Recommendation of Hotel Place Selection in London

Zhen Wang

1. Introduction

   1.1. Background

   London, is a world-famous tourist city. It attracted 20.42 million international visitors in 2018 (Tourism in London, 2019). Tourism plays an important role in UK, tourism has been the fastest growing sector in employment terms in UK since 2010 (Britain's visitor economy facts, 2019). Britain is forecast to have a tourism industry worth over £257 billion by 2025 (Britain's visitor economy facts, 2019). The fast-growing tourism market means more tourists, which means more places are required for tourists to stay over the night.

   1.2. Problem

   This project will focus on finding an optimal location for a hotel. Specifically, since London is a large city, this report is targeted to stakeholders who are interested in opening a hotel in Inner London, UK.

2. Data Description and Cleansing

   2.1. Data Components

   Lee et al (2010) point out that in individual evaluation of hotel location model, three indexes mainly matter: safety, ease of access to transportation portals and close connection to area attractions (historic, businesses, and pleasure). Yang, Y et al (2012) point out that public service infrastructure, road accessibility, subway accessibility, hotel agglomeration effect and accessibility to tourism sites are important determinants of hotel location choice. Hotel agglomeration effect means that hotel can receive benefits when they cluster together (Yang et al, 2012). Aksoy and Ozbuk (2017) mention location-related attributes that influence travelers' impressions of hotel, such as traffic, security, general convenience, tourism attractions and parking. After reviewing several literatures, **Hotel agglomeration effect**, **Accessibility to transports**, **Tourist Attractions** are selected as the key factors for selecting a hotel location in this project.

   2.2. Data Sources

1

To be able to get three key analyzing factors, the websites were scraped as follows:

2.2.1. London Geography Related

1) London areas & postal code from **Wikipedia**

2) London areas & postal code location data from **Geocoder**

3) Inner London boroughs from **Wikipedia**

4) Download London shapefile from **London Datastore**. By using QGIS (a desktop geographic information system application), London shapefile can be transformed to GeoJson data.

2.2.2. Hotel Agglomeration Effect

Hotel location data will be obtained by using **Foursquare API**

2.2.3. Accessibility to Transports

Transports location (bus stop, bus station, light rail station, metro station, train station) data will be obtained by using **Foursquare API**

2.2.4. Tourist Attractions

Tourist attractions data will be obtained by using **TripAdvisor**

2.3. Data Cleansing

2.3.1. Foursquare Data Cleansing

After combining the London Geography data from Wikipedia and location data from Geocoder, data was scraped from Foursquare. The Foursquare data was divided into two datasets: Hotel location data and Transports location data. Yet, there were some problem with the data, lots of data points were missing neighborhood and postal code (See **Figure 1**).

Figure 1

formattedAddress	labeledLatLngs	lat	lng	neighborhood	postalCode	state
[19 Amhurst Road,, Hackney, Greater London, E8...	[{'label': 'display', 'lat': 51.54730441740626...	51.547304	-0.056107	NaN	E8 1LK	Greater London
[United Kingdom]	[{'label': 'display', 'lat': 51.538753390577779...	51.538753	-0.043925	NaN	NaN	NaN
[24-28 Charnley Road, Blackpool, FY1 4PF, Unit...	[{'label': 'display', 'lat': 51.533615, 'lng':...	51.533615	-0.039024	NaN	FY1 4PF	NaN
[106 Wick Road, Hackney, Greater London, Unite...	[{'label': 'display', 'lat': 51.53875853593553...	51.538759	-0.042126	NaN	NaN	Greater London

To make the best use of data points and make this project more accurate, QGIS was used to get this missing information. First, load the London shapefile and data points from Foursquare into QGIS, then use "Join Attributes by Location" method (Performing Spatial Joins, 2018). Finally, neighborhood (area) and district were obtained (See **Figure 2**).

Figure 2

Foursquare Data:

formattedAddress	labeledLatLngs	lat	lng	neighborhood	postalCode	state
[19 Amhurst Road,, Hackney, Greater London, E8...	[{'label': 'display', 'lat': 51.54730441740626...	51.547304	-0.056107	NaN	E8 1LK	Greater London
[United Kingdom]	[{'label': 'display', 'lat': 51.53875339057779...	51.538753	-0.043925	NaN	NaN	NaN

Foursquare Data Join QGIS :

	area	district	hotels	categories	lat	lng
0	Hackney Central	Hackney	KIP Hotel	Hotel	51.547304	-0.056107
1	St. Mary's	Islington	The Hotel Alternative Ltd	Hotel	51.541891	-0.100508

2.3.2. TripAdvisor Data Cleansing

After scraping the Top 300 Tourist Attractions in London from TripAdvisor, Geocoder was used to get these attractions' coordinates. Also, QGIS was used to obtain the attractions' neighborhood (area) and district (See **Figure 3**). Finally, attractions located in Inner London were selected as this report targets to Inner London area.

Figure 3

TripAdvisor Data:

	Attractions
0	Tower of London
1	Churchill War Rooms
2	Tower Bridge

Geocoder

	name	lat	lng	categories
0	Tower of London	51.509740	-0.075230	Attractions
1	Churchill War Rooms	51.502250	-0.129030	Attractions
2	Tower Bridge	51.505480	-0.075380	Attractions

TripAdvisor Data Join QGIS:

	area	district	attractions	categories	lat	lng
0	St. Katharine's & Wapping	Tower Hamlets	Tower of London	Attractions	51.509740	-0.075230
1	St. James's	Westminster	Churchill War Rooms	Attractions	51.502250	-0.129030
2	North Bermondsey	Southwark	Tower Bridge	Attractions	51.505480	-0.075380

After dropping the duplicates in these datasets, the total number for hotel is 342, transport is 574, tourist attraction is 251 in Inner London area.

This concludes the data gathering phase – the data is ready for analysis to produce the report on optimal locations for a new hotel in Inner London.

References

Aksoy, S., & Ozbuk, M. Y. (2017). Multiple criteria decision making in hotel location: Does it relate to postpurchase consumer evaluations?. *Tourism management perspectives*, *22*, 73-81.

Britain's visitor economy facts. (2019). Retrieved from https://www.visitbritain.org/visitor-economy-facts.

Lee, K. W., Kim, H. B., Kim, H. S., & Lee, D. S. (2010). The determinants of factors in FIT guests' perception of hotel location. *Journal of Hospitality and Tourism Management*, *17*(1), 167-174.

Performing Spatial Joins. (2018). Retrieved from https://www.qgistutorials.com/en/docs/performing_spatial_joins.html

Tourism in London. (2019). Retrieved from https://en.wikipedia.org/wiki/Tourism_in_London.

Yang, Y., Wong, K. K., & Wang, T. (2012). How do hotels choose their location? Evidence from hotels in Beijing. *International Journal of Hospitality Management*, *31*(3), 675-685.