

Final Project IBM Data Science Capstone:

By: Mitadru Ghosh

Title: Construction of a new Walmart, the best location for a sprawling city

Introduction:

In this work we are going to discuss where should be the optimal location within my city (León, Gto, México) to construct another Walmart Supermarket, this will be done by analysing the different areas and the presence of another supermarkets and considering that the city is under expansion, where a lot of new zones are emerging. A new location for a supermarket also has to be a very busy area where people from all walks of life benefit from the supermarket

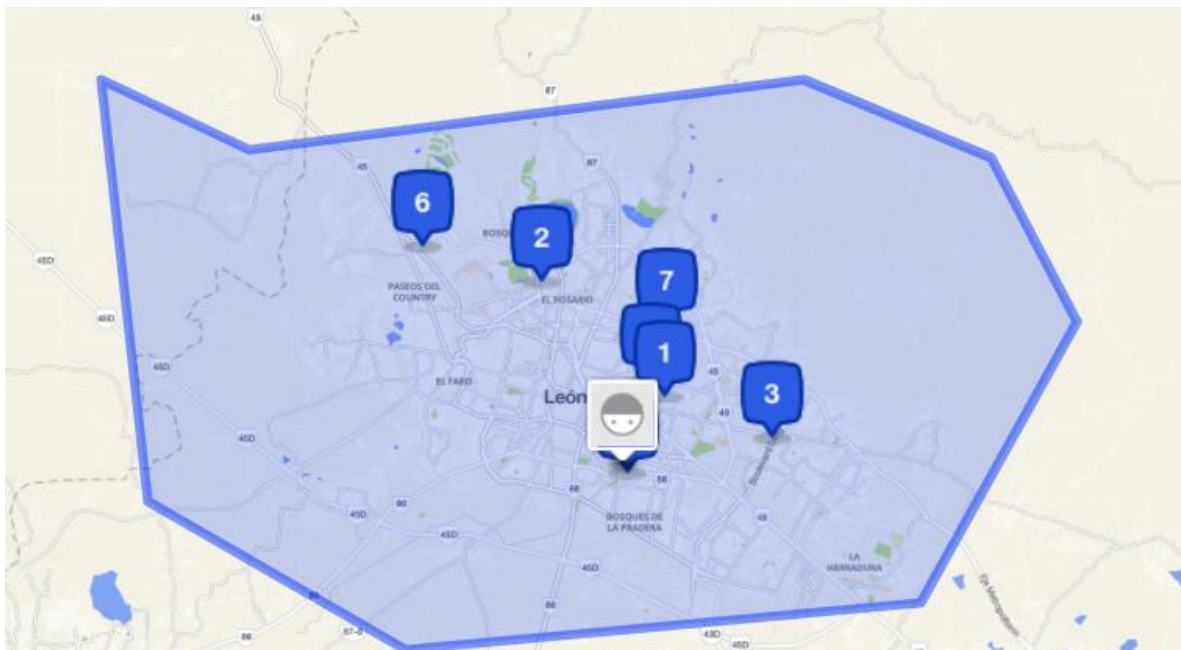


Fig 1. Locations of walmart supermarkets

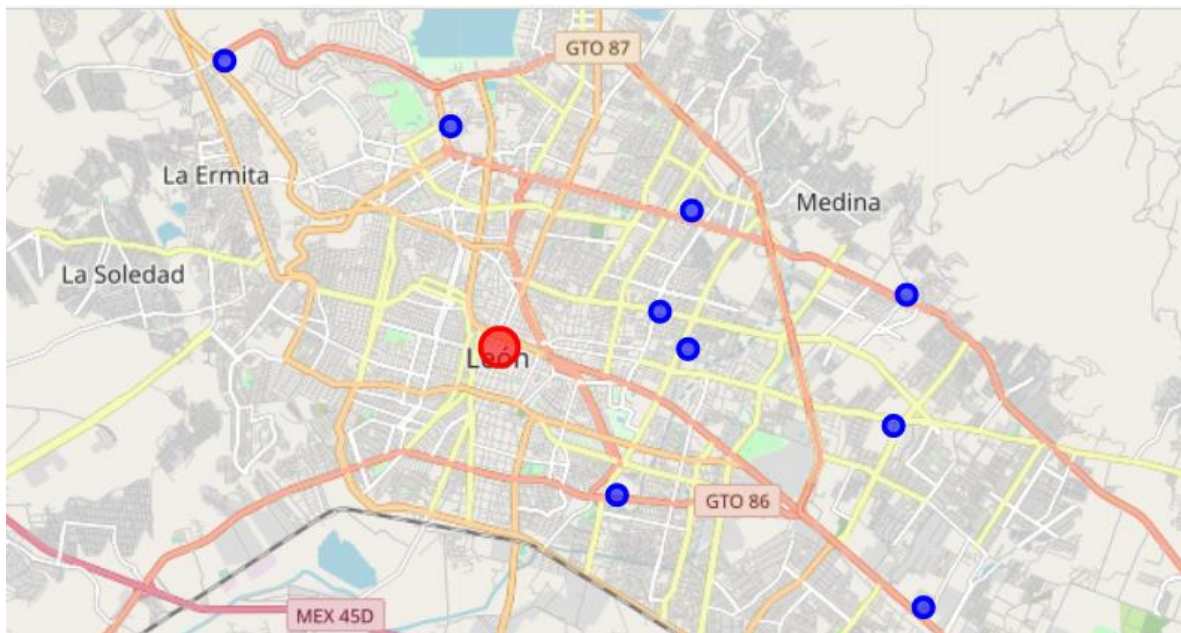
Data Section:

