Project Presentation for Coursera Capstone Project

-The Battle of Neighborhoods

Jinseob Shim

Table of contents

Objective and Problem Statement

Seoul, officially known as the Seoul Special City, is the largest metropolis and capital of South Korea. This megacity is the largest city proper in the developed world and the Seoul Capital Area is the second largest metropolitan in the world with more than 25.6 million people, which is half of all the residents in the country. The estimated population within the city limits for 2016 is 10.29 million. Interestingly, Seoul's population density is almost twice that of New York City, four times higher than Los Angeles and eight times higher than the density of Rome.

People from around the world come to Seoul, Korea with various reasons to make the city more diverse and dynamic. The number of people coming to Seoul has been growing due to Korea’s remarkable economic growth and its cultural power so-called K-wave which includes k-drama, tv show, k-pop, and etc.,

Objective and Problem Statement

Therefore, there are a growing number of potential investors who want to open or operate a unique or trending venue like a restaurant in the city. They might wonder what kind of venue they should choose to invest or open in order to avoid unnecessary risks like getting into red ocean competition. With geospatial data and data science techniques, we can provide information about which type of venue is the most popular and clustered distribution of venues within Seoul to a potential investor. Based on this information, we can suggest a better business investment to the investor. And it helps he or she make an informed decision to invest in Seoul particularly by getting them to know about the geospatial distribution of data by districts related to the economic activity of people in Seoul.

-Target clients: potential clients looking for a business opportunity in Seoul but not sure because of insufficient knowledge about geographical economic features of the city.

Automation Script

Script Language	Python 3.x
Input	Open and private data by wikipedia and Shinhan Bank Google Maps Geocoding API Foursquare location data
Output	1. Economic cluster feature of districts of Seoul 2. Venue cluster feature of districts of Seoul 3. Recommended district to invest or do business 4. Recommended districts in Seoul on Map

Benefits

- Dynamic Recommendations according to socio-economic feature of clusters
- Recommended location plotted on map

Recommeded districts plotted on map

Summary of features of recommended districts

- The strongest buying power and need the most expensive investment in renting places=> "Cluster1(purple) "
- The second strongest buying power and need third most expensive in renting places=> "Cluster 2(light blue)"
- The third strongest buying power but only need 680 dollars for monthly rent cost which is the second cheapest among rent cost =>"Cluster3(light Mint)"
- The second weakest buying power; it is the very attractive since not only cheapest region it is but also have the fourth highest income level: potential growth in buying power is high =>"Cluster4 (light green)"
- The largest population and it has second strongest growth potential in buying power=>"Cluster 5(light ochroid)"
- Not Recommend => cluster 6(Red)

Limitations

- Although the project allows recognizing Seoul’s distinctive features by its districts through machine learning algorithms, it needs more research and factors such as consumption information or trending places and hours at street level should be put into in order to provide more reliable and specific information to customers
- Enable addtional condition checkig with an interactive User Interface.

Q&A