

1. TITLE:

Analysis of Neighbourhoods in Perth, Western Australia, for business, contractors and job seekers to be able make better decisions on where to live, work and to start a business.



Table of Contents

1. Title:

2. Table of Contents and Introduction

3. Links and Literature

4. Data Analysis and Methodology Description

5. Data Analysis and Results

6. Discussion

7. Conclusions

8. References (not already in the links section)

9. Appendices 

2. INTRODUCTION Western Australia

Western Australia has the highest GDP/GSP per capita of any state in Australia, for 2019/2020 this was ~110000 AUD. This GDP/GSP value also makes the global top ten GDP list if Western Australia were a separate country. Western Australia's political stability, attractive Mediterranean type climate of hot summers and mild wet winters and a labour shortage makes the Capital City Perth a desirable location to move too in search of opportunity. Perth's isolation and low population density has also helped control the spread of COVID-19. Apart from the occasional short lockdown, lifestyle is more or less back to the normal pre 2020 level.

[Perth](#) is spread out along the Indian Ocean Coast from Dawesville to Yanchep, a distance of 121 Km ([Haversine](#)). The road transportation network is now well developed There are suburban rail and bus networks. The rail network is mainly used for workers commuting to the CBD. There are continuing developments in these areas including a rail link to Perth International airport; and [extensions](#) of the [Mitchell Freeway](#) and [Tonkin Hwy](#) to the North of the city which is likely to promote further suburban development.

The rapid growth of Perth and Western Australia since the discovery of iron ore in the Pilbara in the 1950s has resulted in considerable opportunity not only in the mining industry but throughout the state. Western Australia is also a significant gold and agricultural producer and is one of the great grain growing areas of the World. The hazards and the remote locations of most of the mines has produced a "fly in fly out" FIFO culture which is very well paid. I have personally heard people who have worked and are working in the mining industry boasting of pay off their family home mortgages in 3 years? The rapid growth of the suburbs has resulted in the shortage of amenities and hence opportunities for business. For, example when I was living in Aveley in 2011, I could only find one petrol station in the entire Ellenbrook/Aveley/TheVines area; these suburbs have a combined population of over 60000.

The methodology to provide information that businesses, contractors, and employees can use to make decisions uses a combination of suburb geolocation data, venue information using the foursquare API, a machine learning algorithm K-means clustering for clustering similar suburbs with similar venues and attributes, and median house price data for randomly selected suburbs. A multiple linear regression analysis will attempt to build a model to predict house median house price (dependant variable) in suburbs using two independent variables, distance from CBD, and distance from the coast. The geolocation data will be used to calculate the distances using the Haversine calculation. This will be explained in more detail in the next section.

3. LINKS and LITERATURE

[List of WA government agencies](#)

[Chamber of Commerce and Industry WA](#)

[Recruitment Agencies in Perth](#)

[Mining and Energy Companies in Perth and WA](#)

[Abattoir's in Western Australia](#)

[Trip Advisor Resorts in Perth and Western Australia](#)

[Backpackers in Western Australia](#)

[Backpacker Jobs in Perth and WA](#)

[Top 20 Perth Development Projects](#)

[Construction Work in Perth](#)

[Truck Driver Jobs in Perth](#)

[Plan WA Map](#)

[Light Industrial Areas Perth](#)

New Residential Housing Subdivisions

[Flynn Sub Divison Experts](#)

[Satterley](#)

[Stockland](#)

[Aigleroyal](#)

[housing.wa.gov.au](#)

[Summit Homes](#)

Infrastructure, Commercial Real Estate and Businesses for Sale

[Aigleroyal](#)

[Commercial RealEstate](#)

[LincProperty](#)

[Infrastructure Project Map](#)

[Major Projects Perth](#)

[Restaurants and Food Outlets for Sale in Perth¶](#)

