Prediction of Under Valued Properties in Singapore

BUSINESS PROBLEM / IDEA

An investment firm is exploring options to invest in property within Singapore due to its central location in Asia. Singapore is one of the largest metropolises in the world where over 5,638,700 (2018 estimate as per Wikipedia) people live and second dense country in the world with 7,804/km2 (20,212.3/sq mi) population density. Investor did not apply any area restriction so the analysis will be applied on all 28 districts of Singapore for predicting most attractive investment options. The overall purpose is to predict underpriced properties based on multiple parameters and avoid over-priced venues.

The property prices based on three major factors will be considered:

1) Type of property (lease or owned) & lease tenure
2) Distance from city central area, and
3) Proximity with major shopping/restaurant areas.

Following step-by-step approach will be applied to achieve desired results:

- Gather above mentioned three factor and current housing prices
- Derive housing prices formulae based on these three factors
- Shortlist top twenty undervalued locations (outliers) for further investigation and top twenty overvalued ones for rejection.

DATA

Based on business problems, following data sources will be used to achieve desired outcomes:

- Recent property price in various neighbors using webscraping from websites https://www.squarefoot.com.sg/latest-transactions/sale/residential/condominium
    - Reading the data
    - Webscrapping it from website
    - Cleaning the Webscrapped data

- Estates' locations and driving distances from city central area using GoogleMap API
    - Using the GoogleMAP API (after creating an account)
    - Calculate distances of the properties from the central locations and adding the information in the data frame

- Proximity with major shopping areas, Restaurants using FourSquare API
    - Using the FourSquare API, finding out the no. of important locations (e.g. Restaurants, Shopping Malls) in the vicinity of properties and adding this data in the data frame

- Understanding all the data using linear regressions

- Building the Model and get the necessary property categories