Applied Data Science Capstone Course – Final Project

1. Introduction (Business Case)

Most of the tourists who visit Peru first land in Lima, the capital city, but most of them immediately take an airplane to Cusco, where the famous Machu Picchu is located. This is due to the abundant advertisement that highlights Machu Picchu as the most important touristic spot, reducing the chances of tourists visiting other interesting spots in Lima or in Cusco region too. However, tourists might not have the same interests. For example, some of them might be more interested in visiting historic places such Machu Picchu or similar Inca fortress, while others might find dynamic cities as Lima more exciting to visit. If we compare both cities, based on the website Versus.com [1] that gives unbiased comparisons using objective information from many categories, Lima has a higher average temperature (19.4°C vs. 12°C), has seaside beach, 4 more big sports facilities (stadiums, arenas, etc.), and has a bike-sharing system; on the other side, Cusco has lakes and smaller humidity rate (55.4% vs. 83%).
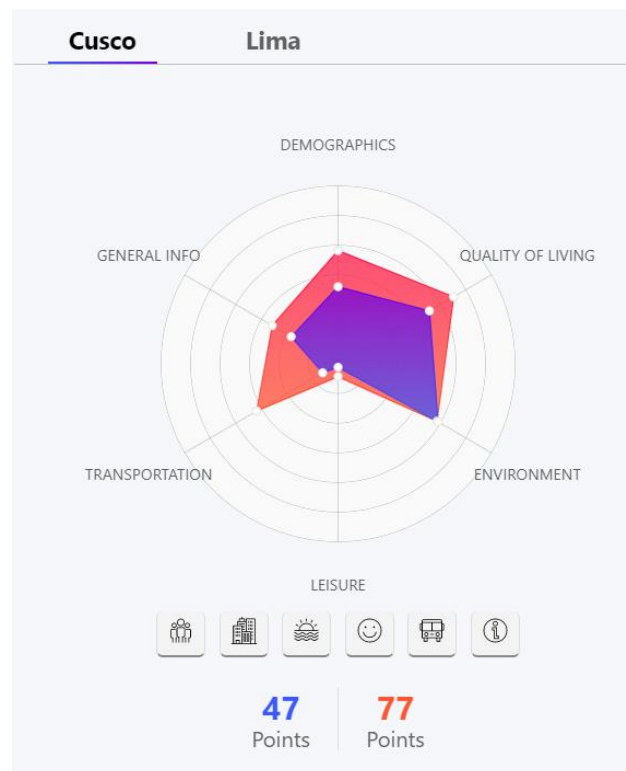

Fig. 1 – Screenshot of comparison between Lima and Cusco from Versus.com
(Cusco in purple, Lima in red)

As we can see, Lima can offer some advantages or may have other attractions than the historical spots located in Cusco. Still, visitors might not be aware of this, trusting blindly in touristic agencies, then getting disappointed when they get to know from other sources that they missed attractions in Cusco or Lima. In consequence, it would be important to explore both cities to find other interesting and beautiful hidden spots around already famous touristic spots. Additionally, it would be important to segment the touristic sites in both cities to create a profile for each cluster, so it would be easier to provide accurate information

to future visitors, helping them to choose where to visit according to their interests. This project aims to fulfill both activities (explore and segmenting), which can be helpful for touristic agencies to extend their packages or programs adding other interesting spots that are near already included sites and create packages according to clients' profiles or interests, resulting in an increased profit. On the other hand, it might result in a higher number of tourists visiting both cities which will help the local government to increase income from touristic activities.

2. Data Sources and Usage

Before introducing the list of data sources, it is important to indicate that Peru is divided in regions and each region is divided in provinces, which are also divided in districts. In this scenario, the first data source would be a list of all districts in Peru grouped by province and region [2], which has the format shown in Fig. 2. The website will be scrapped and then using PositionStack API [3] the coordinates of each district will be found.

Ubigeo ⬍	Departamento ⬍	Provincia ⬍	Distrito ⬍	Población 2017 ⬍
010101	Amazonas	Chachapoyas	Chachapoyas	32 589
010102	Amazonas	Chachapoyas	Asunción	262
010103	Amazonas	Chachapoyas	Balsas	1136
010104	Amazonas	Chachapoyas	Cheto	642
010105	Amazonas	Chachapoyas	Chiliquín	585
010106	Amazonas	Chachapoyas	Chuquibamba	1781
010107	Amazonas	Chachapoyas	Granada	480
010108	Amazonas	Chachapoyas	Huancas	1258
010109	Amazonas	Chachapoyas	La Jalca	3978
010110	Amazonas	Chachapoyas	Leimebamba	3620
010111	Amazonas	Chachapoyas	Levanto	794
010112	Amazonas	Chachapoyas	Magdalena	852
010113	Amazonas	Chachapoyas	Mariscal Castilla	1367
010114	Amazonas	Chachapoyas	Molinopampa	2176
010115	Amazonas	Chachapoyas	Montevideo	496
010116	Amazonas	Chachapoyas	Olleros	375
010117	Amazonas	Chachapoyas	Quinjalca	769
010118	Amazonas	Chachapoyas	San Francisco de Daguas	295
010119	Amazonas	Chachapoyas	San Isidro de Maino	580
010120	Amazonas	Chachapoyas	Soloco	1224
010121	Amazonas	Chachapoyas	Sonche	247
010201	Amazonas	Bagua	Bagua	27 102
010202	Amazonas	Bagua	Aramango	9765
010203	Amazonas	Bagua	Copallín	4595
010204	Amazonas	Bagua	El Parco	1216
010205	Amazonas	Bagua	Imaza	25 162
010206	Amazonas	Bagua	La Peca	6260
010301	Amazonas	Bongará	Jumbilla	1337
010302	Amazonas	Bongará	Chisquilla	306

Fig. 2 - List of all districts in Peru from Wikipedia website (only available in Spanish)

For the first objective, using the Foursquare API we will explore attractions (and their categories) located in the same district where are the popular touristic spots in each city listed by the Ministry of Foreign Trade and Tourism in its report of tourism development elaborated in 2018 [4] [5]. This will show other interesting sites that might be included in touristic packages and be visited while going to the already included famous spots.

