Business problem

The COVID-19 pandemic has hit the large urban areas in the U.S. the hardest. New York has had the largest outbreak in the country, although the number of new cases has declined over the recent weeks. Detroit's Wayne County and Chicago's Cook County are not far behind the epicenter in the number of coronavirus-related deaths.

San Francisco is the second most densely populated large U.S. city, behind New York City, and the fifth most densely populated U.S. county, behind only four of the five New York City boroughs. High population density is one of the factors that could potentially make San Francisco one of the U.S. cities that are most vulnerable to the novel coronavirus.

The objective of this project is to provide a helicopter view on the COVID-19 situation in San Francisco by clustering neighborhoods based on their capacity to fight the pandemic. The results of the analysis could be of interest to people living in San Francisco area. The analysis will use a clustering algorithm that takes into account four variables for each neighborhood: number of hospitals per 1000 people (fetched from Foursquare API), number of hospital beds per 1000 people, number of intensive care unit (ICU) hospital beds per 1000 people, and the population density.

Data

I will collect data from several sources:

List of 41 San Francisco neighborhoods created by grouping 2010 Census tracts:

- Data content: Neighborhood, Geographic coordinates (multipolygon)
- Data format: GeoJSON
- Data source: https://data.sfgov.org/Geographic-Locations-and-Boundaries/Analysis-Neighborhoods/p5b7-5n3h

List of hospitals in San Francisco:

- Data content: Name, Category, Geographic coordinates (point)
- Data format: JSON
- Data source: Foursquare API

USA Hospital Beds:

- Data content: Hospital name, Geographic coordinates (point), Hospital type, Address, Number of licensed beds, Number of staffed beds, Number of ICU beds
- Data format: CSV
- Data source: https://coronavirus-resources.esri.com/datasets/1044bb19da8d4dbfb6a96eb1b4ebf629_0/data?geometry=83.672%2C-16.820%2C-108.633%2C72.123

San Francisco population data:

- Data content: Neighborhood, Population as of 2016
- Data format: PDF
- Data source: https://sfgov.org/sfplanningarchive/article/socio-economic-profiles-2016-five-year-acs-released