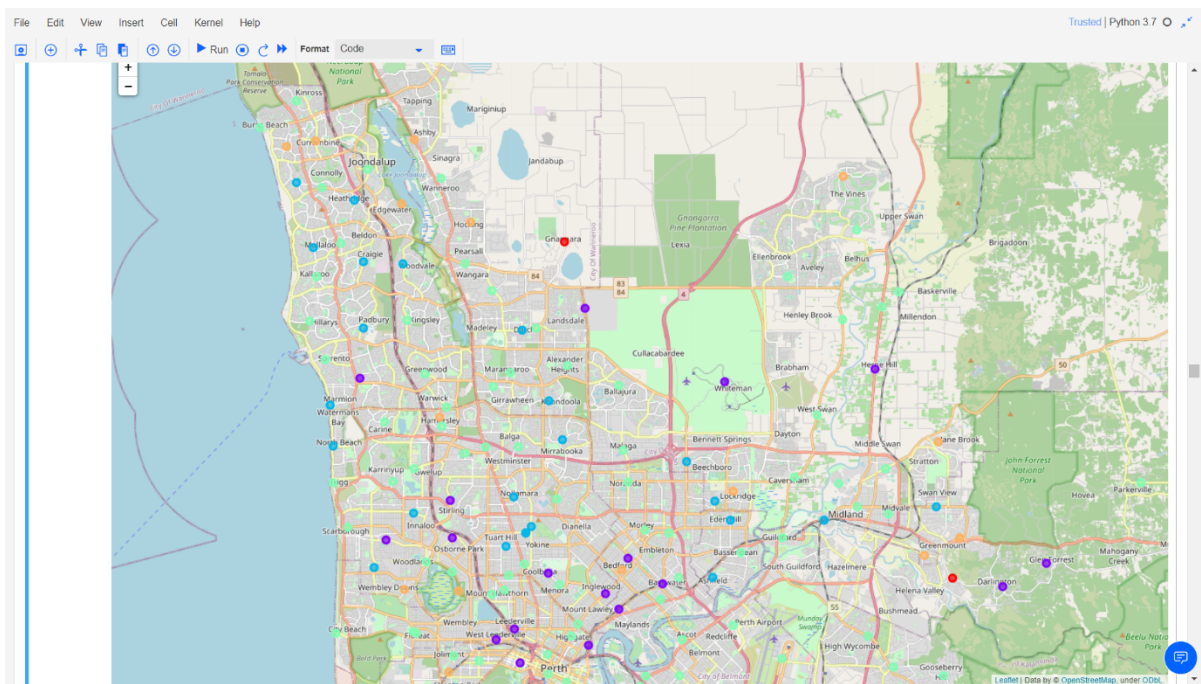
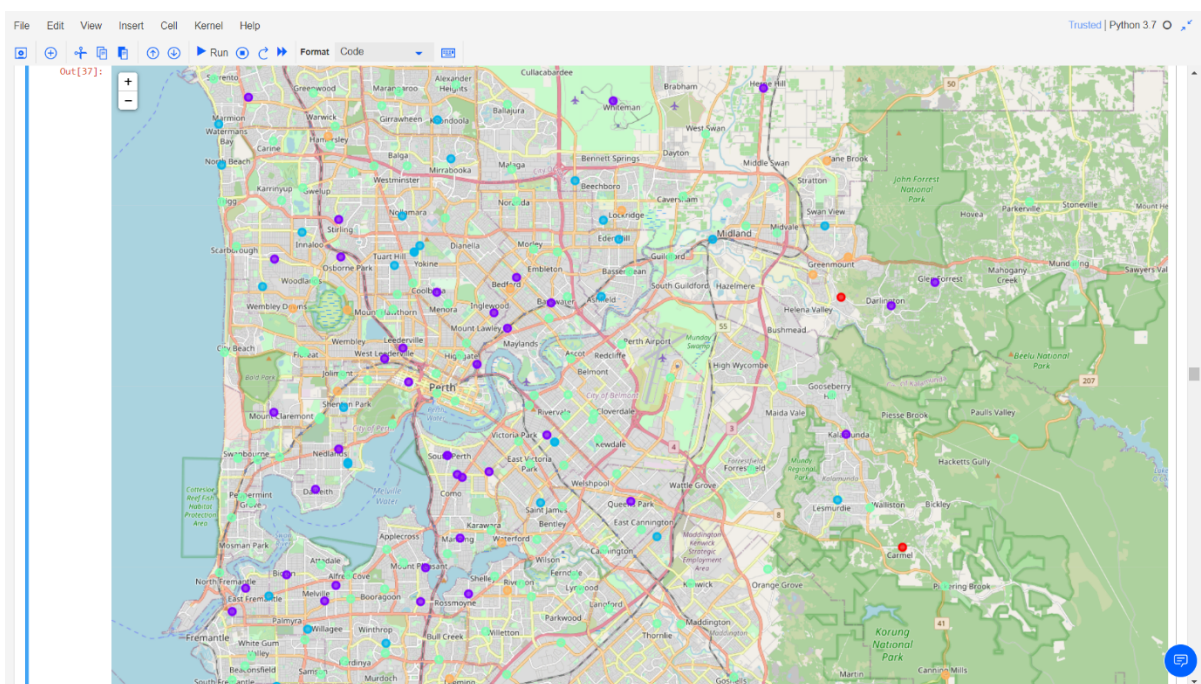
4. Data Analysis and Methodology

