Capstone Project

1. Introduction

1.1 Background

In the past decade, the lifestyle of urban people has changed with the trends and habits of drinking coffee. Coffee, which was ancient, is identical to drinks commonly used by older men, now women and men of all ages are accustomed to drinking coffee. And not just enjoying coffee, but many people are looking for a place to drink coffee. The coffee shop has finally become a cool hangout with an internet connection while enjoying a variety of steeping coffee beans.

This coffee drinking trend will become a big business opportunity. The business world is starting to work on places that serve specialty coffee. With this trend in Hong Kong, it is possible for a coffee shop to get a good profit. However, getting into the business world is not as easy as one might imagine, especially for Hong Kong, where coffee shop is very common.

If you already have the capital to open a coffee shop, then you must have the courage, start designing strategies and seeing the market. If you have long been in love with coffee and a hobby of drinking coffee, it means you can start a business with the right passion. Therefore, I try to practice my learning at Coursera to answer relevant questions, namely designing strategies to determine which areas are suitable for opening coffee shops.

1.2 Problem

Finding data about the area in Hong Kong is a challenge that must be resolved as Hong Kong does not divide area into neighbourhood like some countries. Therefore, this project will use the list of districts in wikipedia to define the area. The price of renting a place to determine the exact location of a coffee shop is also one of the problems that must be resolved.

1.3 Interest

I believe this is a relevant challenge with a valid question for anyone who wants to open a coffee shop and determine the right location. The same methodology can be applied according to demands as applicable. This case also applies to anyone interested in exploring starting or finding new business in any city. Finally, this can also serve as a good practical exercise for developing Data Science skills.

2. Data Acquisition and Cleaning

2.1. Data Acquisition

The data acquired for this project is a combination of data from two sources. The first data source of data is scraped from a wikipedia page that contains the list of districts in Hong Kong ---> https://en.wikipedia.org/wiki/Districts_of_Hong_Kong. The following are the columns:

	Districts	Regions
0	Central and Western	Hong Kong Island
1	Eastern	Hong Kong Island
2	Southern	Hong Kong Island
3	Wan Chai	Hong Kong Island
4	Sham Shui Po	Kowloon

District : Name of the district

Region: Name of the region

The Second data source is the list of Longitude & Latitude from website latlong.net, the following are columns:

	Districts	Latitude	Longitude
0	Tsuen Wan	22.374630	114.115100
1	Sha Tin	22.383381	114.198517
2	Tuen Mun	22.396910	113.974411
3	Tai Po	22.445400	114.167709
4	Yuen Long	22.445570	114.022290

District : Name of the district Latitude : Latitude of the town

Longitude : Longitude of the town.