Recommending System for Opening Businesses in Bogota, Colombia

1. Introduction

1.1. Background

Due to the Covid Pandemic that rapidly spread all over the world in the year 2020, various restrictions were put in place in order to slow down the spreading of the virus. Bogota, Colombia was one of the cities that entered into a lockdown in March 2020, and due to this, various businesses had to shut down because of the lack of clients and people began losing their jobs. Now that some restrictions are being lifted in 2021, people are considering reopening their businesses or even creating new ones.

1.2. Problem

In order to open a business, people would be interested in a solution that helps them determine the best location by taking into account multiple factors such as: Trending venues in the localities, information on similar localities access to outdoor spaces (a plus in order to maintain the covid protocols established by the government) and covid cases in the locality. These are factors that must be taken into account at the moment because a locality with a peak in covid cases and a lack of covid protocol could be sent into lockdown again and impact the new business.

1.3. Possible Stakeholders

This project or recommendation system would be of interest to those persons who want to open a business and want to determine which are some of the best location options taking into account the factors mentioned previously.

2. Data

2.1. Bogota Localities

In order to get information regarding each locality in Bogota, the names and coordinates of each locality are needed. We can find this and more information in this website provided by Opendatasoft. Since there is additional information from the one that is needed and the location information is contained within a table, we will need to perform web scraping on the site and some data cleansing.

We can visualize an example of how the website supplies the data on the next figure:

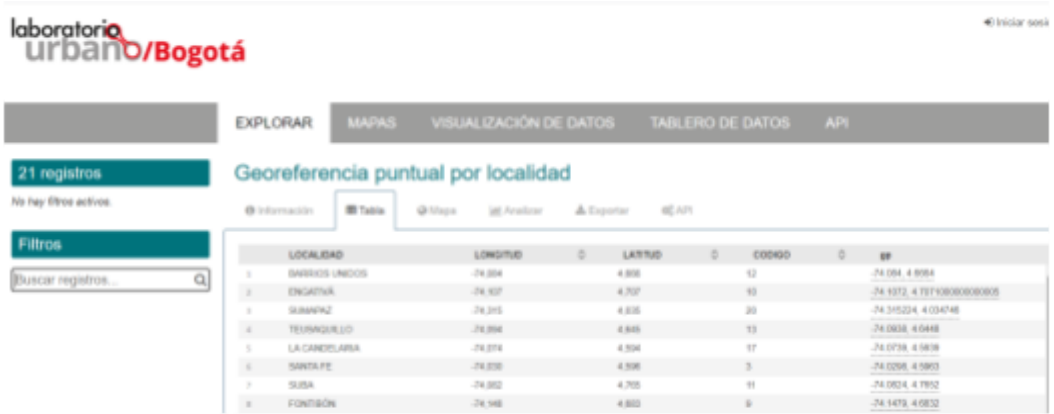

Figure 1. Preview of website with Bogota localities data

2.2. FourSquare API

By having the location information of Bogota's localities, the FourSquare API can be used to get information on the venues located within these localities. Depending on the kind of request sent to the API more information on the venues can be returned. For the recommender system in the project the information which could be considered valuable is: Name, Location, Category, Rating.

	name	categories	lat	lng
0	Roselle Desserts	Bakery	43.653447	-79.362017
1	Tandem Coffee	Coffee Shop	43.653559	-79.361809
2	Cooper Koo Family YMCA	Distribution Center	43.653249	-79.358008
3	Body Blitz Spa East	Spa	43.654735	-79.359874
4	Impact Kitchen	Restaurant	43.656369	-79.356980

Figure 2. Foursquare API data example

2.3. Bogota Covid Cases by Localities

In order to determine which localities have a high probability of going into lockdown again or to implement stronger restrictions, it is necessary to have recent data on the amount of covid cases in the localities in Bogota. This information can also be found in different Municipal Government websites, from which we'll be using this one. On this website a database of covid cases in Bogota can be downloaded and used for the recommender system. This database includes the following information: Case number, Date of the initial symptoms, Date of the diagnosis, City, Locality, Age, Gender, Contagion source, State of the patient. This database has over 1.2M records.

_id	CASO	FECHA_...	FECHA_...	CIUDAD	LOCALI...	EDAD	UNI_MED	SEXO	FUENTE...	UBICAC...	ESTADO
292359	292359	2020-10-...	2020-10-...	Bogotá	Engativá	43	1	F	En estudio	Casa	Recuper...
292360	292360	2020-10-...	2020-10-...	Bogotá	Suba	41	1	F	En estudio	Casa	Recuper...
292361	292361	2020-10-...	2020-10-...	Fuera de...	Fuera de...	41	1	F	En estudio	Casa	Recuper...
292362	292362	2020-09-...	2020-10-...	Bogotá	Suba	40	1	F	En estudio	Casa	Recuper...
292363	292363	2020-10-...	2020-10-...	Bogotá	Suba	40	1	F	En estudio	Casa	Recuper...
292364	292364	2020-10-11	2020-10-...	Bogotá	Engativá	40	1	F	Relacion...	Casa	Recuper...
292365	292365	2020-09-...	2020-10-...	Bogotá	Suba	45	1	F	Relacion...	Casa	Recuper...
292366	292366	2020-10-11	2020-10-...	Bogotá	Usaquén	45	1	F	Relacion...	Casa	Recuper...

Figure 3. Bogota Covid cases database preview