Project Introduction

After the pandemic in 2020, Australia's house market surged instead of diving as initially predicted by economists and analysts. Meanwhile, the apartment market suffered from migration loss but is recovering mildly. Both homebuyers and property investors are keen to know how the market is going on in different suburbs and where the next opportunity is.

Under this circumstance, the project aims to find out the target suburbs for potential investment opportunities in the residential property market of Australia’s biggest cities. In this report we're making the features more granularly than the state level, we're analysing at the per-state level.

Data Collection

Suburbs from Top 5 cities in population

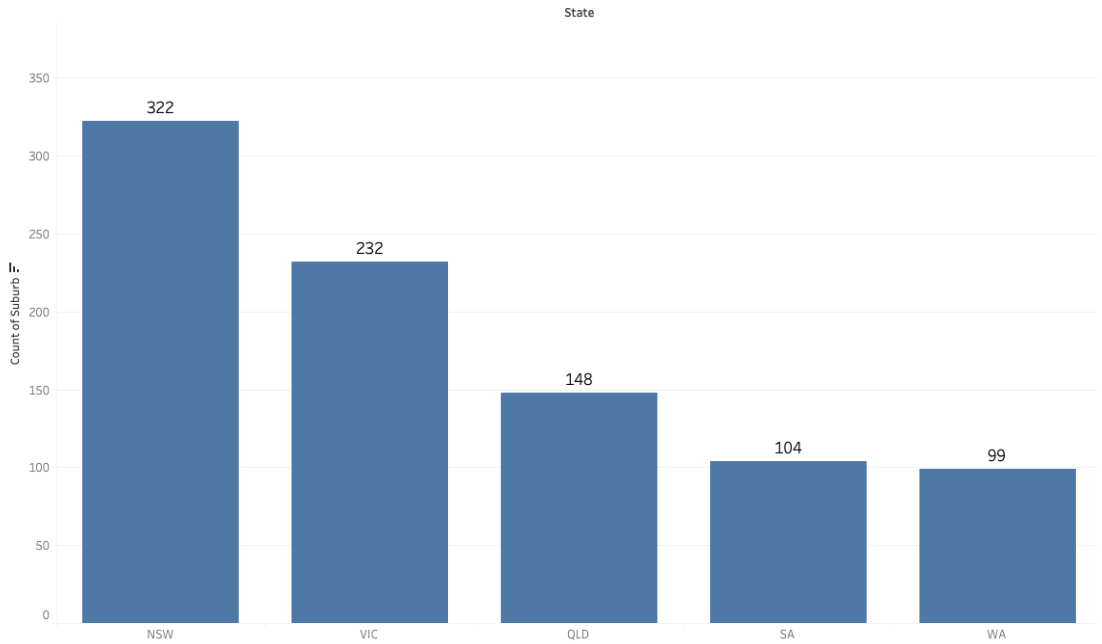
The suburbs selected are in the top 5 cities of Australia by population: Sydney(4.8m), Melbourne(4.7m), Brisbane(2.3m), Perth(2.2m), and Adelaide(1.2m) by 2021. [^1]

In each city, the suburbs selected are by sequential postcodes from the CBD.

Total number of suburbs: 905

- Sydney: 322 (postcode: 2000-2155)
- Melbourne: 232 (postcode: 3000-3140)
- Brisbane: 148 (postcode: 4000-4110)
- Perth: 99 (postcode: 6000-6050)
- Adelaide: 104 (postcode: 5000-5040)

Sample Suburbs by State



Sample properties from [domain.com.au (domain.com.au)]

The sample properties are extracted from domain.com.au by 23 Jul 2021, with 653,022 records in total.

It covers all the sold houses and apartments(incl. apartment, unit, flat) in history in each selected suburb.

The properties with a sold price under $200,000 are removed from our sample set.

The suburbs with less than 30 sold properties in history is removed due to lack of samples.

