

IBM CAPSTONE PROJECT - DS

Entering the coffee shop market in Jakarta, Indonesia

By: Ega Kurnia Yazid



TABLE OF CONTENTS

INTRODUCTION	3
Background	3
Business Problem	3
DATA	4

INTRODUCTION

Background

Indonesia is one of the world largest coffee exporters and drinking coffee has been Indonesian morning mantra for centuries. Therefore, considering Indonesia potential as the home of many coffee varieties; then, bringing Indonesia coffee culture a step further will be a great movement for the industry. Moreover, on-the-go delivery coffee or beverages is becoming a new hype these days. Utilizing these opportunities, then opening a new coffee shop at the right time will be really challenging. Then, it is important to be careful to pick a strategic location to open a new coffee shop.

Business Problem

This report aims to figure out the best location to enter the coffee market in Jakarta, Indonesia. Using data science and machine learning capacities such as clustering, this project will answer the less coffee shop district around Jakarta, so that it will be less competitive to enter the coffee shop market.

The smell of fresh-made coffee is one of the world's greatest inventions. - Hugh Jackman

Jakarta, March 14, 2020

Ega Kurnia Yazid
Project's Data Scientist

DATA

To respond to the problem, hence, this project will use multiple datasets, containing:

- List of neighborhoods (districts) in Jakarta, Indonesian capital city;
- Latitude and longitude data that showing the coordinate of its neighborhood;
- Venue data which informs the nearby venue around the neighborhoods.

Sources of data and method to extract them



The main neighborhood data is retrieved from the Wikipedia page which contains information about Jakarta's districts (https://en.wikipedia.org/wiki/Category:Districts_of_Jakarta), with total of 44 districts. This data retrieved using Python, particularly by scraping with BeautifulSoup package. Furthermore, geographical coordinates dataset is retrieved by using Geocoder package which allow me to get the geolocation of Jakarta's districts. Along with it, I used Foursquare API to get the venue data around the districts and, finally, map the results using Folium.