1. Business Problem

In this project we will try to find an optimal location for investing in a rental property in Melbourne. This report is targeted for investors looking to invest in residential rental property in Melbourne at a suitable location which can generate a high rental yield.

Rental yield measures the profit you generated investment property as a percentage of its value. A good rental yield is dependent of two factors a) profit from the property i.e. earnings minus all expenditures and b) purchase price. In other words, rental yield will be high if earnings i.e. rentals are maximised with minimum investment.

A property can be rented through rental agencies for long term rentals and also short-term letting services like Airbnb. Short term rentals have risen in popularity in recent years and with over 20,000 active listings across Metropolitan Melbourne has experienced one of Australia’s most robust growths in short-term rentals over the past two years. Hence, we will be looking at Airbnb listing data as an indicator of rental yield prospect.

Melbourne is a very big city with lot of property investment options. Property prices vary quite a bit in different parts of the city depending on various factors. Airbnb listings also varies greatly across the city along with rental tariffs or rates. Finally, number of venues like cafes, restaurants, pubs, shops and other places of interest, tourist locations also vary across the city. A high count of venues in a neighbourhood would imply popularity of that neighbourhood with renters.

Hence, to achieve a best rental yield, we need to look at areas with -

* relatively high Airbnb rents
* a high number of venues
* with not so high property prices
* less competition from other properties / listings in that area i.e. lower count of listings will indicate less competition.

We will use data science tools and methodology to identify a few most promising investment locations based on the above criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

1. Data Requirements

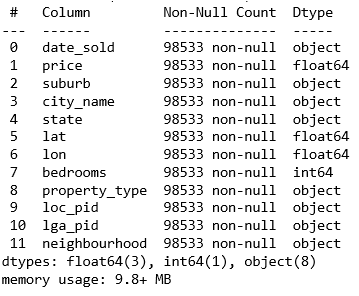
Based on the business problem, we will be looking at below data –

* Current residential property prices in Melbourne metropolitan region from Residential sales [data](#ResSalData)
* Median daily rent of Airbnb [listing](#AirbnbData) in a particular area
* Number of Airbnb [listings](#AirbnbData) in a particular area
* Venues and amenities (restaurants, cafes, pubs) in and around the area from Foursquare [data](#FourSqrData).
  1. Residential sales data -

We will extract Melbourne residential property prices from Australian real estate sales Sep’18-Jun’20 dataset as uploaded on Kaggle by HtAG Holdings.

<https://www.kaggle.com/htagholdings/aus-real-estate-sales-march-2019-to-april-2020>

We will filter Melbourne data from this dataset. We will also cleanse the data by removing any row where any of the cells are Null or NAN. We will review the data for any outliers and remove those as well to get a correct perspective of the data. Below is a snapshot of fields in the data -



We would be primarily be using price, suburb, city\_name, property\_type, latitude (lat), longitude(lon) and neighbourhood in our assessment. The rest of the fields are not required.

After data cleansing, we will then find median price of properties for all suburbs in Melbourne metropolitan region based on the property type (house, town house and unit/ apartment).

Note

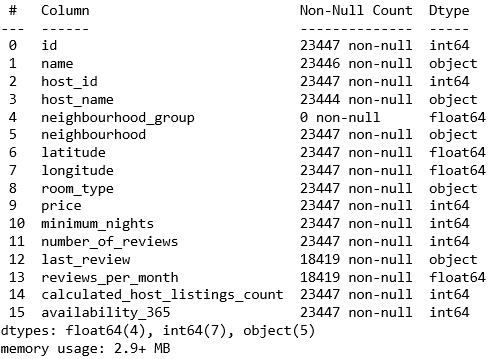
The residential sales data is available for suburbs while the Airbnb listing data is available for neighbourhoods or local government areas (LGAs). An LGA will have multiple suburbs. Hence, we will have to extract corresponding neighbourhoods for all suburbs in the residential sales data. This is available on Wikipedia at <https://en.wikipedia.org/wiki/List_of_Melbourne_suburbs>.

* 1. Airbnb listings

We will extract Airbnb listings data from ‘Inside Airbnb’ dataset

<http://insideairbnb.com/melbourne/?neighbourhood=&filterEntireHomes=false&filterHighlyAvailable=false&filterRecentReviews=false&filterMultiListings=false>

Similar to property prices, we will extract Airbnb listing data for all suburbs. As with residential sales data, we will also cleanse the data by removing any row where any of the cells are Null or NAN. We will review the data for any outliers and remove those as well to get a correct perspective of the data. Below is a snapshot of the Airbnb data –



We would be primarily be using price, neighbourhood, room\_type, price in our assessment. The rest of the fields are not required.

After data cleansing, analyse, average daily rentals across Melbourne neighbourhoods for different accommodation types and find out total listings per neighbourhood.

We will assess what accommodation types are most sought after i.e. entire home/ apartment, a private room, a shared room or a hotel room.

* 1. Foursquare data

Finally, we will extract venues and amenities data from Four square. An area or a location with high number of venues will indicate a popular area which short term renters might prefer.