

Introduction

This project aims to build a preliminary recommendation system for hotel site selection in Hong Kong, which will provide the real estate companies with basic ideas about where to build a new hotel and which class of hotel is better.

When choosing the hotel location, the land price and construction cost must be considered. However, due to complexity of these data, they will be excluded from the analysis.

The project will only focus on surrounding venues such as restaurants, parks, shopping malls and other points of interest.

Data

Source

The Hong Kong Tourism Board (HKTb, <https://partnernet.hktb.com/en/home/index.html>) provides reports of [Hotel Supply Situation\(Seasonally Published\)](#) and [Hotel Room Occupancy Report\(Monthly Published\)](#) in .xls form.

By 30 Sept 2018, there're a total of 287 hotels and 1,504 guesthouses authenticated in Hong Kong. During the month, 206 hotels and 79 guesthouses responded the HKTb's monthly survey, which will be the main data source of this project.

And the project will scrape a part of these two reports and store the data in .csv file, then uploaded into DB2 for further analysis.

Parameters

The project will use the following three datasets:

1. HOTEL_CATEGORY

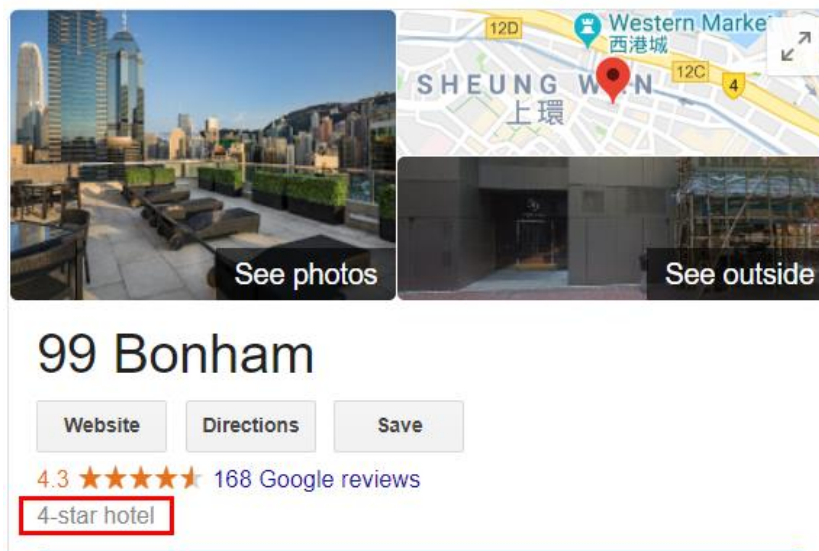
	HOTEL_NAME VARCHAR(80)	LOCATION VARCHAR(18)		TOTAL_ROOMS SMALLINT	CATEGORY VARCHAR(19)
1	338 Apartment	Central & Western		62	Unclassified Hotels
2	60 West Suites Hotel	Central & Western		60	3-star
3	99 Bonham	Central & Western		84	4-star
4	AKVO Hotel	Central & Western		30	Unclassified Hotels
5	Aveny	Central & Western		11	Unclassified Hotels
6	Best Western Hotel Harbour View	Central & Western		432	4-star
7	Best Western Plus Hong Kong	Central & Western		307	4-star
8	Bishop Lei International House	Central & Western		227	4-star
9	Butterfly On LKF	Central & Western		34	4-star
10	Butterfly On Waterfront	Central & Western		90	4-star

HOTEL_NAME --- Names of all authenticated hotels in Hong Kong. Guesthouses will not be included in the analysis.

LOCATION --- There're totally 18 District Council Districts in Hong Kong: Central & Western, Eastern Hong Kong, Southern Hong Kong, Wan Chai, Kowloon City, Kwun Tong, Sham Shui Po, Wong Tai Sin, Yau Tsim Mong, Kwai Tsing, North, Sai Kung, Sha Tin, Tai Po, Tsuen Wan, Tuen Mun, Yuen Long and Islands. Among the districts, North and Tai po aren't in the list due to there's no authenticated hotel in these areas. (See https://en.wikipedia.org/wiki/Districts_of_Hong_Kong for more information)

TOTAL_ROOMS --- Total number of rooms of each hotel, rooms under repair or being refurbished are excluded

CATEGORY --- Hotel ratings collected from Google



2. OCC_BY_CATEGORY

	CATEGORY VARCHAR(20)	NO._OF_HOTELS SMALLINT	NO._OF_ROOMS SMALLINT	HOTEL_ROOM_OCCUPANCY_RATE SMALLINT	AVERAGE_ACHIEVED_HOTEL_ROOM_RATE SMALLINT
1	High Tariff A Hotels	36	18839	88	2073
2	High Tariff B Hotels	101	30800	91	1110

CATEGORY --- Instead hotel star-rating, the Hong Kong Hotel Classification System classifies hotels into 3 categories: High Tariff A Hotel (similar to 5-star hotel), High Tariff B Hotel (similar to 4-star hotel) and Medium Tariff Hotel. However, the category of each hotel is not publicly available, only individual hotels are informed of their respective category so that they can compare their own performance against their category averages in the hotel industry. Thus, this project will use the star ratings for analysis.

NO._OF_HOTELS --- Total numbers of hotels of each category

NO._OF_ROOMS --- Total numbers of rooms of each category. Rooms under repair or being refurbished are excluded.

HOTEL_ROOM_OCCUPANCY_RATE --- Based on daily rooms occupied against daily rooms available for sales of responded hotels and guesthouses to HKTB Monthly Hotel Room Occupancy Survey.

AVERAGE_ACHIEVED_HOTEL_ROOM_RATE (ARR) --- The average of ARR of responded

hotels and guesthouses to HKTB Monthly Hotel Room Occupancy Survey. ARR excluded Government tax and all non-room related components such as F&B, laundry, airport transfer, service charges, etc., which have been built into the room rate.

3.OCC_BY_DISTRICT

	DISTRICT VARCHAR(29)	NO_OF_HOTELS SMALLINT	NO_OF_ROOMS SMALLINT	HOTEL_ROOM_OCCUPANCY SMALLINT
1	Central & Western	53	9037	89
2	Wan Chai	52	11301	91
3	Eastern & Southern Hong Kong	20	6245	87
4	Tsim Sha Tsui	57	17088	93
5	Yau Ma Tei & Mong Kok	40	6810	95

DISTRICT --- Based on the 18 District Council Districts, Eastern and Southern Hong Kong are merged together, Kowloon City is merged into Other Kowloon along with Kwun Tong, Sham Shui Po and Wong Tai Sin. New Territories includes Kwai Tsing, Sai Kung, Sha Tin, Tsuen Wan, Tuen Mun and Yuen Long. And due to large number of hotels, Yau Tsim Mong is separated into Tsim Sha Tsui and Yau Ma Tei & Mong Kok.

NO._OF_HOTELS --- Total numbers of hotels of each district

NO._OF_ROOMS --- Total numbers of rooms of each district. Rooms under repair or being refurbished are excluded.

HOTEL_ROOM_OCCUPANCY_RATE --- Based on daily rooms occupied against daily rooms available for sales of responded hotels and guesthouses to HKTB Monthly Hotel Room Occupancy Survey.