Capstone Project –The Battle of the Neighborhoods

for Applied Data Science Capstone Course

Albert K

April 22nd 2020

DATA

Here is the key data needed to complete this project:

- population by Korea city information downloaded from **worldpopulationreview.com**
- Different hospitals located in Korea name and latitude/longitude gathered using **Foursquare API**
- Different hospitals located in Korea hospital name, city and number of beds, test center and others scraped from **Korean Hospital Directory website** www.amc.Korea.co.kr
- COVID-19 overall hospitalization rate calculated from data downloaded from **Korea Government healthcare website** http://ncov.mohw.go.kr/en/ (This website gives you access to all the detail information happening in Korea.

NOTE: City name was chosen as the core measure (rather than city or county) to provide the most granular information. This will allow kyung gi do district to be split into its far-ranging geographic communities (e.g. Seoul, Busan, Hangdang) rather than just lumping all Seoul districts population data into one wide spread blob called Seoul.

Will also need this supplemental data to aid presentation and ease comprehension: cross-reference of Korea's cities name to their corresponding city using scraped & filtered data from

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

1. Calculate population and number of beds for each city
2. Convert population to max # of beds needed for each city/community using the overall hospitalization rate from Korea national report
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, test centers' for that city or community
5. Plot the 'bed supply and test centers ratio' for each city on a thematic map of Korea visually highlighting the high risk areas
6. Plot each hospital on that Korea map with its name, city and number of beds, test centers and others appearing when clicked

7. Tabulate the 10 most and 10 least at risk communities