

Coursera Capstone

IBM Data Science Professional Certificate

Opening a new hotel in Hanoi, Vietnam

Duc Vu - Feb 2020



Business Problem

Vietnam is world renowned for its natural beauty, rich history and amazing cuisine. As described by [Lonelyplanet](#): *“A land of staggering natural beauty and cultural complexities, of dynamic megacities and hill-tribe villages, Vietnam is both exotic and compelling”*. With this in mind, it is no surprise that tourism is a booming sector in Vietnam, with millions of travellers from all over the world choosing the country as their holiday destination every year. As a result, the hotel industry of Vietnam is a very attractive field for investors. However, it is a very competitive field given the number of hotels already existing and the ones being built. This is especially true for Hanoi, the capital city of Vietnam.

As with any business decision, opening a new hotel requires serious planning beforehand, one of which is choosing the location for the venue. This is crucial in determining whether the hotel will be a successful or failed investment.

This capstone aims to analyse and select the best possible locations to open a new hotel in Hanoi, Vietnam using data science.

Data

List of neighbourhoods in Hanoi: This can be scraped from [Wikipedia](#). This data defines the scope of this project which is confined to the city of Hanoi, Vietnam. The aim for now is to explore the data on a Ward level, where English resources are not readily available thus the use of Vietnamese sources. If this doesn't work out during the project we can move one level up to the District level where more English data is available.

Latitude and longitude coordinates of those neighbourhoods: This is required for data to be plotted on a map. The data can be retrieved through Python's Geocode package .

Venue data, especially those related to hotels: Clustering will be performed based on this data. The data can be retrieved using Foursquare API.