For the second objective, the complete list of attractions in each city will be segmented and clustered, resulting in a division for each city composed by clusters with defined characteristics which can become new touristic packages that will be offered targeting specific visitors according to their interests.

3. Methodology (exploratory data analysis, machine learnings)

The first step was getting the complete list of districts in Peru from the Wikipedia site using the library BeautifulSoup, then the data was filtered to get Cusco and Lima region information and the result was saved in a pandas dataframe for each region.

```
      Region  Province        District
696   Cusco     Cusco           Cusco
697   Cusco     Cusco          Ccorca
698   Cusco     Cusco           Poroy
699   Cusco     Cusco     San Jerónimo
700   Cusco     Cusco    San Sebastián
```

Fig. 3 – First 5 rows of Cusco region dataframe

```
       Region  Province   District
1278    Lima      Lima       Lima
1279    Lima      Lima      Ancón
1280    Lima      Lima        Ate
1281    Lima      Lima   Barranco
1282    Lima      Lima      Breña
```

Fig. 4 – First 5 rows of Lima region dataframe

The dataframes were completed with coordinates of each district, which were retrieved from the PositionStack API. However, some of the district's information could not be extracted so those rows were removed.

	Region	Province	District	Latitude	Longitude
696	Cusco	Cusco	Cusco	-13.551969	-72.069168
697	Cusco	Cusco	Ccorca	-13.585000	-72.059720
698	Cusco	Cusco	Poroy	-13.495000	-72.043330
699	Cusco	Cusco	San Jerónimo	-13.545940	-71.884780
700	Cusco	Cusco	San Sebastián	-13.531380	-71.931430
...	...	...	...	...	...
803	Cusco	Urubamba	Huayllabamba	-13.338060	-72.065280
804	Cusco	Urubamba	Machupicchu	-13.158330	-72.531390
805	Cusco	Urubamba	Maras	-13.335000	-72.156670
806	Cusco	Urubamba	Ollantaytambo	-13.265008	-72.267208
807	Cusco	Urubamba	Yucay	-13.321390	-72.082220

112 rows × 5 columns

	Region	Province	District	Latitude	Longitude
696	Cusco	Cusco	Cusco	-13.551969	-72.069168
697	Cusco	Cusco	Ccorca	-13.585000	-72.059720
698	Cusco	Cusco	Poroy	-13.495000	-72.043330
699	Cusco	Cusco	San Jerónimo	-13.545940	-71.884780
700	Cusco	Cusco	San Sebastián	-13.531380	-71.931430
...	...	...	...	...	...
803	Cusco	Urubamba	Huayllabamba	-13.338060	-72.065280
804	Cusco	Urubamba	Machupicchu	-13.158330	-72.531390
805	Cusco	Urubamba	Maras	-13.335000	-72.156670
806	Cusco	Urubamba	Ollantaytambo	-13.265008	-72.267208
807	Cusco	Urubamba	Yucay	-13.321390	-72.082220

111 rows × 5 columns

Fig. 5 – Cusco dataframe: before cleaning (left), after cleaning (right)

	Region	Province	District	Latitude	Longitude
1278	Lima	Lima	Lima	-12.049141	-77.021217
1279	Lima	Lima	Ancón	-11.774270	-77.170060
1280	Lima	Lima	Ate	-12.050000	-76.966670
1281	Lima	Lima	Barranco	NaN	NaN
1282	Lima	Lima	Breña	NaN	NaN
...	...	...	...	...	...
1444	Lima	Yauyos	Tauripampa	-12.583330	-76.116670
1445	Lima	Yauyos	Tomas	-12.238610	-75.745280
1446	Lima	Yauyos	Tupe	-12.740830	-75.808330
1447	Lima	Yauyos	Viñac	-12.931110	-75.778610
1448	Lima	Yauyos	Vitis	-12.223610	-75.806390

171 rows × 5 columns

	Region	Province	District	Latitude	Longitude
1278	Lima	Lima	Lima	-12.049141	-77.021217
1279	Lima	Lima	Ancón	-11.774270	-77.170060
1280	Lima	Lima	Ate	-12.050000	-76.966670
1283	Lima	Lima	Carabayllo	-11.853330	-77.037780
1284	Lima	Lima	Chaclacayo	-12.068195	-76.861999
...	...	...	...	...	...
1444	Lima	Yauyos	Tauripampa	-12.583330	-76.116670
1445	Lima	Yauyos	Tomas	-12.238610	-75.745280
1446	Lima	Yauyos	Tupe	-12.740830	-75.808330
1447	Lima	Yauyos	Viñac	-12.931110	-75.778610
1448	Lima	Yauyos	Vitis	-12.223610	-75.806390

146 rows × 5 columns

Fig. 6 - Cusco dataframe: before cleaning (left), after cleaning (right)