Report Assumption

The report assumes that domain.com.au represents the residential property sales market in Australia. Domain is one of the two dominant players in the residential property marketplace in Australia (incl. online and offline). It has the most listings and potential customers.[^2]

The report takes the buyers' budget as the key factor for making a property investment. The other influencing factors such as rental yield, train station, shopping centre, schools, etc. are not included at this stage.

The report takes the distance to CBD as an important consideration. Though the covid virus pandemic is influencing home buyers' decisions, the location still has its value in the long term from the author's view.

Investors are target clients for this report.

The median prices in 2021 are based on the sold properties up until 23 Jul 2021.

Technologies for Analytics

1. **Python**: for scraping, parsing, cleansing, and loading data
2. **Pandas**: for Data ingestion and simple transformation
3. **SQL**: for data querying in database, transforming and analysing
4. **Jupyter Notebook**: for data analysis and reporting
5. **PowerBI/Tableau**: for data visualisation and ad-hoc analysis
6. **Docker**: deploying analysing environment

For technical details please see this notebook: [domain.ipynb (https://github.com/helengtt/properties-investment/blob/main/ipynb/domain.ipynb)](https://github.com/helengtt/properties-investment/blob/main/ipynb/domain.ipynb)

House Market Analysis

As of the pandemic from 2020, the House market and Apartment market are two different stories in Australia's residential property market. While house prices keep surging after the first lockdown, apartment prices bounced mildly after reaching the bottom. Therefore, the report will analyse the two markets separately for investors.

Everyone knows picking up suitable suburbs from all the suburbs can be an overwhelming problem for all investors. So we'll start with something every single investor would consider, which is budgeting. Different investors with different budgeting will have very different choices of suburbs.

Profile Assumption: House Investor with $2M Budget

For house market, we'll take the following steps to shorten the suburbs list and assist investors to find out the target suburbs. We assume a budget of $2M for the house investor here.

1. Suburbs' Affordability

Here we measure affordability by the percentage of the houses sold within the budget range in each suburb.

The calculation is based on the houses sold in the past year because the price grew fast after 1 Jul 2020.

For the house investors, we select the house sold at the price between $1.5M and $2M to meet their budget.

In [2]:

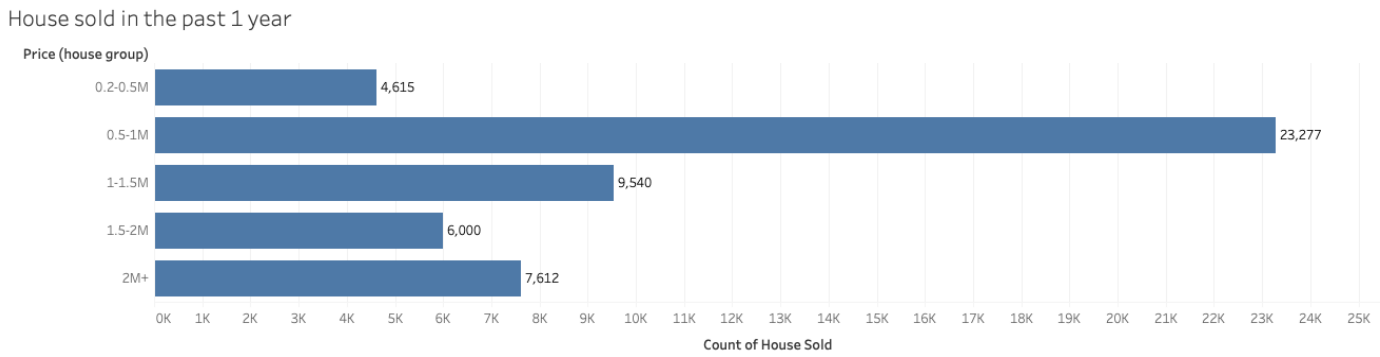
```
# %load_ext sql
# %sql postgresql://postgres:password@this_postgres
%sql SELECT count(suburb_id) as suburbs, sum(ttl_house_sold) as total_house_
```

```
 * postgresql://postgres:***@this_postgres
1 rows affected.
```

Out[2]:

suburbs	total_house_sold	house_sold_2m
905	51044	6000

We've found 6,000 out of 51,044 houses sold between $1.5M and $2M in the 905 selected suburbs.



