



## SIMILAR CITIES IN SOUTH AMERICA

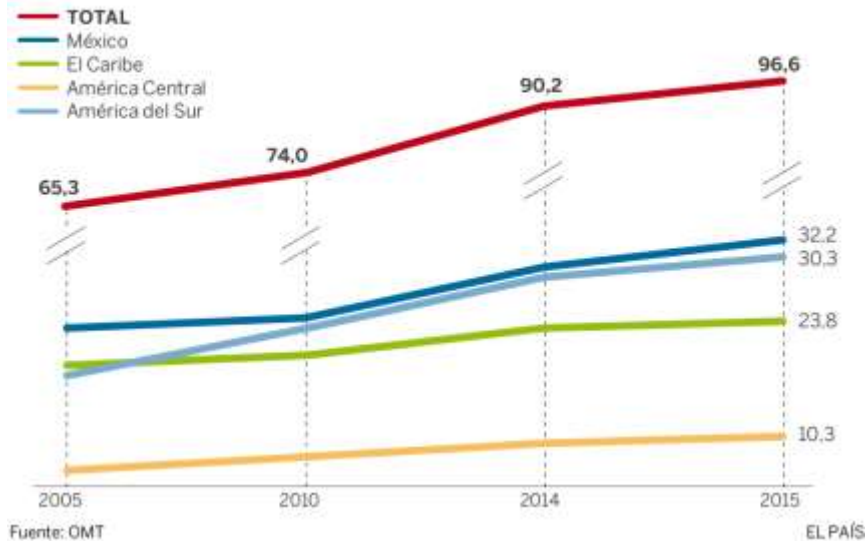


# INTRODUCTION

The problem would be to determine which cities in South America are similar, so that it is an additional information for tourists or agencies, when making a trip.

## LLEGADAS DE TURISTAS INTERNACIONALES A AMÉRICA LATINA Y EL CARIBE

En millones de personas



According to the graph in recent years the trend in the number of tourists visiting South America is increasing, which is why it is necessary to present additional information that may be of interest when visiting a city.

# DATA

- **Foursquare:**

<https://developer.foursquare.com/>

- **GeoData**

<https://www.geodatos.net/poblacion>

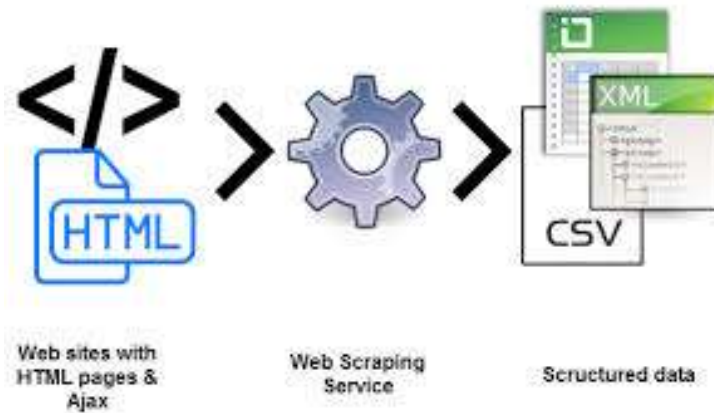
For example, the ten most populated cities in Peru

País	Ciudad	Población	Coordenadas
Perú	Lima	7,737,002 habitantes	-12.043, -77.028
Perú	Arequipa	841,130 habitantes	-16.399, -71.535
Perú	El Callao	813,264 habitantes	-12.057, -77.118
Perú	Trujillo	747,450 habitantes	-8.116, -79.03
Perú	Chiclayo	577,375 habitantes	-6.771, -79.841
Perú	Iquitos	437,620 habitantes	-3.749, -73.254
Perú	Huancayo	376,657 habitantes	-12.065, -75.205
Perú	Piura	325,466 habitantes	-5.194, -80.633
Perú	Chimbote	316,966 habitantes	-9.085, -78.578
Perú	Cusco	312,140 habitantes	-13.523, -71.967

