**IBM Capstone Project - The Battle of Neighborhoods**

**The Selection of Neighborhoods for Staying in a suitable location in Madrid**

Applying Data Science using Python & Machine learning to identify the suitable place within surrounding areas

By Debesh Roy

17Aug2020

1. Introduction

Due to the present era of Globalization, people are moving around the globe, knowing or without knowing much about the destination and surrounding neighborhoods, where they are planning to stay for some time – months to years.

1. Background

Madrid is the capital and most populous city of Spain. The city has almost 3.3 million inhabitants and a metropolitan area population of approximately 6.5 million. It is the second-largest city in the European Union (EU), surpassed only by Berlin, and its monocentric metropolitan area is the second-largest in the EU, surpassed only by Paris. The municipality covers 604.3 km2 (233.3 sq mi).

Madrid is a city of elegant boulevards and expansive, manicured parks and many more. It’s renowned for its rich repositories of European art, including the Prado Museum’s works by Goya, Velázquez and other Spanish masters. The heart of old Hapsburg Madrid is the portico-lined Plaza Mayor, and nearby is the baroque Royal Palace and Armory, displaying historic weaponry.

In this project we will try to find a suitable locations for staying, who is planning to relocate to Madrid for some time, based on his/her personal preferences of the surrounding neighborhoods and available venues.

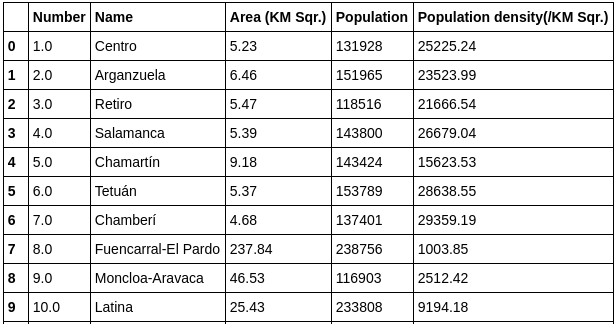
We will use the data science mechanism to generate more information about the nearby venues and whereabouts, which will help stakeholders or interested people to take the decision and plan accordingly.

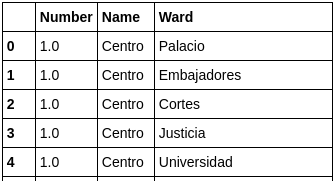
1. Data

The related information about Madrid got downloaded from Wikipedia, i.e. <https://en.wikipedia.org/wiki/Districts_of_Madrid>

We extracted info like District, Area, Population, Density, corresponding Wards from the Wiki, and definitely the Venue & Category details using Foursquare URL [https://foursquare.com](https://foursquare.com/) and corresponding APIs.

* 1. District list



* 1. District wise ward count
  2. **District wise ward list**

Also using the “geopy” library and got the Latitude & Longitude for all districts and wards.

1. Methodology

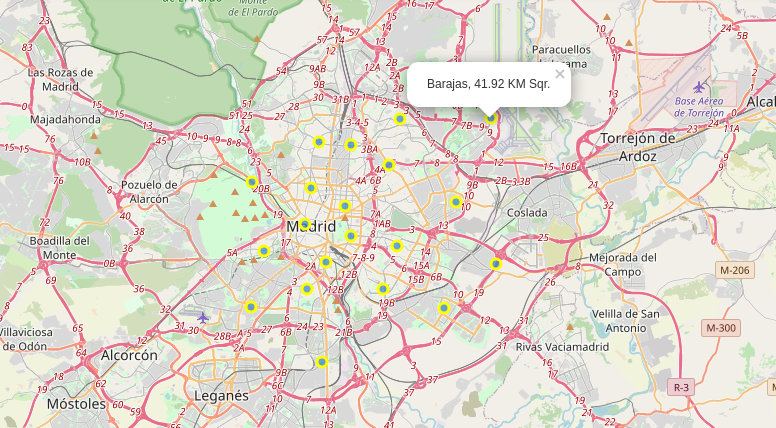
The base of the project is Python, where we extensively used multiple important libraries, i.e. pandas, numpy, geocoder, folium, and many more. Each one of the used library is having dedicated purpose.

Used pandas’ “read\_html” to read the information from wiki URL. Used DataFrame feature extensively to handle data, i.e. transforming, cleansing, etc.

This data handling helps to identify the District wise Ward wise Venue and the Category.

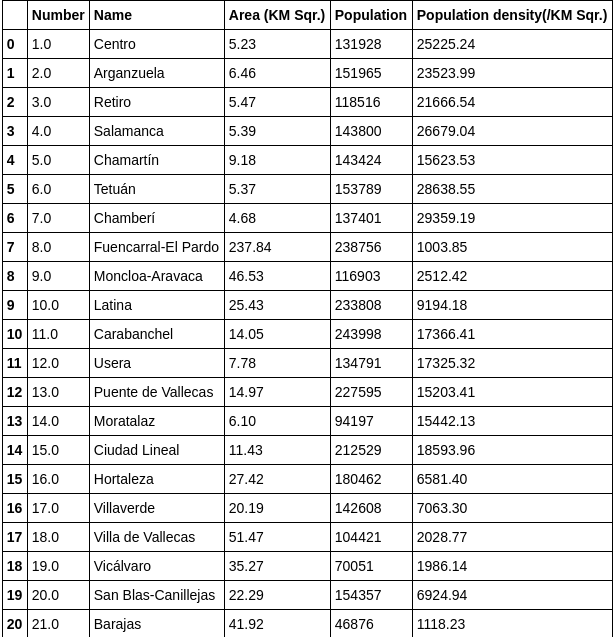


Used Florium map feature, which will help interested stakeholder to identify the surrounding venues, categories, neighborhoods, etc.

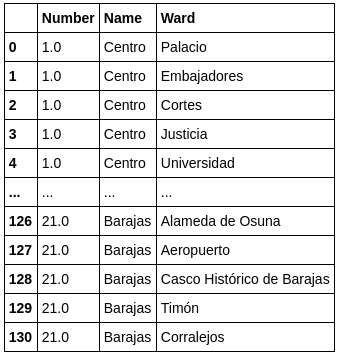


5. Results

Using the Python and corresponding libraries, we identified that there are 21 districts in Madrid, along with the corresponding area, population and density.



 We also identified that each district is having on an average 6~7 wards.

 And the corresponding wards for each district

6. Discussion

Well, the purpose of this Python notebook is to identify the suitable location to stay. The tool will not suggest the Good or Bad neighbor, but based on personal interest, whether you are looking for School or College, trade off between Shopping Mall and Supermarket, preference between Candy Store and Ice Cream Shop, looking for Restaurant or Bar, and many more.

Along with the data point, from the Florium map one can also identify the near around venues and approx. distances.

7. Conclusion

It’s always preferred to have a careful consideration before selecting a neighborhood, where one is going to stay for some time. Though this report will not decide the best place but it’s up to individual choice to finalize.

After going through the detailed analysis between districts, wards and areas, once can easily get a fair knowledge on the surroundings and can conclude. Based on that, s/he can either put on advertise on the local media, or just can pay a visit to feel the ambience of the place.