Source: Domain

Next, we'll filter the suburbs in the following criteria:

- Affordability over 30%. From the author's perspective, it doesn't make sense to choose the suburbs where you can only buy less than 30% of the properties because you wouldn't have many choices.
- Minimum 10 houses were sold in the price range in the past year.

In [4]:

```
combined_suburbs where house_affordability_tier2 >= 0.3 and ttl_house_tier2
```

```
 * postgresql://postgres:***@this_postgres
5 rows affected.
```

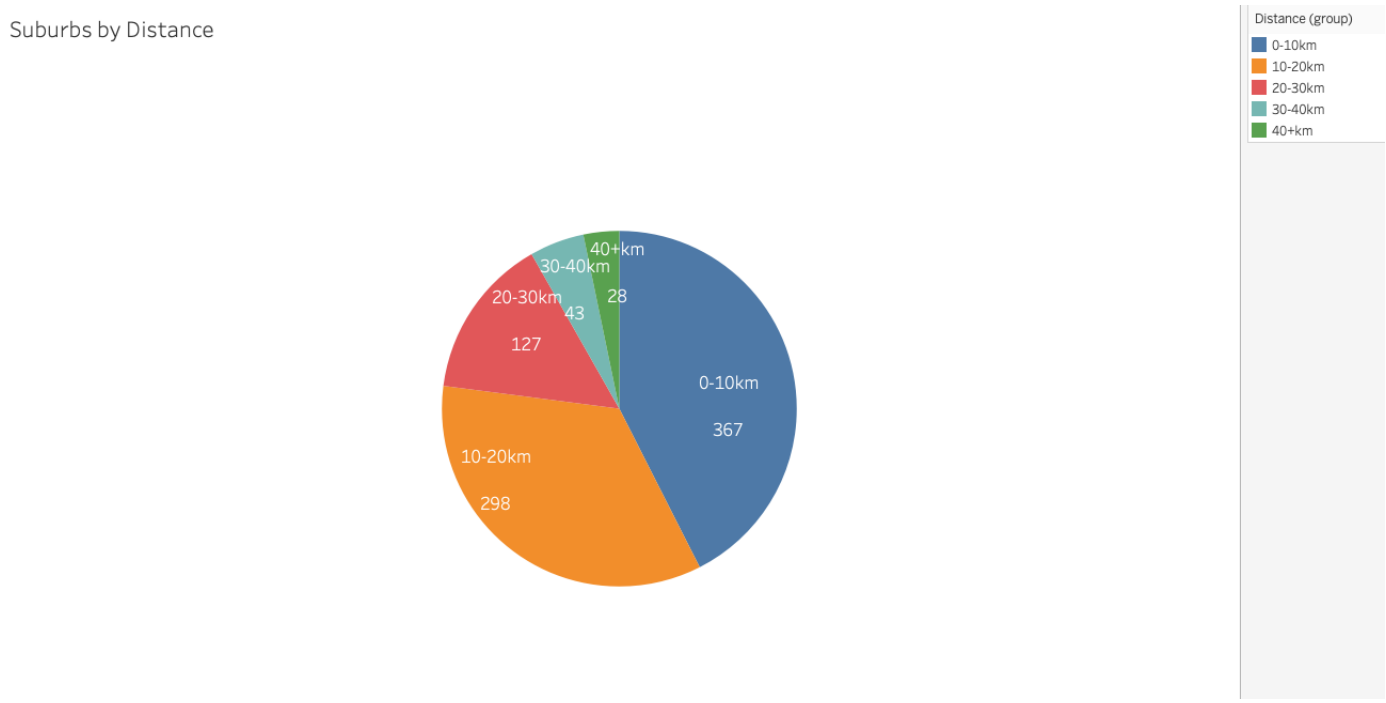
Out[4]:

suburb_id	ttl_house_tier2	house_affordability_tier2
castle-hill-nsw-2154	159	0.473
kellyville-nsw-2155	111	0.329
baulkham-hills-nsw-2153	103	0.346
wahroonga-nsw-2076	80	0.351
leichhardt-nsw-2040	78	0.506

We get 97 suburbs out of 905 with at least 30% houses within the investor's budget.

