athletics & Sports	Pub	Soccer Stadium	Trail	Donut Shop	Bakery	Music Venue

	Area	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
3	Santiago de Surco_15054	2	Asian Restaurant	Fried Chicken Joint	BBQ Joint	Gym / Fitness Center	Burger Joint	Park	Peruvian Restaurant	Dog Run

Regarding the analysis of the district of Callao (Callao), It can be seen that there are categories associated with: restaurant, gym, fast food, hotel, market, airport.

	Area	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
4	Callao_7041	0	Bakery	Train Station	Trail	BBQ Joint	Big Box Store	Boxing Gym	Breakfast Spot	Burger Joint

	Area	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	Callao_7001	1	Hotel	Coffee Shop	Mediterranean Restaurant	Bakery	Boxing Gym	Breakfast Spot	Café	Chinese Restaurant
1	Callao_7026	1	BBQ Joint	Big Box Store	Electronics Store	Market	Fish Market	Train Station	Gym / Fitness Center	Bakery
3	Callao_7036	1	Plaza	Burger Joint	Sushi Restaurant	Bakery	Restaurant	Pizza Place	Chinese Restaurant	Peruvian Restaurant

	Area	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
2	Callao_7031	2	Airport	Breakfast Spot	Trail	BBQ Joint	Bakery	Big Box Store	Boxing Gym	Burger Joint

7. Conclusion

As a result of the analysis it can be concluded that both cities have categories of places in common associated with business, gastronomy and tourism.

Lima is the capital of Peru.

Callao is a great international exchange center for having airports and ports.