Now, the districts where the most popular sites are located will be explored. First, Cusco region will be explored. The districts to be considered are where Machu Picchu, Cusco city, Aguas Calientes, Ollantaytambo, Valle Sagrado and Pisac are:

District	Province	Region
Machupicchu	Urubamba	Cusco
Ollantaytambo	Urubamba	Cusco
Yucay	Urubamba	Cusco
Cusco	Cusco	Cusco
Písac	Calca	Cusco

Fig. 7 – List of selected districts in Cusco

In case of Lima region, the districts to be considered are:

District	Province	Region
Lima	Lima	Lima
Miraflores	Lima	Lima
San Isidro	Lima	Lima
Barranco	Lima	Lima

Fig. 8 – List of selected districts in Lima

However, the coordinates for Barranco couldn't be retrieved from PositionStack API, so only 3 districts will be considered: Lima, Miraflores and San Isidro.

Using the Foursquare API, the information from venues located in each district will be retrieved, including venue name, venue coordinates and venue category. After, we list the top 10 most common venue categories in each district. Knowing the categories, also using the Foursquare API, we can get venues list from each category and select the ones with high rating to be a list of recommended spots. These steps were applied to data from each region separately.

For the second objective of this project, the total list of venues in each region was used. According to the number of venues in each category, K-Means algorithm was used to segment the districts in each region so we can know which type of venues can be encountered there.

4. Results

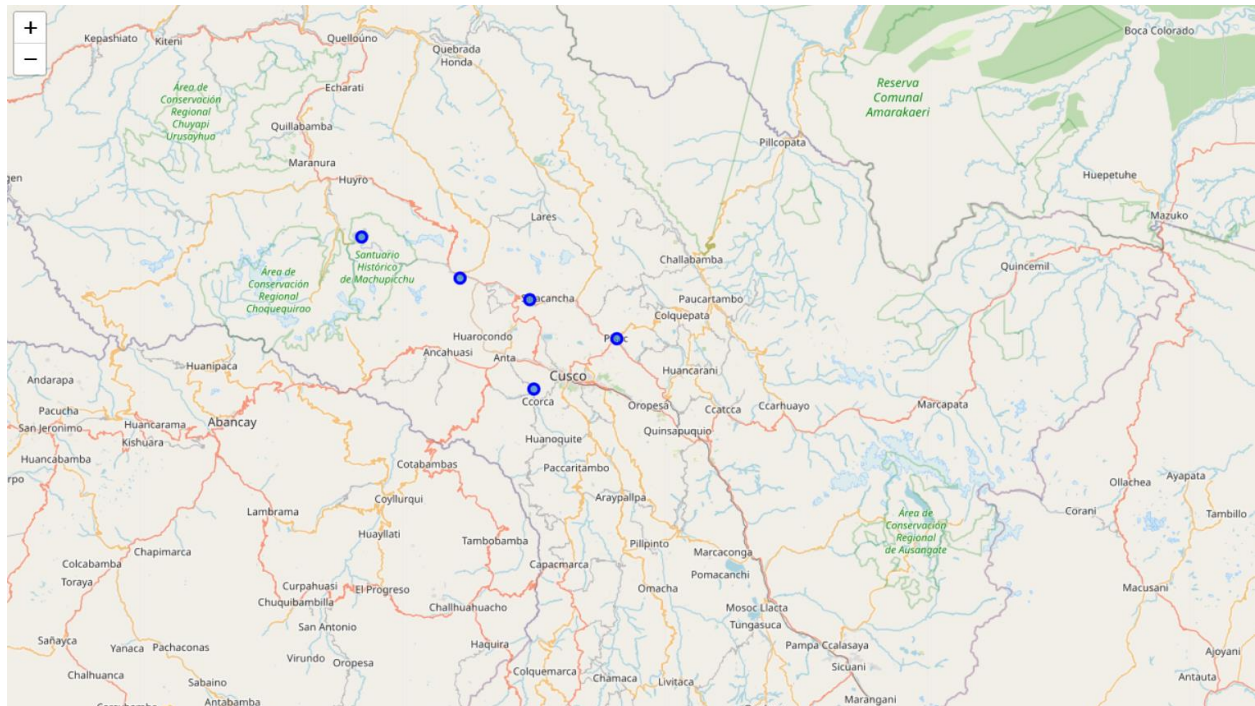The selected districts in Cusco where popular spots are located are shown in the Fig. 9.

Fig. 9 – Map with selected districts in Cusco

After analyzing the venues located in the selected districts, the Fig. 10 shows the top 10 popular venue categories in Cusco.

	District	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	9th Most Common Venue	10th Most Common Venue
0	Machupicchu	Hotel	Trail	Peruvian Restaurant	Hostel	Bed & Breakfast	Restaurant	Zoo	Italian Restaurant	Buffet	Food
1	Ollantaytambo	Hotel	Café	Train Station	Breakfast Spot	Peruvian Restaurant	Zoo	Italian Restaurant	Bed & Breakfast	Buffet	Food
2	Písac	Café	Jewelry Store	Peruvian Restaurant	Hotel	Latin American Restaurant	Market	Bakery	Plaza	Restaurant	Pizza Place
3	Yucay	Peruvian Restaurant	Café	Pizza Place	Hotel	Zoo	Italian Restaurant	Bed & Breakfast	Breakfast Spot	Buffet	Food