2. Distance to CBD

Now we'll take the distance to CBD as another criterion to filter the suburbs.



Source: Domain, World Bank, Calculation

To simplify the model, we take 15km from each city's CBD as a reasonable measure at this stage.

We can take different distance to CBD for different cities to make it more reasonable in further studies.

In [3]:

```
%%sql
select suburb_id, ttl_house_tier2, house_affordability_tier2, distance_km
from combined_suburbs
where house_affordability_tier2 >= 0.3 and ttl_house_tier2 >=10 and distance
order by distance_km
limit 5
```

```
 * postgresql://postgres:***@this_postgres
5 rows affected.
```

Out[3]:

suburb_id	ttl_house_tier2	house_affordability_tier2	distance_km
surry-hills-nsw-2010	18	0.327	1.838
new-farm-qld-4005	20	0.333	2.305
glebe-nsw-2037	32	0.478	2.450
darlington-nsw-2008	10	0.357	2.979
forest-lodge-nsw-2037	13	0.361	3.026

We get 65 suburbs within 15km in total.

3. Investment by Median Price Growth

Usually, it is reasonable for property investors to expect a minimum 30% growth in 5 years. So we shorten the suburb list with this criteria.

In [5]:

```
%%sql
select suburb_id, house_5yr_growth, house_2017_growth, house_2018_growth, ho
from combined_suburbs
where house_affordability_tier2 >= 0.3 and ttl_house_tier2 >=10 and distance
order by house_5yr_growth desc
limit 5
```

```
 * postgresql://postgres:***@this_postgres
5 rows affected.
```

Out[5]:

suburb_id	house_5yr_growth	house_2017_growth	house_2018_growth	house_2019_g
city-beach-wa-6015	0.892	0.559	-0.061	
allambie-heights-nsw-2100	0.680	0.127	-0.024	
st-lucia-qld-4067	0.632	0.274	-0.116	
killarney-heights-nsw-2087	0.588	0.109	-0.162	
narraweena-nsw-2099	0.538	0.104	-0.119	

Now we get 27 suburbs.

Next, we need to do some risk control. We know Australia's house prices have experienced the largest downturn on record between July 2017 and May 2019 as of the tightened lending conditions and changes of loan limits to overseas investors by the government. Investors would avoid losing money while seeking a high ROI suburb for investment. Thus we remove the suburbs with a decrease of over 10% in 2018 and 2019.

In [6]:

```
%%sql
select suburb_id, house_5yr_growth, house_2017_growth, house_2018_growth, ho
from combined_suburbs
where house_affordability_tier2 >= 0.3 and ttl_house_tier2 >=10 and distance
order by house_5yr_growth desc
limit 5
```

```
 * postgresql://postgres:***@this_postgres
5 rows affected.
```

Out[6]:

suburb_id	house_5yr_growth	house_2017_growth	house_2018_growth	house_2019_grc
city-beach-wa-6015	0.892	0.559	-0.061	-0
allambie-heights-nsw-2100	0.680	0.127	-0.024	-0
floreat-wa-6014	0.538	0.172	0.059	-0
beacon-hill-nsw-2100	0.483	0.167	-0.080	-0
marsfield-nsw-2122	0.442	0.043	-0.071	-0

Now 12 suburbs are left for the investors from 905 suburbs in the top 5 cities.

There are 8 out of 12 suburbs from NSW, 2 from West Australia, 1 from Victoria and 1 from Queensland.

