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. Although the Indonesian people are not addicted to coffee, which means they have to drink every day, like in Melbourne. And the Coffe shop industry is still relatively new, but in big cities like Jakarta, Coffe shop has the opportunity to get a gross profit of Rp 100 million to Rp 1 billion. However, getting into the business world is not as easy as one might imagine.

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 and postcode in South Jakarta is a challenge that must be resolved. 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 Neighboorhod Jakarta Selatan. This page contains additional information about the boroughs, the following are the columns:

	Kelurahan	Kecamatan	Kota
1	Cilandak Barat	Cilandak	Jakarta Selatan
2	Cipete Selatan	Cilandak	Jakarta Selatan
3	Gandaria Selatan	Cilandak	Jakarta Selatan
4	Lebak Bulus	Cilandak	Jakarta Selatan
5	Pondok Labu	Cilandak	Jakarta Selatan

- **Kelurahan** : Name of the urban village
- **Kecamatan** : Name of the sub-district
- **Kota / Provinsi** : Name of the province

The Second data source is the list of Logitude & Latitude from website longlat.net, and list of postcode from Wikipedia in Jakarta Selatan, the following are columns:

	Kelurahan	PostCode	Latitude	Longitude
0	Bangka	12730	-6.260838	106.820788
1	Kuningan Barat	12170	-6.236810	106.827220
2	Mampang Prapatan	14470	-6.242965	106.828490
3	Pela Mampang	12720	-6.250210	106.816750
4	Tegal Parang	12790	-6.247590	106.831810

- **Kelurahan** : Name of the urban village.
- **PostCode** : Number of the posccode area.
- **Latitude**   : Latitude of the urban village.
- **Longitude** : Longitude of the urban village.