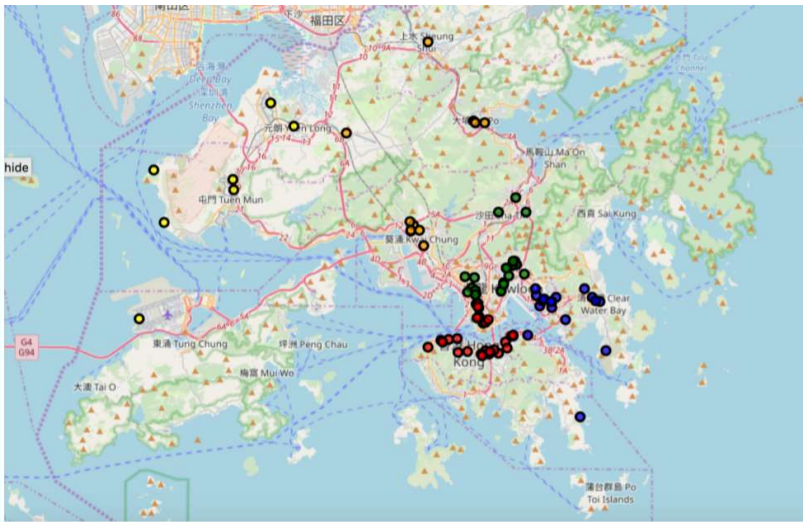
Final Presentation

Capstone Project - The Battle of Neighborhoods

Data Understanding

- The dataset was imported into a Notebook (Watson Studio is used to run the notebook) to understand the data.
- Since the dataset consists of geospatial location of the clinics, the latitudes and longitudes of the clinics are associated with the clinic names and addresses

Data Visualization



Recommendation Approach

- Each clinic data can be fitted into FourSquare location API to explore the nearby locations
- Radius of 1000 meters is used. 1 kilometer is usually within 10-minute walk which is generally acceptable
- From the consumer's interest, 5 categories are summarized as entertainment, hotels, restaurants, shops and supermarket
- A weighting of the categories to calculate the overall score is necessary because the primary concerns of a clinic location should be the nearby hotels

Results and Discussion

- The top five clinics are shown in the Folium Map next page
- Further looking into the locations, all these clinics are located at the center of the Hong Kong Island which are also the high dense population districts in Hong Kong
- The recommendation is generally good to match with the clustering result in the first data visualization of the clinics
- The weighting parameter can be easily adjusted to reveal a more custom scoring for recommendation

Recommendation Visualization

Conclusion

- The business problem of the influx of Mainland Chinese consumers into Hong Kong clinics will require a data analysis of the clinic locations with nearby interested necessities to provide recommendation.
- Top five clinic locations are suggested and all are matched with the Hong Kong situation that located in the center of Hong Kong Island