Analysis of the comorbidities associated with the COVID-19 pandemic in Mexico and hospital availability in CDMX

José Frías

July 8, 2020

1 Description of the problem

1.1 Background

The COVID-19 pandemic, caused by SARS-CoV-2 virus, has affected the lives of many people around the world. Governments, organizations and general population are concerned about the evolution of the pandemic and any other information that may be useful in reducing the affectations in their personal and community lives. Governments around the world have taken different preventive and corrective measures to reduce the impact of the pandemic in their societies according to their particular conditions. Among the wide variety of studies concerned with the COVID-19 pandemic, those which analyze the comorbidities associated to the main disease are very useful in protecting vulnerable populations. Another important matter to take into account in order to improve the COVID-19 patients care is the hospital capacity of any city or region under consideration.

1.2 Problem

In the present project we propose the use of public data to analyze two important aspects of the COVID-19 pandemic in Mexico: first, the analysis of the main comorbidities associated to COVID-19 in all the country and, secondly, an analysis of the hospital infrastructure available in Mexico City (CDMX).

With data on the main comorbidities related to COVID-19 disease in Mexico, we aim to use clustering techniques to bunch the 32 states in the country into groups characterized by the affectations of the different comorbidities among their populations. This analysis may be very useful for the federal and local governments to take preventive policies to reduce the risk of health complications among vulnerable populations. The information we will obtain could also be useful for individuals with one or more comorbidities to take their own precautions.

In the second part of this analysis, we aim to use public data and Foursquare to analyze the hospital infrastructure in the 16 municipalities in CDMX. The distribution of the health institutions in Mexico City is not homogeneous and depends on socioeconomic variables. The same is true for the COVID-19 pandemic spread. We aim to compare the total number of COVID-19-positive cases in each municipality in Mexico City with the estimated number of hospitals reported by Foursquare in a representative sector of this municipality. The results of this analysis may be of interest to the city government for the development of strategies in the care of the infected population.

2 Description of data

The first part of the project is based on a dataset named *'casos-asociados-a-covid-19.csv'* that can be found at *https://www.gob.mx/salud/documentos/datos-abiertos-152127* and was updated on June 23, 2020. This dataset contains 384,283 records of persons tested for COVID-19 in Mexico and includes 40 fields containing relevant information for the pandemic statistics. In our analysis we focus on the comorbidities for COVID-19, which constitute nine fields in the dataset, and the final status of each tested person (positive test, hospitalized and dead). Our aim is to use this dataset to analyze the prevalence of the comorbilities among the populations of the Mexican states and cluster them according to this criteria.

For the second part of the project we intend to compare the number of positive cases for COVID-19 in the municipalities of CDMX to the population and the number of hospitals by municipality. The data of coronavirus disease in Mexico City that we will analyze can be found at *https://datos.cdmx.gob.mx/* and was updated on July 1, 2020. This dataset contains 155,722 records of persons tested for COVID-19 in Mexico City and several fields, among which we shall focus on the municipality of origin and status of the records. We also get demographic data of Mexico City municipalities from *https://en.wikipedia.org/wiki/Municipalities_of_Mexico_City*. Finally, we shall use Foursquare to find hospitals and medical clinics in a radius of 2 km around the center of each municipality using the search parameter *categoryId = '4bf58dd8d48988d196941735'*, which corresponds to the *"hospital"* venue category.

For both parts of this project we aim to use the *geocoder* library to find the geographical coordinates of the Mexican states and the municipalities in Mexico City to visualize the results we obtain in our analysis by using the *folium* library.