Fig. 10 – Top 10 popular venue categories in Cusco

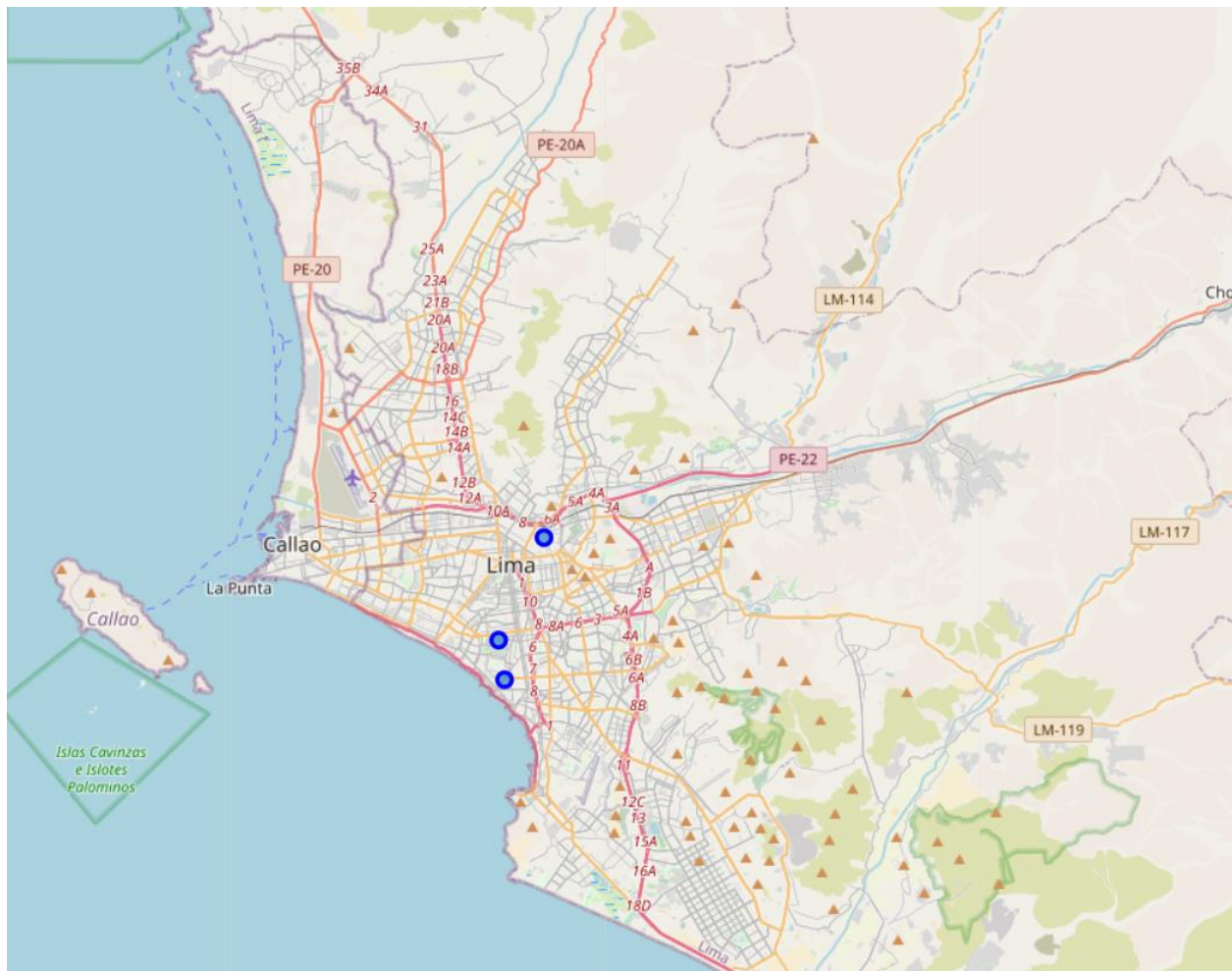The next figure shows the selected districts in Lima:

Fig. 11 – Map with selected districts in Lima

The resulted top 10 venue categories in Lima are shown in the next figure:

	District	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	9th Most Common Venue	10th Most Common Venue
0	Lima	Chinese Restaurant	Seafood Restaurant	Bakery	History Museum	Furniture / Home Store	Market	Dim Sum Restaurant	Park	College Academic Building	Plaza
1	Miraflores	Park	Seafood Restaurant	Café	Coffee Shop	Italian Restaurant	Restaurant	Hostel	Peruvian Restaurant	Chinese Restaurant	Ice Cream Shop
2	San Isidro	Hotel	Peruvian Restaurant	Coffee Shop	Restaurant	Italian Restaurant	Steakhouse	Yoga Studio	Park	Bistro	Hotel Bar

Fig. 12 – Top 10 venue categories in Lima

For the segmentation, the list of all venues in Cusco was retrieved using the Foursquare API, containing in total 53 venue categories:

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Ccorca	-13.585	-72.05972	Plaza Ccorca	-13.582261	-72.060280	Plaza
1	Poroy	-13.495	-72.04333	PeruRail - Estación Poroy | Poroy Station (Est...	-13.494642	-72.042468	Train Station
2	Poroy	-13.495	-72.04333	La Olla Peruana	-13.493507	-72.042513	Peruvian Restaurant
3	Poroy	-13.495	-72.04333	Arco de Ticatica	-13.494541	-72.042780	Historic Site
4	Poroy	-13.495	-72.04333	Chicharroneria Santa Barbara	-13.493830	-72.045506	BBQ Joint

Fig. 13 – List of all venues in Cusco region (first 5 rows)