Source: Domain, World Bank, Calculation

4. Result Table

In [7]:

```
%%sql
select c.suburb_id, s.suburb, s.state, s.population, s.median_income
    , house_5yr_growth, house_2017_growth, house_2018_growth, house_2019_gro
from combined_suburbs c
left join suburbs s on c.suburb_id = s.suburb_id
where house_affordability_tier2 >= 0.3 and ttl_house_tier2 >=10 and distance
order by house_5yr_growth desc
limit 5
```

```
 * postgresql://postgres:***@this_postgres
5 rows affected.
```

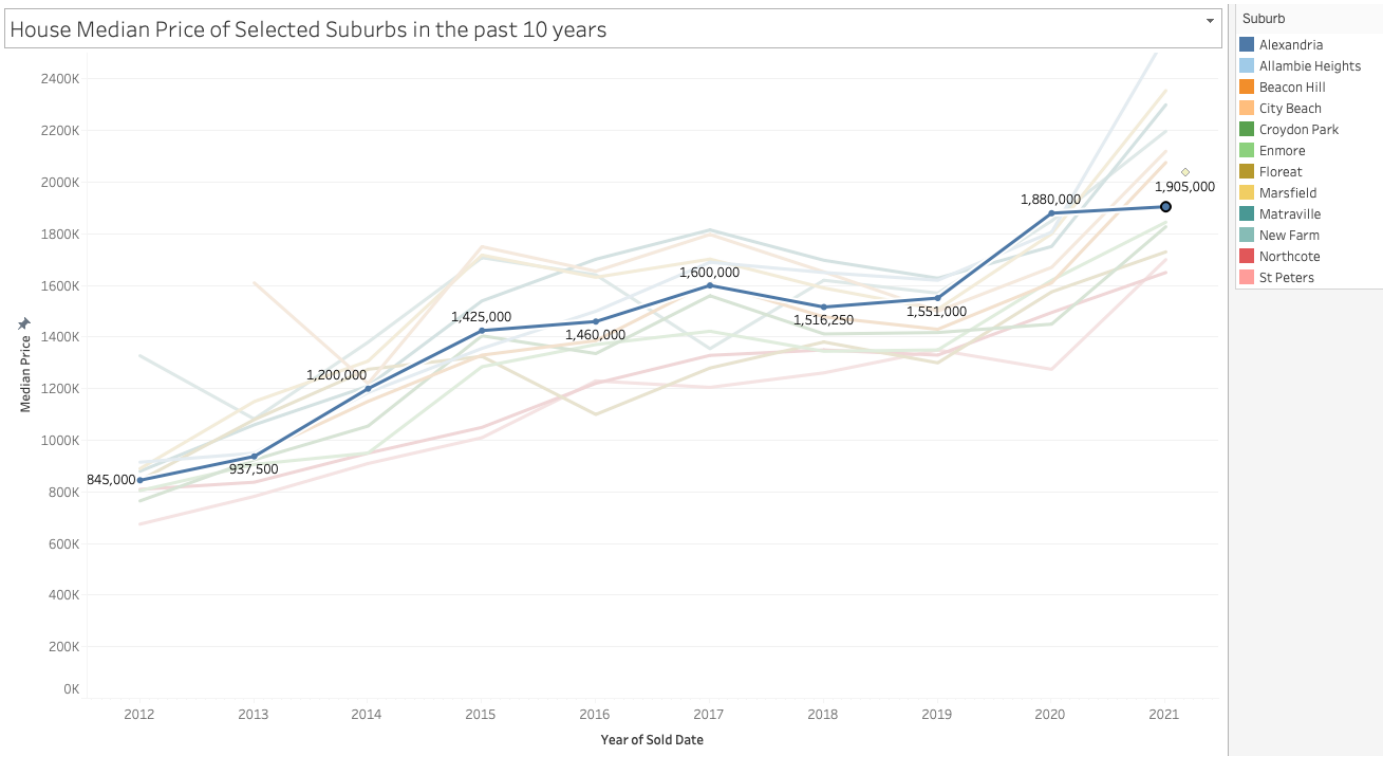
Out[7]:

suburb_id	suburb	state	population	median_income	house_5yr_growth	house_201
city-beach-wa-6015	City Beach	WA	6609	53196	0.892	
allambie-heights-nsw-2100	Allambie Heights	NSW	7009	40404	0.680	
floreat-wa-6014	Floreat	WA	7899	54184	0.538	
beacon-hill-nsw-2100	Beacon Hill	NSW	7457	41288	0.483	
marsfield-nsw-2122	Marsfield	NSW	13297	32448	0.442	

Source: Domain, World Bank, Calculation

As the table above, we add the most features we think are needed to support as many investors as possible for their decision-making.

At this stage, different investors have different preferences. For example, the line chart below shows Alexandria's house median price trend. From 2020 to 2021, the median price didn't grow as fast as the other selected suburbs. Thus, aggressive investors may not invest in this suburb, while defensive investors would see it as a safer investment option and expect higher growth later.



Conclusion

The house selecting model successfully picked up 12 suburbs finally. However, this is just an example. We still have more features like rental yield, percentage by bedrooms and monthly growth, etc. in [domain.ipynb (https://github.com/helengtt/properties-investment/blob/main/ipynb/domain.ipynb)](https://github.com/helengtt/properties-investment/blob/main/ipynb/domain.ipynb). Anyone who knows SQL querying can adjust the parameters above or add their own features to select their desirable suburb(s). Then the investors can decide in which suburbs they'll inspect properties. If we enrich the dataset with other relevant features and data, we would be able to add more value.

Apartment Market Analysis

The apartment market has experienced a heavy strike after the pandemic, in contrast to the white-hot house market. However, family-friendly apartments in medium and low-rise buildings still have their value as an affordable alternative to houses in popular areas. With capital growth in houses outperforming apartments so far this year, the increasing demand for these more affordable apartments is moving forward.

For investors, the apartment price is at its low level at the moment. We could find some cost-effective opportunities if we select carefully and patiently.

Profile Assumption: Apartment Investor with $800K Budget

We assume a budget of $800K for the apartment investor here. Let's look at the affordability first.

1. Suburbs' Affordability

For the apartment investors, we select the apartments from $500K to $800K to meet their budget. The affordability is based on the apartments sold in the past year.

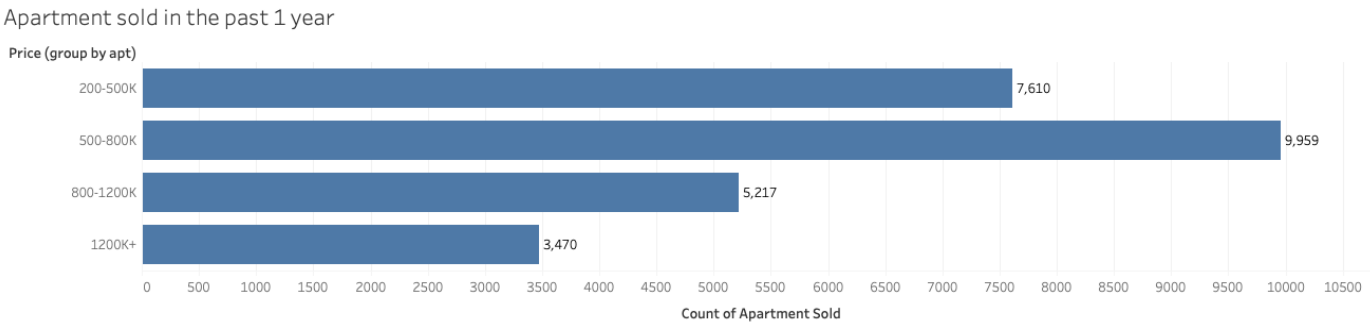
In [6]:

```
# %sql SELECT count(suburb_id) as suburbs, sum(ttl_apt_sold) as total_apt_sc
aff_apt = ('SELECT count(suburb_id) as suburbs, sum(ttl_apt_sold) as total_a
```

```
 * postgresql://postgres:***@this_postgres
1 rows affected.
```

Out[6]:

suburbs	total_apt_sold	apt_sold_tier1
905	26256	9959



Source: Domain

We've found 9,959 out of 26,256 apartments sold in $500K - $800K in the 905 selected suburbs.