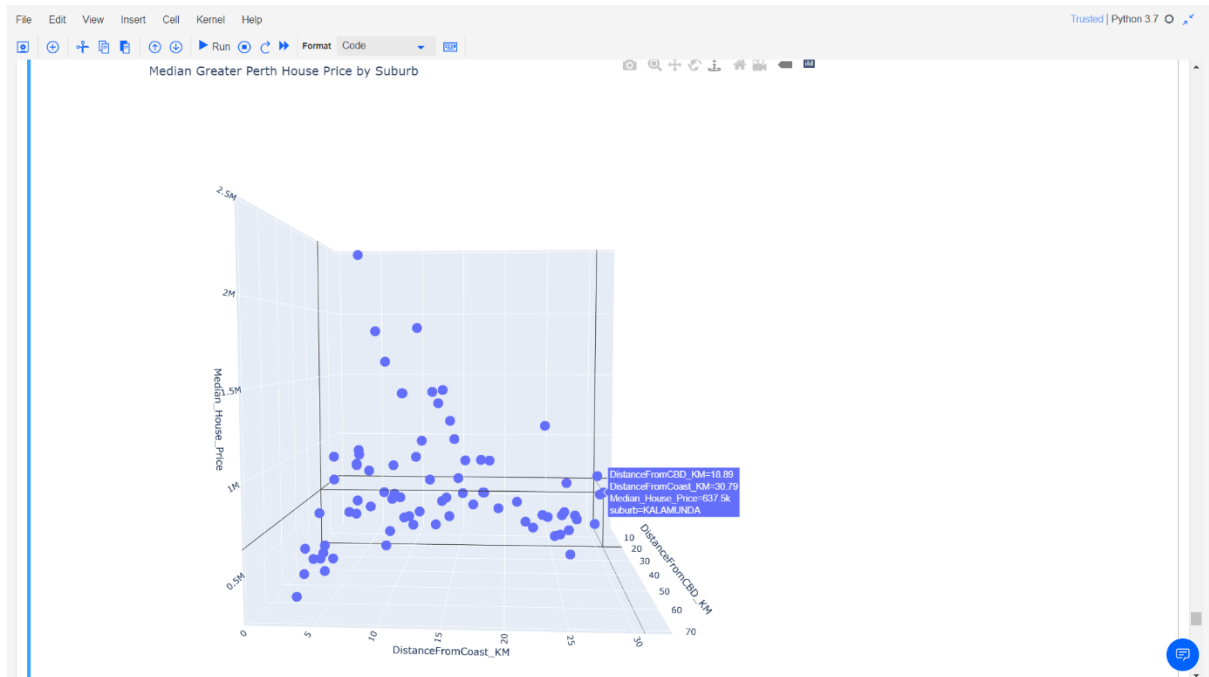
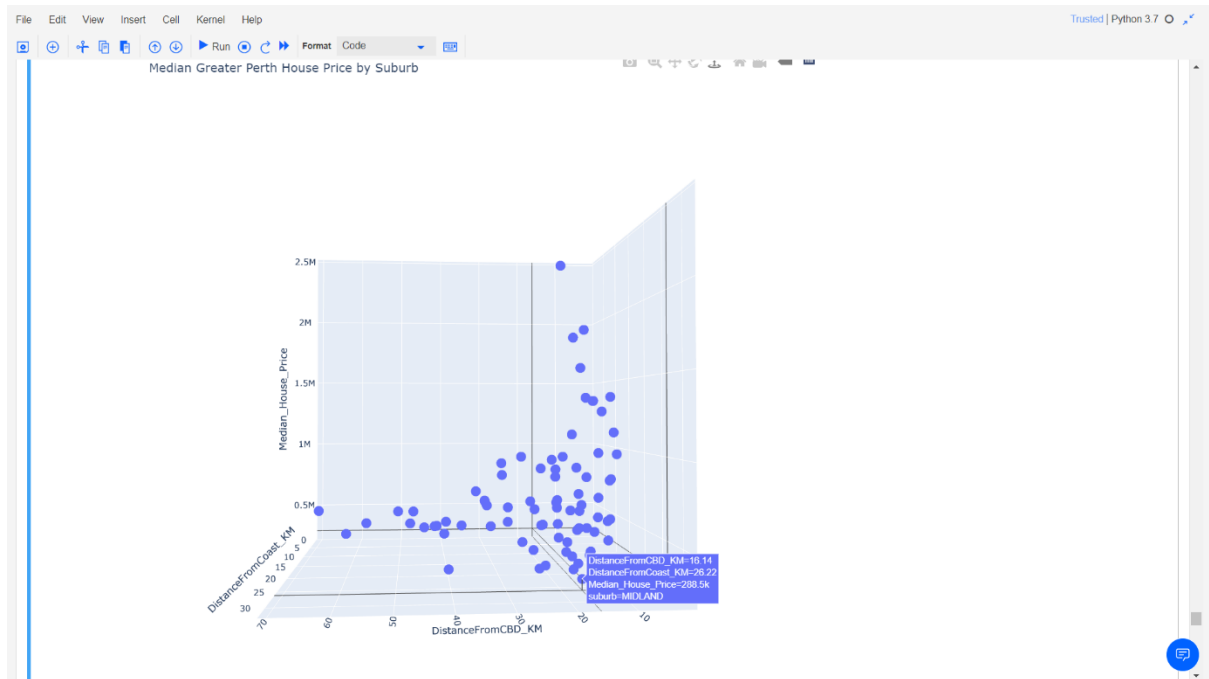
The geolocation data was obtained from <https://www.corra.com.au/australian-postcode-location-data/>

