Capstone Project –The Battle of the Neighborhoods

for Applied Data Science Capstone Course

Karl D Huber
April 2020

DATA

Here is the key data needed to complete this project:

- population by Southern California (SoCal) zip code downloaded from **US Census using their public API**
- SoCal hospital name and latitude/longitude gathered using **Foursquare API**
- SoCal hospital name, city and number of beds scraped from **American Hospital Directory website** www.amd.com
- COVID-19 overall hospitalization rate calculated from data downloaded from **NYC Health website** https://www1.nyc.gov/site/doh/covid/covid-19-data.page (New York City is similar in size to the SoCal region and has the most COVID data)
- geocoordinates from **geopy and Nominatum** for Los Angeles, CA as an appropriate centerpoint to map SoCal using **Folium**

NOTE: zip code was chosen as the core measure (rather than city or county) to provide the most granular information. This will allow Los Angeles to be split into its far-ranging geographic communities (e.g. Century City & Westwood & Downtown Los Angeles) rather than just lumping all LA population data into one wide spread blob called Los Angeles.

Will also need this supplemental data to aid presentation and ease comprehension:
cross-reference of SoCal zipcodes to their corresponding city using scraped & filtered data from https://www.laalmanac.com/communications/cm02_communities.php

Here's how the data will be used once gathered:

1. Calculate population and number of beds for each city/community based on zip codes
2. Convert population to max # of beds needed for each city/community using the overall hospitalization rate from NYC
3. Calculate the 'bed supply ratio' for each city/community = max # of beds/number of beds (a value < 1 indicates there are enough hospital beds to cover the max # of beds needed; a value > 1 indicates a possible shortage of beds)
4. Assign each zip code in a city or community the 'bed supply ratio' for that city or community
5. Plot the 'bed supply ratio' for each zip code on a thematic map of SoCal visually highlighting the high risk areas
6. Plot each hospital on that SoCal map with its name, city and number of beds appearing when clicked
7. Tabulate the 10 most and 10 least at risk communities