MELBOURNE HOUSING MARKET ANALYSIS

Intergrating a business case to help a client find the most suitable house for their preference

Background

According to the annual Demographia international housing affordability survey, Sydney ranked third and Melbourne fourth compared to 309 housing markets across eight countries including some of the most expensive cities in the world such as Hong Kong or Vancouver. Melbourne's housing market is heating up, as spring is coming. “Spring is a traditional strong sales period, and with fresh demand from buyers, now is the time for anyone thinking about selling a house or a unit to act”, Real Estate Institute of Victoria CEO Gil King says. However, in such a scenario, this situation imposes a big concern for homebuyers: Among hundred thousand available real estate on sale, which one is both most suitable and affordable for my preferences and financial status?

Business Insights

This report aims at two target audience: First, anyone who is planning to buy/rent a new house/apartment in Melbourne. Second, real estate agents whose responsibility is to find the most suitable one for their clients, yet still profitable for the company.

In my opinion, it is necesarry to adopt machine learning and data analysis to provide recommendations for them to decide or purchase a suitable and reasonably priced real estate in Melbourne.

To solve this business problem, we are going to cluster Melbourne neighborhoods in order to recommend venues and the current average price of real estate where homebuyers can make a real estate investment. We will recommend profitable venues based on amenities and essential facilities surrounding such venues i.e. elementary schools, high schools, hospitals & grocery stores.

Business Case

Imagining there is a potential client coming to your door. The couples are Vietnamese migrants and have 2 kids, 1 boy 1 girl. They express the preference for duplex house with at least 3 rooms. They want some Vietnamese restaurants within their proximity. Other essential amenities and recreational places, especially for sports, are also highly welcomed.

So, as a real estate agent, how can you help them to find their dream house?

Data Collection

Data on Melbourne properties and price was scraped from publicly available results posted every week from DomainGroup < https://www.domain.com.au/ >. The dataset includes almost every information you request of a property, such as Address, Type of Real estate, Suburb, Method of Selling, Rooms, Price, Real Estate Agent, Date of Sale. The dataset even contains location coordinates of properties for the sake of clustering and visualizing.

Methodology

The original data set consists of **34857 rows** and **21 columns,** which means that we have the data of total 34,857 properties all over Melbourne. This dataset is too big and general, besides, there are quite many missing values. We need to cleanse and preprocess it before taking further analysis.

We drop all the rows containing missing values related to the price, house type, number of rooms, as they are prioritized requirements of the client. In order to understand the dataset clearly, there are some key details to check out:

Key/Method	**Meaning**
Suburb	Suburb
Rooms	Number of rooms
Price	Price in Australian dollars
Br	Bedroom(s)
H	House, cottage, villa, semi, terrace
U	Unit, duplex
T	Townhouse
Dev site	Development site
O res	Other residential
Distance	Distance from Central Business District (City Centre) in kilometres
Regionname	General Region (West, North West, North, etc.)
Propertycount	Number of properties that exist in the suburb
CouncilArea	Governing council for the area

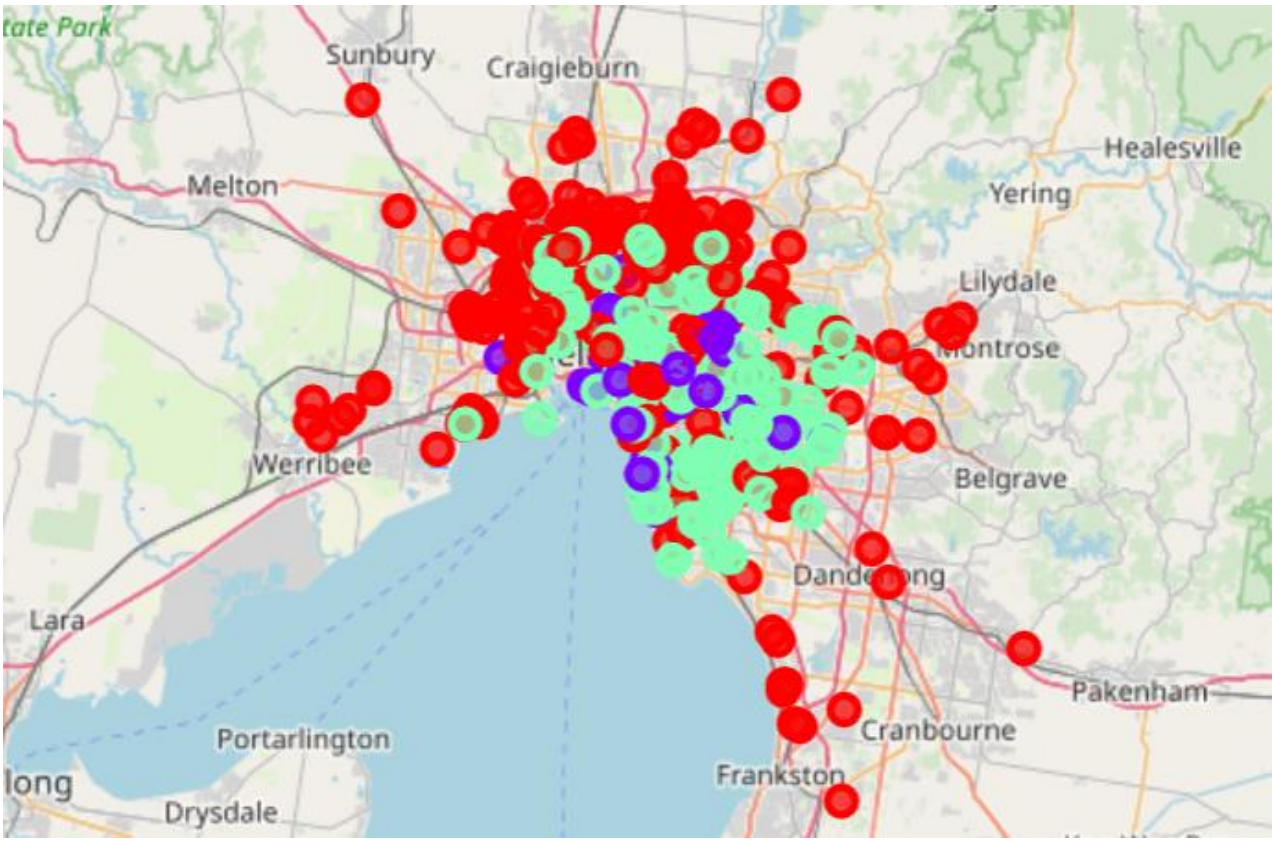
Then we start sorting out all properties. Only the properties that are unit or duplex ("u") with at least 3 rooms are chosen. After sorting out, 576 out of 34,857 properties meet the requirements. The dataset's size shrinks significantly, and it is ready to be clustered.

