Melbourne Housing Price and Neighbourhood Venues Analysis

Applied Data Science Project by IBM

Table of content

Introduction: Business Problem

Due to high cost of living, Melbourne housing can be a nightmare for most. Melbourne is currently experiencing a housing bubble (some experts say it may burst soon). A potential client aspiring to buy a suitable property would like to become knowledgeable about the ongoing pricing to make a conscious decision. Furthermore, he/she would like to consider several factors like proximity to schools, medical care, restuarants, other liesure amenities to accomodate needs.

With Melbourne housing market data coupled with data science techniques, one can derive useful insights and information about current pricing in different suburbs of Melbourne while considering other factors of his/her choice. This would help to make an informed decision about owning a property in a suitable location in Melbourne.

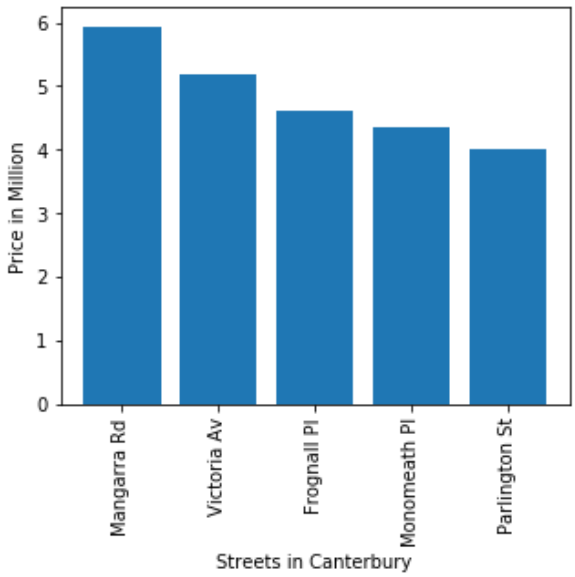
Data

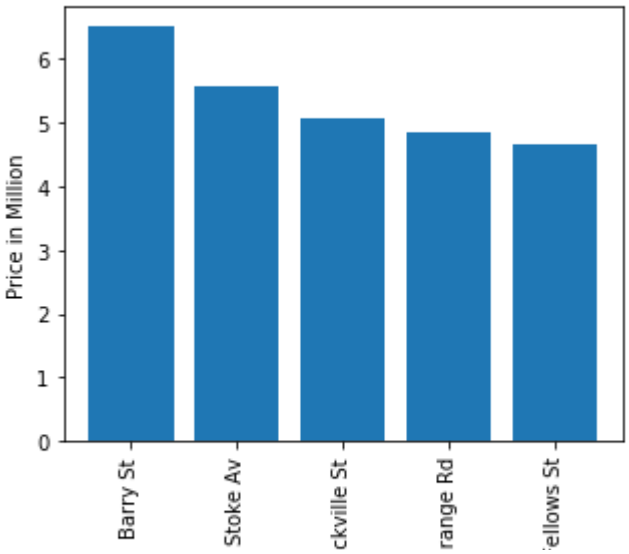
- **Mebourne Housing Dataset**
  This data was scraped from publicly available results posted every week from *Domain.com.au*. The dataset has been cleaned, and now it available for us folks (data analysts) to do some data analysis magic.
  Author: Tony Pino
  License: Attribution-NonCommercial-ShareAlike 4.0 (CC BY-NC-SA 4.0)
  Description:
  Suburb: Suburb, Address: Address, Rooms: Number of rooms, Price: Price in Australian dollars, Method: S - property sold; SP - property sold prior; PI - property passed in; PN - sold prior not disclosed; SN - sold not disclosed; NB - no bid; VB - vendor bid; W - withdrawn prior to auction; SA - sold after auction; SS - sold after auction price not disclosed; N/A - price or highest bid not available, Type: br - bedroom(s); h - house,cottage,villa, semi,terrace; u - unit, duplex; t - townhouse; dev site - development site; o res - other residential, SellerG: Real Estate Agent, Date: Date sold, Distance: Distance