The data to be used will be the locations of the current walmart supercenters, and we will also include the location of the main (and biggest) competitors in my city as Soriana, Bodega Aurrera, Chedraui, etc.

We will also review data from my city to know the number of inhabitants per locality and the main avenues

Methodology:

To know where the current supermarkets are localized we have searched using the Foursquare API and the coordinates of my city (León, Gto). In the Jupyter Notebook we have added the multiple locations and filtered by using knowledge of all locations



*Fig2. Locations of all **walmart** supermarkets by using Folium*

Distribution of Soriana Supermarkets:

The main competitors in my city are Soriana Supermarkets, let's see the distribution of this supermarkets using Folium:



Fig.3 Locations of all Soriana Supermarkets

If we compare the zones, there are no Walmart Supermarkets in the southwest zone, this gives us an insight of the WHERE we can locate our new supermarket. All of this has been done by using python and Foursquare, not Google Maps.

Let's compare with the other main supermarkets in León: Bodega Aurrera

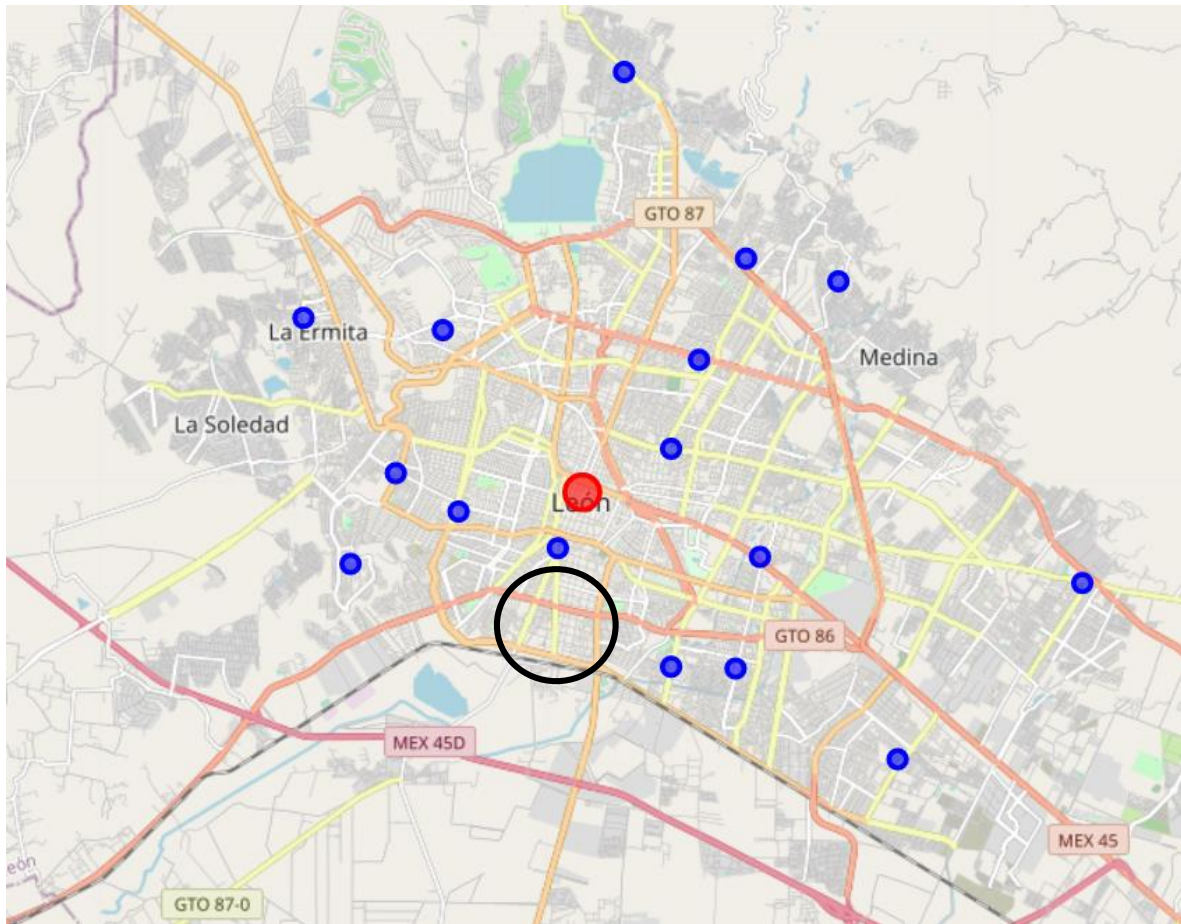


Fig.4. Locations of all Bodega Aurrera Supermarkets