Section 1: Clustering house price into 3 price groups

We choose the k-Means algorithm to perform unsupervised learning because the number of clusters is prespecified already. We want the price to fall into 3 categories: Low Price, Medium Price, High Price. After running the model, we have this result:

Cluster	Range (AUD)	Category	Count
0	780,000 - 1,296,000	Medium Price	238
1	1,320,000 - 3,610,000	High Price	60
2	301,000 - 777,000	Low Price	278

Next, we visualize the clustering on the map to see the distribution of each price group. Running the Folium map, we have the following result, with Red, Green, Purple respectively representing Low, Medium and High Price.



Section 2: Get nearby venues of each suburb

To explore and target recommended locations across different venues according to the presence of amenities and essential facilities, we will access data through FourSquare API interface and arrange them as a dataframe for visualization. By merging data on Melbourne properties and price and data on amenities and essential facilities surrounding such properties from FourSquare API interface, we will be able to recommend suitable real estate investments.

We use location data to find nearby venues within the radius of 800 meters in each neighborhood. After that, at each suburb we find top 5 most common venues by the frequency of location. Here are some of examples:

	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Abbotsford	Vietnamese Restaurant	Café	Thai Restaurant	Pub	Korean Restaurant
1	Aberfeldie	Athletics & Sports	Café	Park	Playground	Golf Course
2	Airport West	Café	Grocery Store	Airport	Thai Restaurant	Supermarket
3	Albion	Café	Music Store	Bus Stop	Park	Vietnamese Restaurant
4	Altona	Café	Pizza Place	Beach	Park	Train Station
5	Altona Meadows	Fish & Chips Shop	Cricket Ground	Restaurant	Zoo Exhibit	Event Space
6	Altona North	Pizza Place	Athletics & Sports	Grocery Store	Mexican Restaurant	Gym / Fitness Center
7	Armadale	Café	Light Rail Station	Convenience Store	Japanese Restaurant	Grocery Store
8	Ascot Vale	Café	Pizza Place	Light Rail Station	Bakery	Italian Restaurant
9	Ashwood	Café	Thai Restaurant	Pharmacy	Korean Restaurant	Liquor Store
10	Avondale Heights	Bakery	Electronics Store	Café	Mexican Restaurant	Restaurant
11	Balaclava	Café	Breakfast Spot	Coffee Shop	Pizza Place	Bar
12	Balwyn	Café	Park	Grocery Store	Gym	Light Rail Station
13	Balwyn North	Café	Park	Pharmacy	Thai Restaurant	Supermarket
14	Bayswater North	Golf Course	Sporting Goods Shop	Convenience Store	Gas Station	Fast Food Restaurant
15	Beaconsfield	Café	Park	Train Station	Steakhouse	Supermarket
16	Bellfield	Café	Grocery Store	Food Court	Fast Food Restaurant	Gas Station
17	Bentleigh	Café	Grocery Store	Park	Fast Food Restaurant	Train Station
18	Bentleigh East	Café	Bakery	Grocery Store	Hotel	Golf Course

The final thing is sorting out which suburb has most Vietnamese restaurants, as this is one of the client's preferences. The results are as follow:

	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Abbotsford	Vietnamese Restaurant	Café	Thai Restaurant	Pub	Korean Restaurant
20	Box Hill	Vietnamese Restaurant	Athletics & Sports	Baseball Field	Gourmet Shop	Supermarket
63	Footscray	Vietnamese Restaurant	Café	Asian Restaurant	Bakery	Bar

CONCLUSION AND DISCUSSION

- Houses are clustered reasonably: Low-priced ones are still the majority, followed by the middle-priced. High price range is quite out of their capability, so we don't discuss about it.
- As we can see from the map, the more expensive the property is, the closer it is to CBD (Central Business District) of Melbourne.
- There are 3 suburb areas that have Vietnamese restaurants within their proximity: Abbotsford, Box Hill, and Footscray. Of these three potential suburbs, Box Hill also has sports venues and one supermarket nearby, which is suitable for households. Moreover, Box Hill is in Low Price cluster where housing price ranges from $301,000 to $777,000, which fits their financial capability.

In conclusion, the final recommendation that we give to our Vietnamese client is **BOX HILL, MELBOURNE**