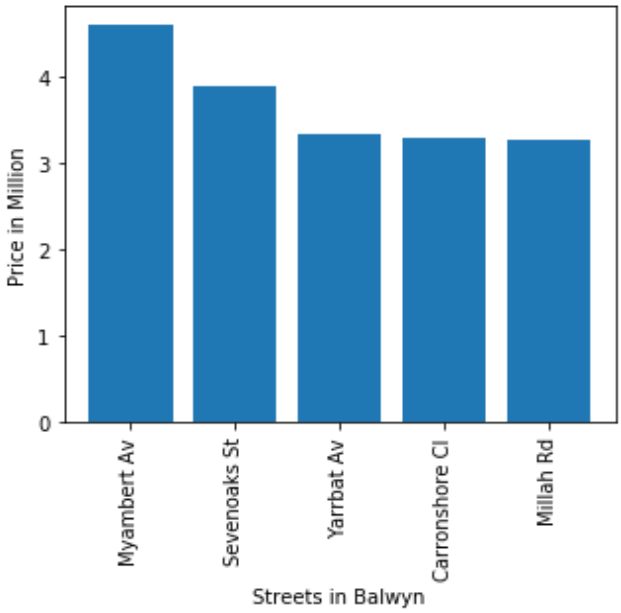
from CBD in Kilometres, Regionname: General Region (West, North West, North, North east ...etc), Propertycount: Number of properties that exist in the suburb, Bedroom2 : Scraped # of Bedrooms (from different source), Bathroom: Number of Bathrooms, Car: Number of carspots, Landsize: Land Size in Metres, BuildingArea: Building Size in Metres, YearBuilt: Year the house was built, CouncilArea: Governing council for the area, Lattitude: Self explanitory, Longtitude: Self explanitory
Duration: January 2016 - October 2018
Link: Kaggle link

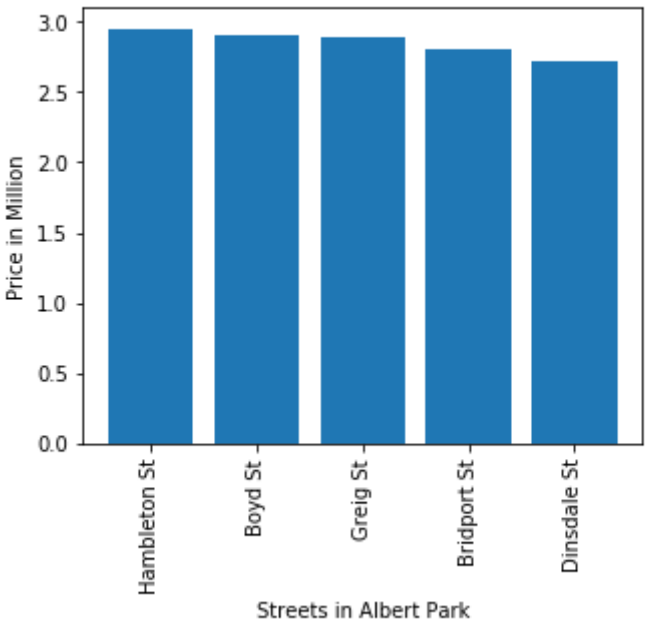
- **Foursquare Locatation Data**
  Description: To determine the various amenities in the proximity of a desired location, Foursquare location data is used.
  Link: Foursquare website

