2. Data

To solve the problem defined in section 1, we will only analyze most popular neighborhoods, these places correspond to Lima Center and is composed by 14 neighborhoods whose information (names, latitude, longitude, etc.) will be gathered from Wikipedia. The data downloaded contain 43 neighborhoods so we will clear the data to only have the 14 neighborhoods to analyze. From our data the most popular neighborhoods are "Miraflores" and "San Isidro", the first one because of their venues like bars and restaurants and for the presence of many tourist, the second one is better known as a business neighborhood but also with many well-known restaurants.

We will use Geopy as a geocoding tool to get latitude and longitude for each selected neighborhood. One problem with this service was that "Miraflores" neighborhood was located far from the center causing the data collected not cover places where many restaurants are, to fix this we will move the location of "Miraflores" closer to center. The venues of each neighborhoods will be gathered using Foursquare API with a limit of 100 venues by neighborhood and a radius of 2 km. Since the distance between each neighborhood are not constant is inevitable that many venues will be duplicated, so after the clean of duplicated venues we have 903 venues in total and 170 unique categories.

Table 1: Neighborhoods venues

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	San Miguel	-12.078656	-77.095283	Do it!	-12.081012	-77.082683	Accessories Store
1	Lima	-12.062107	-77.036526	Tony Roma's	-12.057406	-77.030081	American Restaurant
2	Lima	-12.062107	-77.036526	Chili's	-12.066808	-77.047318	American Restaurant
3	San Miguel	-12.078656	-77.095283	T.G.I. Fridays	-12.076010	-77.084234	American Restaurant
4	San Miguel	-12.078656	-77.095283	T.G.I. Friday's	-12.078503	-77.087624	American Restaurant

Table 2: Neighborhoods information

	Neighborhood	Latitude	Longitude
0	Lima	-12.062107	-77.036526
1	Barranco	-12.143959	-77.020268
2	Breña	-12.059700	-77.050119
3	Jesús María	-12.078186	-77.046412
4	La Victoria	-12.073358	-77.016417
5	Lince	-12.086568	-77.036647
6	Magdalena del Mar	-12.095654	-77.068213
7	Pueblo Libre	-12.076639	-77.067858
8	Miraflores	-12.118814	-77.040097
9	San Borja	-12.096452	-76.995690
10	San Isidro	-12.097902	-77.035367
11	San Miguel	-12.078656	-77.095283
12	Santiago de Surco	-12.125105	-76.981919
13	Surquillo	-12.114198	-77.010475