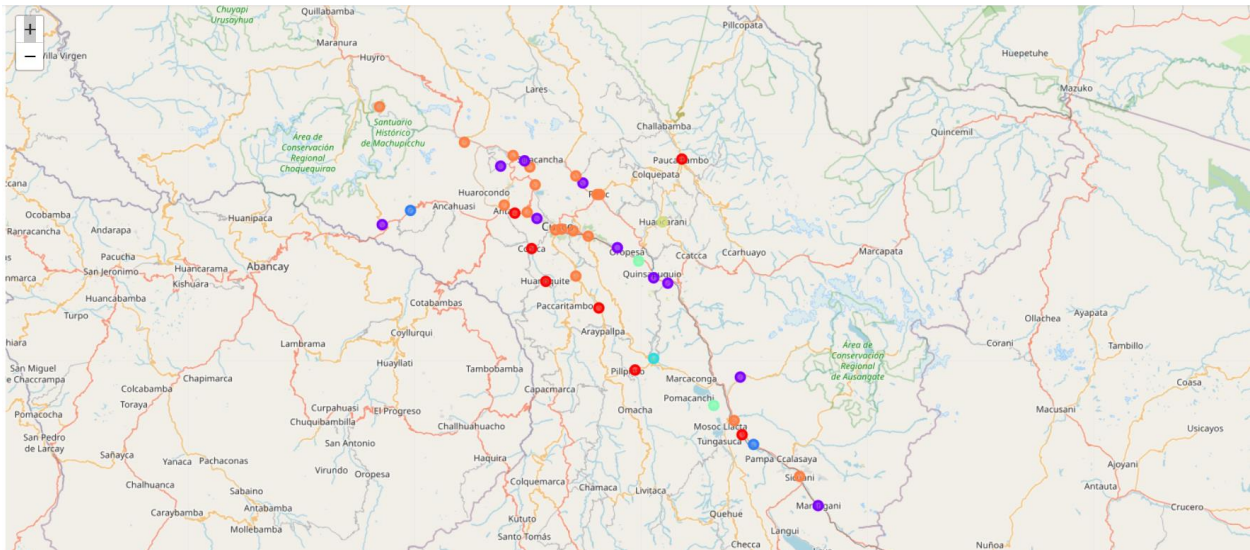The clustering resulted in 7 groups of districts and popular venue categories for each one:

Fig. 14 – Clusters of districts (Cusco region)

Cluster 1:

	District	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	9th Most Common Venue	10th Most Common Venue
697	Ccorca	0.0	Plaza	Zoo	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field
706	Acos	0.0	Plaza	Zoo	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field
718	Pucyura	0.0	Plaza	Zoo	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field
743	Tinta	0.0	History Museum	Plaza	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field
774	Paruro	0.0	Shopping Mall	Plaza	Zoo	Clothing Store	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field
778	Huanoquite	0.0	Bakery	Plaza	Zoo	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market
783	Paucartambo	0.0	Park	Plaza	Zoo	Deli / Bodega	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field

Cluster 2:

	District	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	9th Most Common Venue	10th Most Common Venue
698	Poroy	1.0	Peruvian Restaurant	Historic Site	Train Station	BBQ Joint	Deli / Bodega	Grocery Store	General Entertainment	Gas Station	Food	Flea Market
717	Mollepata	1.0	Peruvian Restaurant	Bed & Breakfast	Restaurant	Farmers Market	Zoo	Deli / Bodega	Grocery Store	General Entertainment	Gas Station	Food
721	Coya	1.0	Trail	Peruvian Restaurant	Plaza	Zoo	Deli / Bodega	Grocery Store	General Entertainment	Gas Station	Food	Flea Market
739	Marangani	1.0	Peruvian Restaurant	Zoo	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field
740	Pitumarca	1.0	Peruvian Restaurant	Farm	Zoo	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market
790	Andahuaylillas	1.0	Peruvian Restaurant	Plaza	History Museum	Zoo	Deli / Bodega	Grocery Store	General Entertainment	Gas Station	Food	Flea Market
795	Huaro	1.0	Lake	Peruvian Restaurant	Plaza	Zoo	History Museum	Grocery Store	General Entertainment	Gas Station	Food	Flea Market
799	Oropesa	1.0	Zoo	Peruvian Restaurant	Restaurant	Plaza	Deli / Bodega	Grocery Store	General Entertainment	Gas Station	Food	Flea Market
805	Maras	1.0	Plaza	Arts & Crafts Store	Restaurant	Peruvian Restaurant	Zoo	Deli / Bodega	Grocery Store	General Entertainment	Gas Station	Food
807	Yucay	1.0	Peruvian Restaurant	Pizza Place	Café	Zoo	Deli / Bodega	Grocery Store	General Entertainment	Gas Station	Food	Flea Market

Cluster 3:

	District	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	9th Most Common Venue	10th Most Common Venue
716	Limatambo	2.0	History Museum	Hostel	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field	Fast Food Restaurant
742	San Pedro	2.0	History Museum	Historic Site	Hostel	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field	Fast Food Restaurant

Cluster 4:

	District	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	9th Most Common Venue	10th Most Common Venue
704	Acomayo	3.0	Park	Zoo	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field

Cluster 5:

	District	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	9th Most Common Venue	10th Most Common Venue
705	Acopía	4.0	Lake	Zoo	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field
796	Lucre	4.0	Lake	Zoo	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field

Cluster 6:

	District	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	9th Most Common Venue	10th Most Common Venue
787	Huancarani	5.0	Café	Zoo	Deli / Bodega	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field