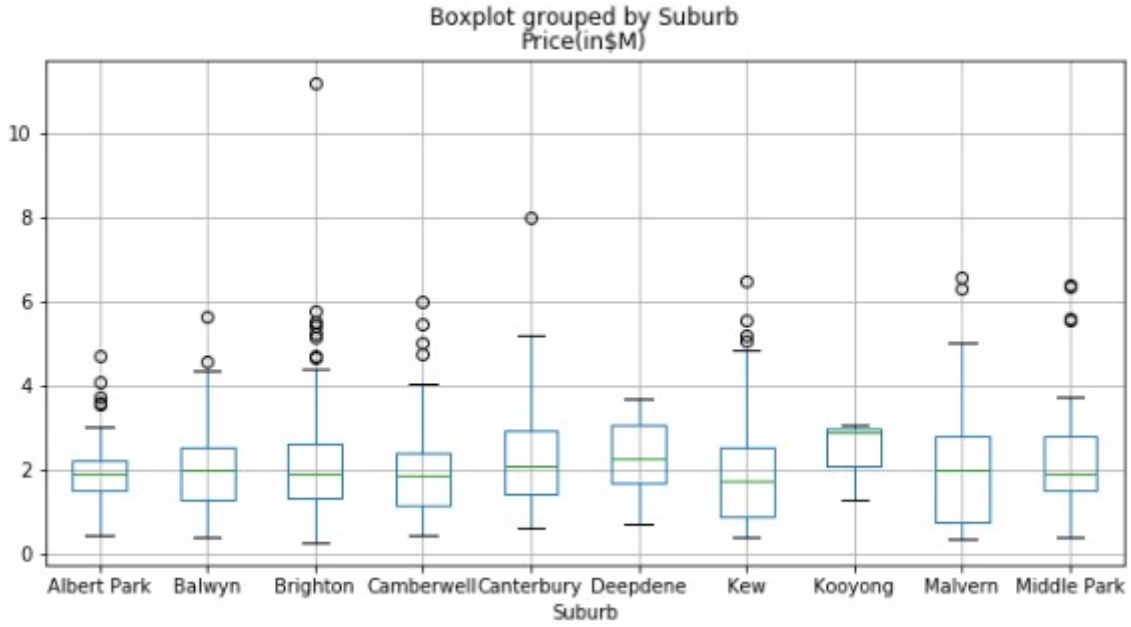
Analysis

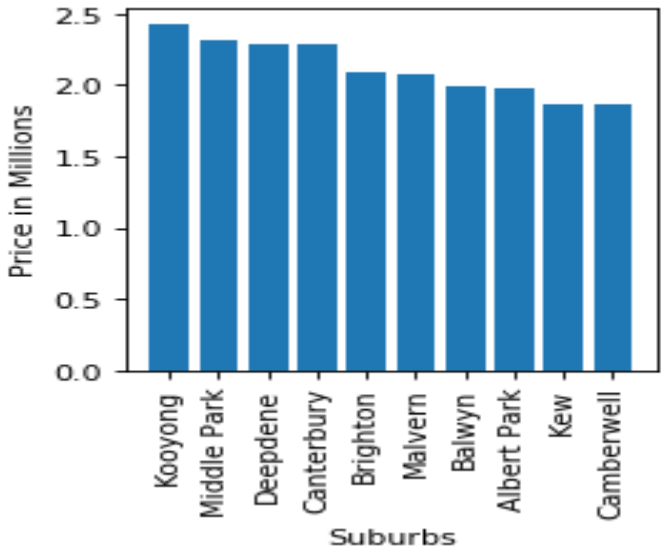









Boxplot grouped by Suburb
Price(in$M)
10
8
6
4
2
0
Albert Park
Balwyn
Brighton
Camberwell
Canterbury
Deepdene
Kew
Kooyong
Malvern
Middle Park
Suburb


2.5
2.0
1.5
1.0
0.5
0.0
Price in Millions
Kooyong
Middle Park
Deepdene
Canterbury
Brighton
Malvern
Balwyn
Albert Park
Kew
Camberwell
Suburbs

Methodology

In this project I will be focusing on investigating the recent (from January 2016 to October 2018) housing market prices of residential properties in the city of Melbourne and to recommend buying at various potential locations based on price.

In first step we will be cleaning, filtering and transforming the data obtained from the Melbourne Housing Market dataset which includes the transactions in the period from 2016 to 2018.

In the second step we will do exploratory data analysis on various suburbs and streets. Unique "street names" in the city of Melbourne in each suburb where recent transactions for sale of property were done are filtered from the dataset.We will caculate the average price of property on each of those streets is determined by taking a mean on recent transactions of sale of property on respective streets.

And the current average prices are compared and all recommendations for the locations are made by plotting them on map of Melbourne. The location popups are labelled with the respective street names and their average property price

In the third step, we will coordinates locations i.e. latitude and longitudes of the streets are fetched from the Melbourne Housing Market dataset And we will build recommended locations determined based on average pricing are further fed into Foursquare API calls to discover various amenities in proximity to them. All reported venues are then tabulated, analysed thoroughly and presented.

Results and Discussion

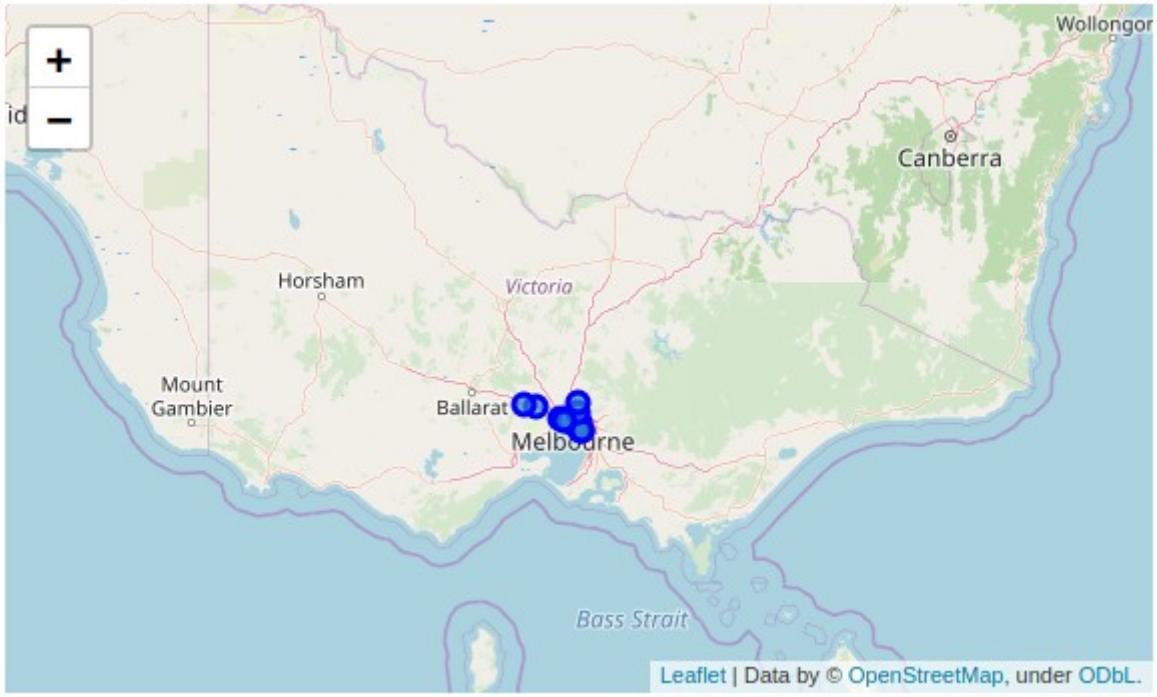
Based on the findings in the results section, the user can take a conscious decision about choosing a street i.e. location based upon his/her specific requirements.

