Opening a new Hotel in Delhi (INDIA)

Aditya Trivedi
https://www.linkedin.com/in/iadi7ya
https://github.com/iadi7ya

Table of Contents

- Introduction
- Dataset Description
- Technologies used
- Implementation
- Observations
- Future Enhancements

Introduction

- Find the best location for opening a new Hotel in Delhi (India)
- Decision making for Investment into Hotel and real estate business
- This project is particularly useful to property developers and investors
- This project is timely as the city is currently suffering from oversupply of Hotel

Dataset Description

The Dataset used in this project:

- A complete list of neighborhoods in New Delhi, India.(Source is Wikipedia)
- Geographical coordinates (latitude and longitude) of those neighborhoods. Source of the data will be FourSquare.

Technologies used

List of tools and technologies used:

- KMeans clustering algorithm used for categorizing the venues.
- Python programming for Implementation of the project.
- Pandas library for preprocessing of the data.
- Scikit-learn for applying machine learning.
- Geocoder for fetching latitude and longitude of the venues.
- Folium for the visualization of venues on the map of Delhi

Implementation

Steps involved in the end-to-end Implementation of the project:

- Exploratory Data Analysis for understanding the dataset.
- Preprocessing for cleaning of the data.
- Feature Selection for selecting the best feature for the model.
- Machine learning model development.
- Developed KMeans clustering model.
- Cluster analysis for selecting best places for opening a new Hotel

Observations

Cluster Name		Suitable for New Hotel	Detailed Description
1	First Cluster	✓	• Lower to moderate availability of Hotels. • Mid size cluster
2	Second Cluster	✓	• Very Low/None availability of Hotels. • Mid size Cluster
3	Third Cluster	✗	• Availability score of the hotels is very high • Cluster consists of only two places
4	Fourth Cluster	✗	• Availability of hotels is good if not the best. • Mid size Cluster
5	Fifth Cluster	✗	• Availability score of the hotels is very high. • Cluster consists of only two places

Future Enhancements

- Collect more data (demographic, income, investments etc)
- Combine multiple models for better decision making
- Consult from domain experts of Real estate and Investments