This data was then wrangled in Excel to create a file of only the Greater Perth suburbs and uploaded to the IBM cloud Watson Studio site and converted into a Pandas data frame in a Python 3.7 Jupyter Notebook. The python packages, for example, pandas, matplotlib, and folium were already imported. The mapping package folium was used to display the Perth suburbs in a Map.

The foursquare API was accessed with my credentials [The foursquare API](#) was tested out retrieving venues within a 500 m radius of a PERTH Postcode 6000 CBD location, Latitude and longitude values of this CITY DELIVERY CENTRE are -31.924073999999997, 115.91223000000001. A new pandas data frame was then created with the top venues for all the Perth suburbs listed in the original data frame.

The Perth Metro Venues were grouped by venue categories, for instance, there were 208 cafes. One hot encoding was then used because machine learning algorithms cannot work with categorical data directly. The top venues categories for each suburb were grouped and then displayed in a data frame. This data frame was then used to cluster similar suburbs by k-means clustering and the result presented in a folium generated map. This included a list of all the suburbs and what cluster they were in. A multiple linear regression plot was developed of two independent variables, distance from CBD, and distance from coast, to potentially predict the median house price of the different suburbs (dependant variable). A sample of the Perth suburb geolocation data (n=79) was taken randomly (random number generator) and a new pandas data frame was constructed with the median house price data. These distances were calculated from the geolocation data for the different suburbs by a haversine calculation. The haversine program used to calculate these distances was checked with known data. The realestate median house price data was obtained from [RealEstate.com.au](https://www.realestate.com.au/). The R² value was calculated for the multiple linear regression plot and 3D plots were generated using MATPLOTLIB and PLOTLY Express