Cluster 7:

	District	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	9th Most Common Venue	10th Most Common Venue
699	San Jerónimo	6.0	Hotel	Latin American Restaurant	Flea Market	Plaza	Deli / Bodega	Grocery Store	General Entertainment	Gas Station	Food	Field
700	San Sebastián	6.0	Hotel	Supermarket	Plaza	Soup Place	Rental Car Location	Grocery Store	Clothing Store	General Entertainment	Gas Station	Food
701	Santiago	6.0	Hotel	Bus Station	Hostel	Plaza	Jewelry Store	Shopping Mall	Farmers Market	Asian Restaurant	Grocery Store	General Entertainment
703	Wánchaq	6.0	Hotel	Bed & Breakfast	Seafood Restaurant	Arts & Crafts Store	Train Station	Bus Station	Pizza Place	Plaza	Breakfast Spot	Asian Restaurant
711	Anta	6.0	Soccer Field	Zoo	Clothing Store	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field
713	Cachimayo	6.0	General Entertainment	Bike Trail	Zoo	Deli / Bodega	Historic Site	Grocery Store	Gas Station	Food	Flea Market	Field
722	Lamay	6.0	Hotel	Latin American Restaurant	Fast Food Restaurant	Peruvian Restaurant	Resort	Deli / Bodega	Grocery Store	General Entertainment	Gas Station	Food
724	Pisac	6.0	Jewelry Store	Peruvian Restaurant	Café	Latin American Restaurant	Market	Restaurant	Hotel	Bakery	Arts & Crafts Store	Plaza
726	Taray	6.0	Jewelry Store	Café	Latin American Restaurant	Peruvian Restaurant	Market	Restaurant	Plaza	Bakery	Arts & Crafts Store	Field
736	Sicuani	6.0	Hotel	Italian Restaurant	Fast Food Restaurant	Peruvian Restaurant	Restaurant	Bus Station	Plaza	Deli / Bodega	General Entertainment	Gas Station
738	Combapata	6.0	Rest Area	Zoo	Clothing Store	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field
752	Espinar	6.0	Restaurant	Hotel	Pizza Place	Clothing Store	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field
766	Kimbiri	6.0	Deli / Bodega	Bus Station	Zoo	Historic Site	Grocery Store	General Entertainment	Gas Station	Food	Flea Market	Field

In case of Lima region, we got in total 165 venue categories:

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Lima	-12.049141	-77.021217	Restaurante La Buena Muerte	-12.048123	-77.022988	Seafood Restaurant
1	Lima	-12.049141	-77.021217	Chifa Kong	-12.051862	-77.024901	Chinese Restaurant
2	Lima	-12.049141	-77.021217	Escuela Nacional Superior Autónoma de Bellas A...	-12.046799	-77.023974	College Academic Building
3	Lima	-12.049141	-77.021217	Wa Lok	-12.052227	-77.025085	Chinese Restaurant
4	Lima	-12.049141	-77.021217	Museo de la Moneda	-12.049476	-77.023991	History Museum

The clustering resulted in 5 groups of districts and popular venue categories for each one:
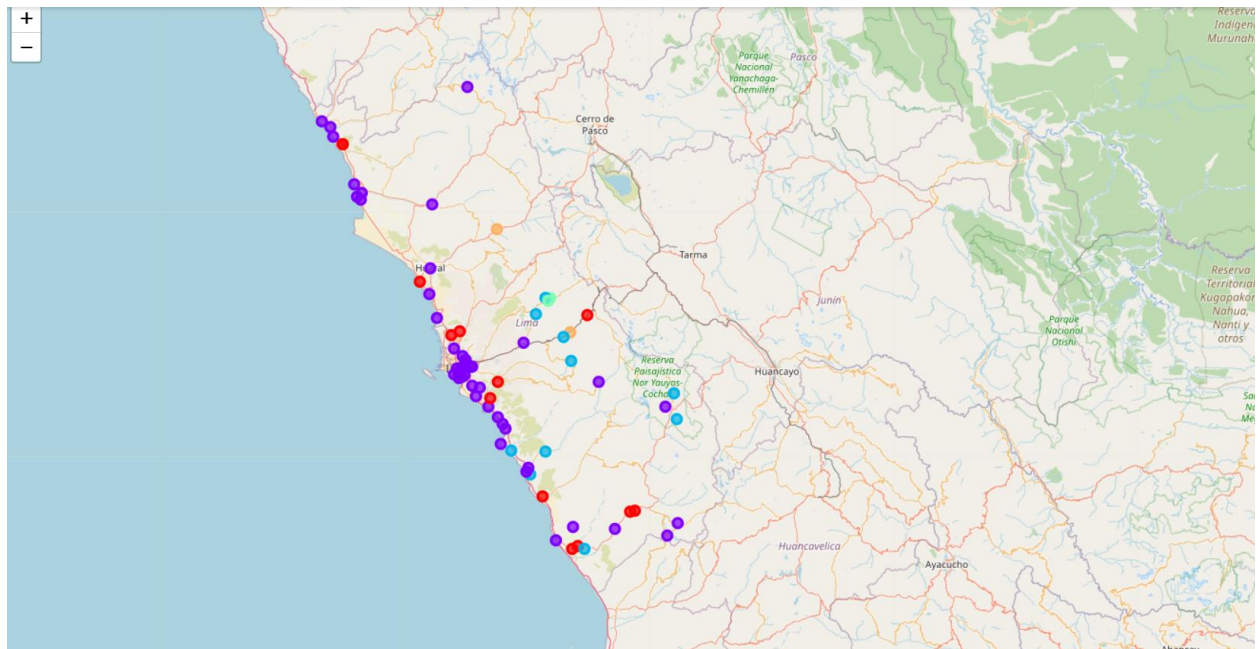
Fig. 15 - Clusters of districts (Lima region)