The results section enlists 26 locations where a prospective client can buy a property based on his needs and choices. Such choices would be affected by the venues and facilities which are close to the property which match against his familial needs.

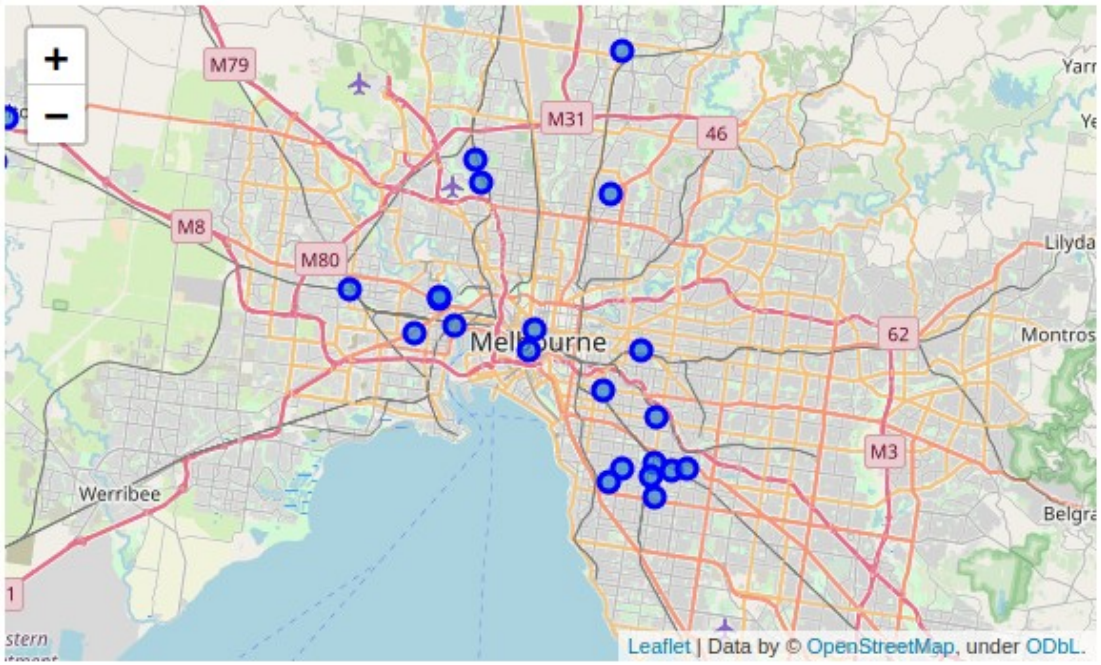
Few possible use-cases are:

1. A prospective client with elders in the family would be inclinded to choose a location where hospitals and grocery stores are located in close proximity
2. A prospective client with kids in the family would be choosing a location where elementary and high schools are nearby. He would also like to choose a place with parks and other venues in the close vicinity
3. A bachelor would be inclined to choose a property which has pubs, bars, entertainment, etc. close to the property

t[21]:



Out[11]:



Conclusion

The decision of a buyer is influenced by the familial needs, personal biases and so on. So, based on the findings summarized in the results and discussion sections, following conclusions can be made:

1. While making recommendations to a prospective client, it is imperative to know requirements besides the budget, which dictates his/her decision of buying the property largely. This would help to catch attention

2. Knowledge about the most recent market prices can be very helpful for the client and can help him take an informed decision