Discussion Neighbourhood Data

North and South of the River Swan

Most of the [light industrial areas](#) I have found listed are south of the Swan River apart from, Bayswater, Osborne Park, Malaga, Joondalup, Wangara, Balcatta and Neerabup. This reinforces the idea that North of the Swan river is more suburban than the South. South of the "River" has a well developed road system including the Kwinana Freeway and the Roe and Tonkin Hwys, especially the Roe Hwy, give quick transport time between suburbs that are not close to the CBD and are easily accessed from most suburbs. Traffic snarl ups around the CBD during rush hours are to be avoided so if your job or business or contracting business is North or South of the River it is advised you live there as well. Obviously, if you work in the CBD it is best you have the option of catching a [train](#) to work; this also expands your options of where to live.

Neighbourhoods South of the River

The suburbs south of the river are generally grouped into four areas. The majority of the suburbs are in a triangle from Fremantle to Maddington and Gosnells and North to Belmont. There are groups of suburbs around the cities of Armadale, Rockingham and Mandurah. Armadale is not recommended as a place to live. The commute times are longer in Mandurah. The first group mentioned has the highest price suburbs, which also have the most amenities so if you can afford a house here live here, for instance, Applecross has food outlets or restaurants in the top 6 venues including a Japanese, Mediterranean and an Italian Restaurant. There are cheaper suburbs such as Gosnells and Maddington, these will be discussed below. These leaves Rockingham and Kwinana (most suburbs in cluster 2). Apart from Eastern European Restaurants and Falafel Restaurants, there appears to be no other ethnic eateries in the top 10 venues in the Rockingham Kwinana area. They do exist since I have seen Asian food offered in the Council Street shopping complex in Rockingham. Falafel Restaurants appear to be everywhere in the Greater Perth area. The top food outlets seem to be of the fast-food variety. [Baldvis](#) which is a rapidly expanding area, after a fast-food outlet, the top venues are a department store and a hardware store. Therefore, a good place to actually buy a house and get set up; Baldvis is also close to the Kwinana Freeway.

Maddington and Gosnells which have median house prices of 335000 and 320000 respectively are just two of the cheapest suburbs for housing price south of the river and closer to the CBD, they also have [train stations](#). Maddington is the better choice it has a sandwich place and Dumpling Restaurant in the top 10 venues which is a good sign, while Gosnells has a liquor outlet as the third top venue which is a bad sign. While talking about liquor out lets the suburbs which have them in the top 10 venues are -

Merriwa, Marmion, Gosnells, North Beach, Scarborough, Pinjarra, High Wycombe, Gooseberry Hill, Como, Manning, Bently South, Mount Pleasant, Kwinana Town Centre, Melville, Duncraig, Kelmscott, Forrestfield, Kenwick, Noranda, Karawara, Bull Creek, Bertram, Baldvis, Willagee Central, Coolbelup, Innaloo, Padbury and Beldon. (11/7/2021)

Melville, West Leederville, Duncraig, Mount Pleasant, Como, Marmion, North beach, Padbury, Wembley Downs, Colbellup, Bertram, Carlisle North, Carlise, Kelmscott, Merriwa, Beldon

South Fremantle, Cottesloe, Peppermint Grove, Wannero, Willagee Central, South Lake, Shelley, Parkwood, Bassendean, Gosnells, Forrestfield, High Wycombe, Middle Swan, Gooseberry Hill, Pinjarra (12/7/2021)

All these suburbs are in clusters 2 and 3 for the 11 July and in cluster 2, 3 and 4 on 12 July. For the 11 of July the two suburbs with liquor stores as the top venue are Mount Pleasant and Kwinana Town Centre, both in cluster 3. For the 12 July the four suburbs with liquor stores as the top venue are Mount Pleasant, Cluster 2, Pinjarra and Shelley, Cluster 4 and Wembley Downs, Cluster 3. Finally, Midland, which has one of the lowest median house prices in Perth and surrounds should not be ignored. Midland is close to the Swan Valley [wineries](#)() and has Japanese and Asian Restaurants in the top 10 venues also a steak house and Movie House. Midlands most famous [son](#).