Cluster 1:

	District	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	9th Most Common Venue	10th Most Common Venue
1283	Carabayllo	0.0	Peruvian Restaurant	Restaurant	Latin American Restaurant	Yoga Studio	Falafel Restaurant	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
1286	Cieneguilla	0.0	Historic Site	Stables	Peruvian Restaurant	Yoga Studio	Exhibit	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
1300	Pachacámac	0.0	Plaza	Peruvian Restaurant	Italian Restaurant	Breakfast Spot	City	Yoga Studio	Falafel Restaurant	Food	Flea Market	Fish & Chips Shop
1302	Puente Piedra	0.0	Peruvian Restaurant	Yoga Studio	Falafel Restaurant	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm
1324	Supe	0.0	History Museum	Peruvian Restaurant	Shopping Mall	Plaza	Falafel Restaurant	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
1325	Supe puerto	0.0	History Museum	Peruvian Restaurant	Shopping Mall	Plaza	Falafel Restaurant	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
1338	San Vicente de Cañete	0.0	Peruvian Restaurant	Bus Station	Market	Restaurant	City	Plaza	Gym / Fitness Center	Donut Shop	Fast Food Restaurant	Farmers Market
1339	Asia	0.0	City	Peruvian Restaurant	Exhibit	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm
1344	Imperial	0.0	Hotel	Bus Station	Peruvian Restaurant	Plaza	Health & Beauty Service	Electronics Store	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
1348	Pacaran	0.0	Peruvian Restaurant	Latin American Restaurant	Yoga Studio	Falafel Restaurant	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
1353	Zuñiga	0.0	Plaza	Peruvian Restaurant	Historic Site	Exhibit	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm
1358	Chancay	0.0	American Restaurant	History Museum	Flea Market	Peruvian Restaurant	Yoga Studio	Falafel Restaurant	Food & Drink Shop	Food	Fish & Chips Shop	Fast Food Restaurant
1387	San Mateo	0.0	Accessories Store	Peruvian Restaurant	Department Store	Motel	Plaza	Electronics Store	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market

Cluster 2:

	District	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	9th Most Common Venue	10th Most Common Venue
1278	Lima	1.0	Chinese Restaurant	Seafood Restaurant	Bakery	Furniture / Home Store	Restaurant	Market	College Academic Building	Dim Sum Restaurant	History Museum	Park
1279	Ancón	1.0	Seafood Restaurant	Beach	Ice Cream Shop	Dessert Shop	Yoga Studio	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant
1280	Ate	1.0	Fast Food Restaurant	Peruvian Restaurant	Chinese Restaurant	Breakfast Spot	Soccer Field	Seafood Restaurant	Coffee Shop	Burger Joint	BBQ Joint	Park
1287	Comas	1.0	Bus Station	Cupcake Shop	Soccer Field	Snack Place	Restaurant	Seafood Restaurant	Falafel Restaurant	Flea Market	Fish & Chips Shop	Fast Food Restaurant
1288	El Agustino	1.0	Fried Chicken Joint	Bus Station	Plaza	Cemetery	Big Box Store	Farmers Market	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop
1289	Independencia	1.0	Home Service	Furniture / Home Store	Falafel Restaurant	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm
1293	Lince	1.0	Seafood Restaurant	Peruvian Restaurant	Park	Bakery	Restaurant	Italian Restaurant	Café	Burger Joint	Latin American Restaurant	Market
1295	Lurigancho	1.0	Fried Chicken Joint	Restaurant	Chinese Restaurant	Gym	Grocery Store	Park	Seafood Restaurant	Climbing Gym	Fast Food Restaurant	Juice Bar
1296	Lurín	1.0	Plaza	Shopping Mall	Market	BBQ Joint	Grocery Store	Supermarket	Juice Bar	Falafel Restaurant	Flea Market	Fish & Chips Shop
1297	Magdalena del Mar	1.0	Café	Italian Restaurant	Peruvian Restaurant	Coffee Shop	Seafood Restaurant	Fried Chicken Joint	Pizza Place	Restaurant	Cupcake Shop	Burger Joint
1299	Miraflores	1.0	Park	Seafood Restaurant	Café	Coffee Shop	Restaurant	Italian Restaurant	Supermarket	Chinese Restaurant	Peruvian Restaurant	Hostel
1301	Pucusana	1.0	Beach	Park	Cajun / Creole Restaurant	Italian Restaurant	Harbor / Marina	Falafel Restaurant	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant
1303	Punta Hermosa	1.0	Beach	Ice Cream Shop	Juice Bar	Scenic Lookout	Beer Garden	Yoga Studio	Farm	Food & Drink Shop	Food	Flea Market

Cluster 3:

	District	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	9th Most Common Venue	10th Most Common Venue
1340	Calango	2.0	Plaza	Seafood Restaurant	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Exhibit
1342	Chilca	2.0	Plaza	Asian Restaurant	Beach	Falafel Restaurant	Food Court	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant
1346	Mala	2.0	Plaza	Recreation Center	Restaurant	Electronics Store	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm
1347	Nuevo imperial	2.0	Plaza	Ice Cream Shop	Campground	Exhibit	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm
1377	Laraos	2.0	Plaza	Mountain	Cupcake Shop	Department Store	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
1383	San Damián	2.0	Plaza	Food Truck	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
1389	San Pedro de Casta	2.0	Campground	Yoga Studio	Food Truck	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm
1397	Surco	2.0	Home Service	Plaza	Mountain	Cupcake Shop	Department Store	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant
1429	Huancaya	2.0	Chocolate Shop	Campground	Yoga Studio	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm
1433	Laraos	2.0	Plaza	Mountain	Cupcake Shop	Department Store	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market

Cluster 4:

	District	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	9th Most Common Venue	10th Most Common Venue
1369	Carampoma	3.0	Mountain	Food Truck	Food Court	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm
1384	San Juan de Iris	3.0	Mountain	Food Truck	Food Court	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Farm

Cluster 5:

	District	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	9th Most Common Venue	10th Most Common Venue
1362	San Miguel de acos	4.0	Latin American Restaurant	Yoga Studio	Farm	Food Court	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
1366	Matucana	4.0	Motel	Latin American Restaurant	Farm	Food Court	Food & Drink Shop	Food	Flea Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market

5. Discussion

Before explaining observations based on the results, it is important to note that some of the provinces and districts have the same name, especially in Cusco region. This causes difficulties when retrieving not only coordinates for each district but also when exploring venues around popular spots, which was also influenced by the venue name containing the district or province name.

When exploring the popular spots' districts in Cusco, we can observe that the most common categories in the 4 considered districts are hotels and restaurants, but also zoos, train stations and trails which can be recommended when visiting the already well-known spots. If we look at the rating of the venues in each category, we can get specific places to be recommended to tourists in each selected district. On the other hand, when exploring Lima region, we can see that most common categories near popular spots are hotel and restaurants too, but the cuisine is more diverse. Additional to those categories, parks are also common and can be of interest for tourists. However, it is important to mention that there was a limitation about the possible requests to the Foursquare API to get the venue details such as the rating.

Talking about the segmentation, in case of Cusco we analyzed 7 clusters. The first cluster is composed of 7 districts (Ccorca, Acos, Pucyura, Tinta, Paruro, Huanoquite, Paucartambo) which have a lot of plazas, zoo, and historic sites. The second cluster where 10 districts are (Poroy, Mollepata, Coya, Maranganí, Pitumarca, Andahuaylillas, Huaro, Oropesa, Maras, Yucay) has a diversity of restaurants, zoo, and historic sites and museums. The third cluster has only the districts of Limatambo and San Pedro where mostly history museums and historic sites are. The clusters 4, 5 and 6, including Acomayo, Acopia, Lucre, Huancarani districts, are areas for people interested in nature since have parks, lakes, and zoos. Finally, the cluster 7 has 19 districts where hotels and restaurants are the most common venues.

In case of segmentation in Lima, we used 5 clusters. The cluster 1, where Carabayllo, Cieneguilla, Pachacamac, Puente Piedra, Supe, Supe puerto, San Vicente de Canete, Asia, Imperial, Pacaran, Zuniga, Chancay and San Matero are, has diverse venue categories such as restaurants, historic sites, and stores. The cluster 2 includes more than 40 districts that concentrate restaurants of diverse cuisine but also beaches, parks, and hotels. The cluster 3 (Calango, Chilca, Mala, Nuevo Imperial, Laraos, San Damian, San Pedro de Casta, Surco, Huancaya) has mostly plazas. Finally, for tourists interested in climbing or farming, the clusters 4 and 5 could be good options since they included mountains and farms, aside of restaurants.

6. Conclusion

During the project it was possible to explore already popular places in Cusco and Lima region, making possible to recommend new places to future tourists. These new places in Cusco include:

Venue	Venue Category	District
Sumaq Machu Picchu Hotel	Hotel	Machupicchu
Tierra Viva Cusco Machu Picchu	Hotel	Machupicchu
El Albergue de Ollantaytambo	Hotel	Ollantaytambo
Hotel Pakaritampu	Hotel	Ollantaytambo
Café Mullu	Cafe	Pisac
Ulkire's Café	Cafe	Pisac
Añañau	Restaurant	Pisac

In case of Lima, we got the next list:

Venue	Venue Category	District
Chifa Kong	Chinese Restaurant	Lima
Wa Lok	Chinese Restaurant	Lima
Chifa San Joy Lao	Chinese Restaurant	Lima
Parque Guillermo Correa Elias	Park	Miraflores
Parque Eduardo Villena	Park	Miraflores
Parque Federico Blume	Park	Miraflores
Delfines Hotel & Casino	Hotel	San Isidro
Country Club Lima Hotel	Hotel	San Isidro
Hyatt Centric Hotel	Hotel	San Isidro

Additionally, the segmentation shows that it is possible to create clusters of districts in each region and create a profile for each one and in the future those profiles could be matched with tourists' interests, generating increased income not only for the tourist agencies but also for the government, especially when the visited spots are managed by them.

References:

[1] https://versus.com/en/cusco-vs-lima

[2] https://es.wikipedia.org/wiki/Anexo:Distritos_del_Per%C3%BA

[3] https://positionstack.com/

[4] https://www.mincetur.gob.pe/wp-content/uploads/documentos/turismo/estadisticas/ReporteTurismoRegional/RTR_Cusco.pdf

[5] https://www.mincetur.gob.pe/wp-content/uploads/documentos/turismo/estadisticas/ReporteTurismoRegional/RTR_Lima_2018.pdf