Figure 1: 10 most populated cities in Peru  
Source: GeoData

# METHODOLOGY

- **Web scraping:**



- **K-means:**

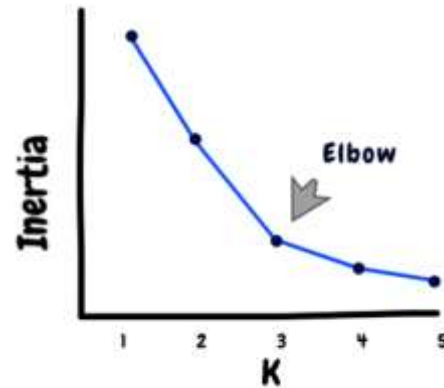


Figure 3: choice of the number of groups through inertia

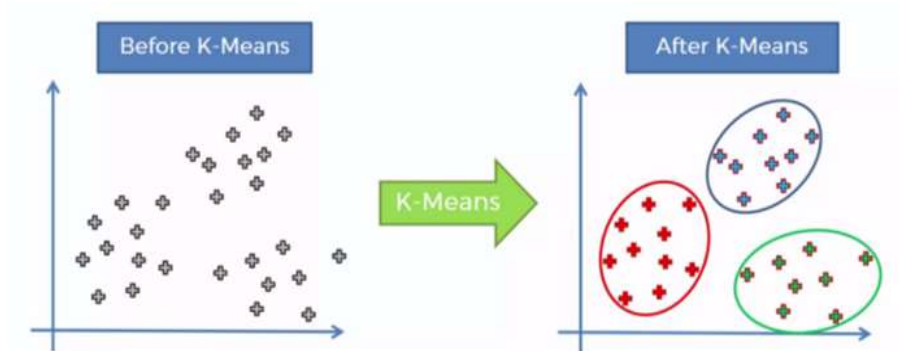


Figure 2: K-means

# EXPLORATORY ANÁLISIS

South American cities



Figure 4: analyzed cities in South America

Population of the cities



Figure 5: Population of the analysed cities

The most populated city is in Argentina followed by Brazil and Peru.

	pais	ciudad	poblacion	coordenadas	num_poblacion	latitude	longitude
80	Argentina	Buenos Aires	13,076,300 habitantes	-34.613, -58.377	13076300	-34.613	-58.377
40	Brasil	São Paulo	10,021,295 habitantes	-23.548, -46.636	10021295	-23.548	-46.636
0	Perú	Lima	7,737,002 habitantes	-12.043, -77.028	7737002	-12.043	-77.028
20	Colombia	Bogotá	7,674,366 habitantes	4.61, -74.082	7674366	4.61	-74.082
41	Brasil	Río de Janeiro	6,023,699 habitantes	-22.906, -43.182	6023699	-22.906	-43.182

# RESULTS



Cluster Labels	
4	69
0	24
3	4
5	1
2	1
1	1

Figure 7: Cluster of the cities

# DISCUSSIONS AND RECOMMENDATIONS

You could reduce the number of clusters and see the results, as we got three cities in one group. In addition, it would be grouped by means of oreo algorithm like DBscan, Hierarchical Cluster.

It would also be possible to compare these cities with the rest of the world and observe similarities between different continents, since the cultural differences are well marked.

According to the results it is recommended, for example, to the tourists, to visit Bolivia since they present cities with different characteristics, for the tourists who are already used to certain places and do not want to change they are recommended Brazil, Chile or Colombia.

It is recommended to generate an application that can compare cities and obtain those that are similar by grouping kmeans clusters and Foursquare data.

# CONCLUSION

In conclusion, there are two well-defined groups, and the rest would indicate cities that are interesting to analyze.

Generating this type of alternatives in more detail would be interesting for tourists who know nothing about South America.