Next, we look into the affordability. We'll filter the suburbs in the following criteria:

- Affordability over 30%.
- Minimum 20 apartments were sold in the price range in the past year.

In [8]:

```
1 from combined_suburbs where apt_affordability_tier1 >= 0.3 and ttl_apt_tier
```

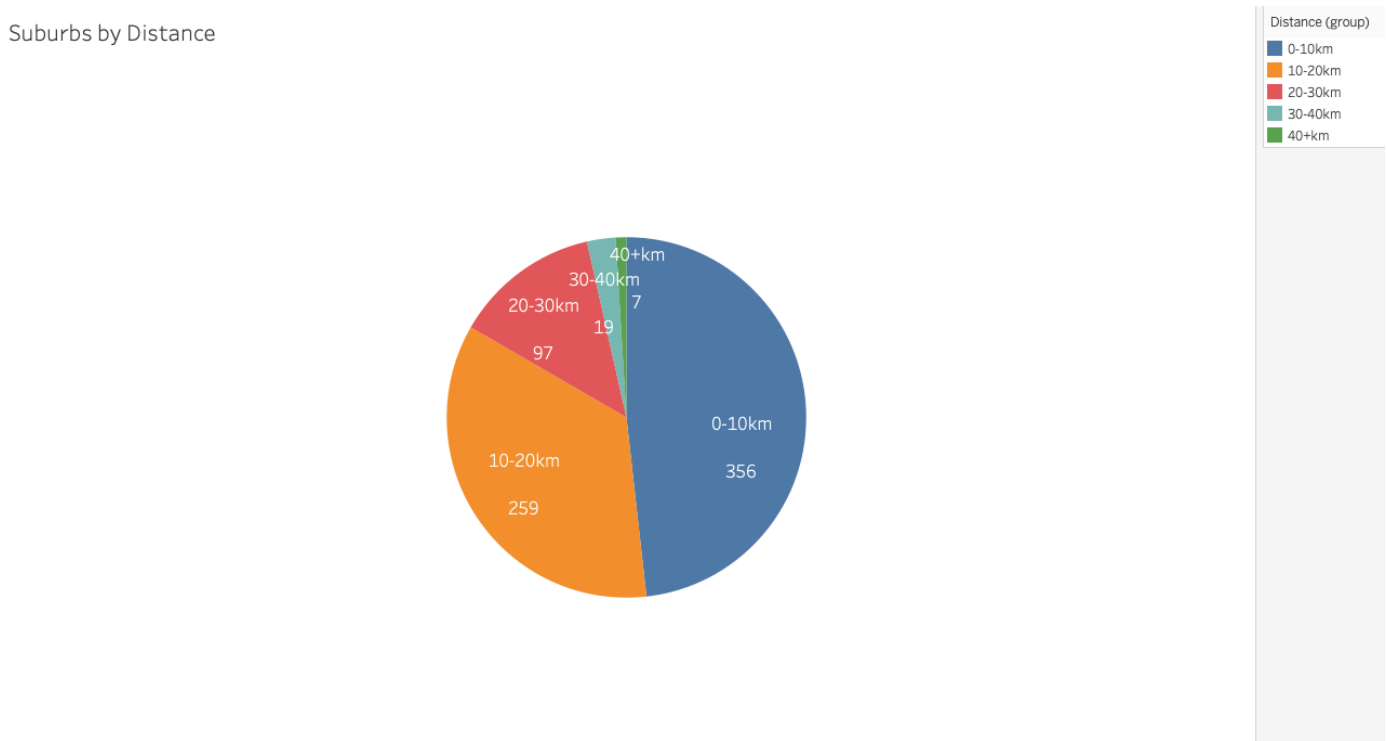
```
 * postgresql://postgres:***@this_postgres
5 rows affected.
```

