Coursera Capstone

IBM Applied Data Science Specialization

Analysis of Opening a New Shopping Mall in Bali, Indonesia

Nelson Agus Kesuma

Introduction

Bali is a heaven for nature enthusiasts and party goers to enjoy and pamper themselves with the wondrous of natures and beaches. This creates a great place for tourism industry to bloom, including resorts, hotels, restaurants, clubs and shopping malls. For sellers and business owners, this is a great place to start businesses to cater for a high demand. There is a need to build more shopping malls to as traffic increases and poses a good business opportunity to build convenient shopping malls for tourists to enjoy not only nature but also shopping in Bali

Business Problem

The objective of thus capstone project is to analyse and select best locations in Bali, to open a new shopping mall. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to recommend a good location to build a new shopping mall

Data

The following data is needed:

- List of neighbourhoods in Bali.
- Latitude and Longitude coordinates of these neighbourhoods.
- Venue data related to shopping malls and cluster neighbourhoods

Target Audience

Targeted for property developers and business owners who are interested in creating new business or expanding portfolios.

Data Source

Wikipedia page (https://en.wikipedia.org/wiki/List_of_districts_of_Bali) will be able to provide a list of neighbourhoods in Bali. Usage of web scraping to extract data from the Wikipedia page with the help of Python library and beautifulsoup packages are required. Then geographical coordinates of the neighbourhoods using Phyton Geocoder will provide latitude and longitude coordinates of the neighbourhoods.

Afterwards, Foursquare API to get the venue data can be used to fetch the real map, particularly in shopping mall category. This will train skills in web scraping, working with API from Foursquare, data cleaning, data wrangling, machine learning (K-clustering) and map visualization (Folium)