Now we analyse the locations of the Bodega Aurrera locations, we can see now that bodega Aurrera is well distributed in my city. The black circle indicates the most profitable place in terms of location, but it is really the best location?

Results

Let's see all supermarkets at the same time

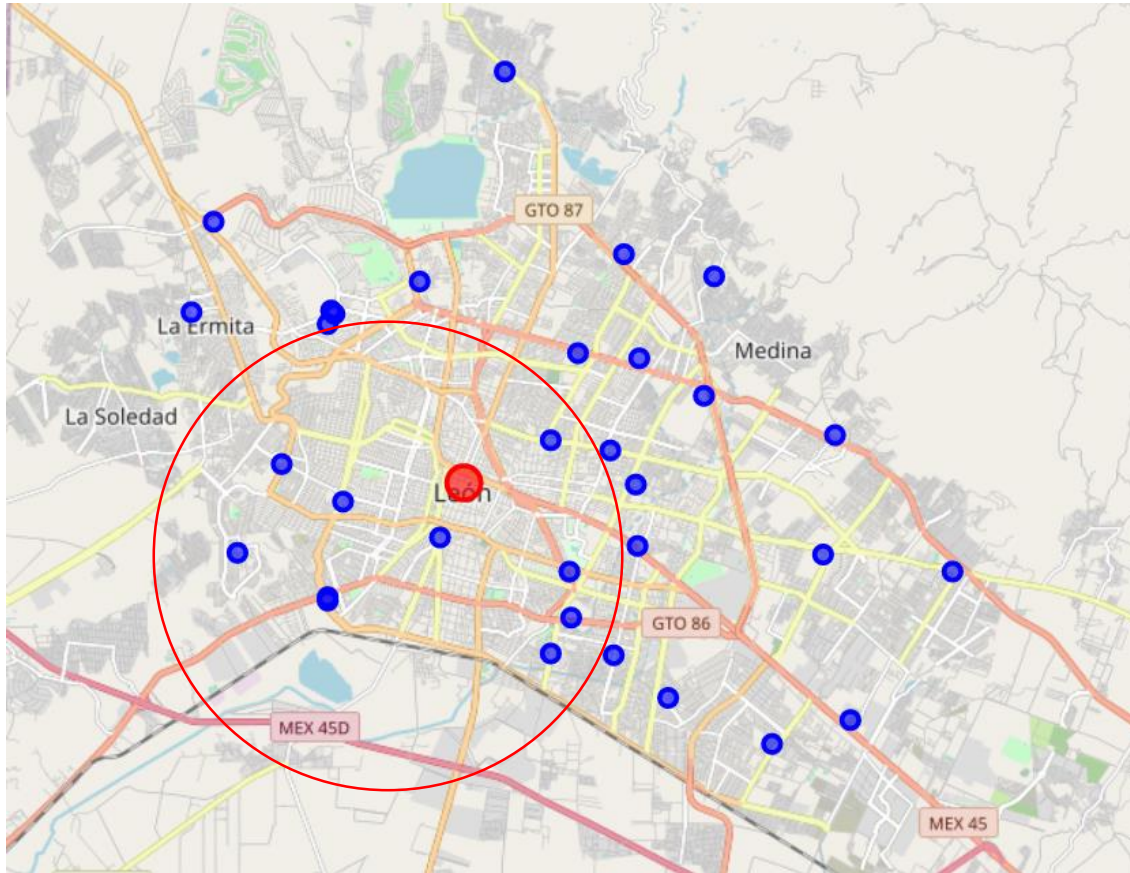


Fig5. All supermarkets in my city

Now we can see that Walmart is not competing in the red zone and the density of supermarkets is also lower in this zone. In this case it could be the best option to place a Walmart supercenter. This southwest zone could be definitively the best zone by looking at the cluster closed in red for the supermarkets in southwest zone.

Discussion

It is important to notice that we also need to include the number of inhabitants in this zone, and to know about another parameters to decide the right location, like availability, profitability among others.

Conclusions

It is important to know the importance of the WHERE statement in a single problem or a machine learning problem because to make a good decision making, for example about a new building or a new retail company or a even a new store we have to analyse all the options and the competitors around the city.