Neighbourhoods North of the River

The far northern suburbs of Perth appears to be where most of the urban development is especially with the extension of the Mitchell Freeway. These suburbs have sprung up rapidly and maybe short of amenities found elsewhere. The suburbs on the coast like Scarborough, Cottesloe, Hillarys and Ocean Reef have expensive realstate. For instance, Hillarys has a medium house price of 900000 AUD. The adjacent suburb away from the coast, Padbury, has a medium house price of about 555000 which is more reasonable. Both these suburbs are in Cluster 2. Hillarys has a super market and a shopping mall as the top venues while Padbury has a Pizza place and a fast food restaurant. Hillarys has a [marina](#) as does [Ocean Reef](#) further North, which is still under development, median house price 775000. The North of Perth is slightly warmer during winter than the southern suburbs and is definitely warmer than the south west. This is balanced by the fact of have personally seen [snakes](#) warning signs on beaches in addition to warnings about rogue waves. This should be checked on a suburb by suburb basis. [Portugese man o' war](#) or bluebottles can also be found washed up on some of the beaches. Further Inland, the Ellenbrook, Vines, Aveley group of suburbs in the Swan Valley is worth mentioning. Ellenbrook has a Thai, Eastern European and Chinese Restaurants in the top 10 venues with a Park being the top venue. In addition there are also [wineries](#) in the Swan Valley. The Ellenbrook medium house price is 395000 AUD. These suburbs are now well serviced by road and it is easy to drive to the coast via the Tonkin and Reid Hwys.

Cluster Number 5

This cluster has construction and landscaping as the 2 top venues in all suburbs. (11 July)

Discussion Housing Data

Housing Price data from [RealEstate.com.au](https://www.realestate.com.au) on the 2 JULY 2021

Linear regression gave very low R^2 values when plotting house prices based on approximate distance from the coast, North and South of the Swan River and distance from the Perth CBD. For instance, in the Midland Delivery area, Midland has the lowest mean house price of all the suburbs in Perth, but it is surrounded by far more affluent suburbs. The most affluent suburb is Cottesloe with a median house price of over 2 million dollars which is on the coast and near the CBD. Traditionally, in Perth they talk about "North and South of the River" North of the River Swan having the more affluent suburbs, this certainly does not seem to be the case in 2021 since "Affluent Suburbs" seem to be equally distributed all over the Greater Perth Metropolitan Area. There are only three suburbs in the whole of Perth (in my $n = 79$ sampled data set) with an average house price of less than 300000 dollars, Midland, Greenfields in Mandurah and COOLOONGUP in Rockingham. I have driven through Greenfields along Lakes Road many times and it does not look poor all the houses are in good condition. Mandurah, which is a popular city for fly in fly out mining workers (FIFO) has most of the cheaper suburbs being the furthest from the CBD it is also on the coast. Mandurah has some very attractive looking suburbs and the most expensive are Wannanup and San Remo with a median house prices of 540000 and 575000 dollars, respectively.

WANNANUP SAN REMO

Some factors that probably influence the low R^2 score are suburbs in the Darling Scarp such as [Kalmunda](#) are elevated and have views over the Swan Plain resulting in higher medium house prices despite being a long way from the coast and Perth CBD. In addition, lifestyle blocks on the fringes of Perth such as [Serpentine](#), which is also more than 20 Km from the coast, also effected the result. There are at least two approaches dealing with these issues. The first approach is just to leave the suburbs that are not on the Swan plain out of the regression analysis and don't include suburbs with lifestyle blocks as well. However, this would leave out some very desirable suburbs that people may be attracted to. The second approach would be to use a different model, some ideas for this would be to have categorical data for "In Darling Scarp, yes or no" and "lifestyle blocks in Suburb yes or no" in some sort of linear multivariate model similar to the one already tried. Another possibility is to try a more complex polynomial model.