Out[8]:

suburb_id	ttl_apt_tier1	apt_affordability_tier1
wentworth-point-nsw-2127	249	0.819
melbourne-vic-3000	227	0.374
southbank-vic-3006	174	0.489
ashfield-nsw-2131	163	0.674
hornsby-nsw-2077	162	0.876

We get 123 affordable suburbs out of 905, each of which has at least 30% apartments affordable.

2. Distance to CBD



Source: Domain, World Bank, Calculation

Then we take 15km from CBD as a reasonable measure at this stage.

In [9]:

```
%%sql
select suburb_id, ttl_apt_tier1, apt_affordability_tier1, distance_km
from combined_suburbs
where apt_affordability_tier1 >= 0.3 and ttl_apt_tier1 >=20 and distance_km
order by distance_km
limit 5
```

```
 * postgresql://postgres:***@this_postgres
5 rows affected.
```

Out[9]:

suburb_id	ttl_apt_tier1	apt_affordability_tier1	distance_km
brisbane-city-qld-4000	86	0.305	0.000
melbourne-vic-3000	227	0.374	0.000
southbank-vic-3006	174	0.489	1.101
south-brisbane-qld-4101	42	0.318	1.167
ultimo-nsw-2007	23	0.511	1.661

We get 92 suburbs within 15km in total.

3. Investment by Median Price Growth

As for the gloomy apartment market from 2017, the first thing we do to help apartment investors is to avoid a big loss. As an investment, we think the annual loss within 5% could be acceptable. So we filter the annual growth under -5% in 2018, 2019 for the policy intervention and 2020, 2021 for the Covid virus impact.

In [10]:

```
%%sql
select suburb_id, apt_5yr_growth, apt_2017_growth, apt_2018_growth, apt_2019
    , ttl_apt_tier1, apt_affordability_tier1, distance_km
from combined_suburbs
where apt_affordability_tier1 >= 0.3 and ttl_apt_tier1 >=20 and distance_km
order by apt_5yr_growth desc
limit 5
```

```
 * postgresql://postgres:***@this_postgres
5 rows affected.
```

Out[10]:

suburb_id	apt_5yr_growth	apt_2017_growth	apt_2018_growth	apt_2019_growth	apt_2
lutwyche-qld-4030	0.446	-0.097	0.168	0.013	
fitzroy-north-vic-3068	0.446	0.196	-0.004	0.040	
reservoir-vic-3073	0.395	0.156	0.060	-0.038	
pascoe-vale-vic-3044	0.394	0.040	0.117	0.010	
glenroy-vic-3046	0.372	0.064	0.133	0.018	

We get 30 suburbs left.

Next, We still seek a minimum 30% capital growth in the past 5 years for the apartment market.

In [11]:

```
%%sql
select suburb_id, apt_5yr_growth, apt_2017_growth, apt_2018_growth, apt_2019
    , ttl_apt_tier1, apt_affordability_tier1, distance_km
from combined_suburbs
where apt_affordability_tier1 >= 0.3 and ttl_apt_tier1 >=20 and distance_km
order by apt_5yr_growth desc
limit 5
```

```
 * postgresql://postgres:***@this_postgres
5 rows affected.
```

Out[11]:

suburb_id	apt_5yr_growth	apt_2017_growth	apt_2018_growth	apt_2019_growth	apt_2
lutwyche-qld-4030	0.446	-0.097	0.168	0.013	
fitzroy-north-vic-3068	0.446	0.196	-0.004	0.040	
reservoir-vic-3073	0.395	0.156	0.060	-0.038	
pascoe-vale-vic-3044	0.394	0.040	0.117	0.010	
glenroy-vic-3046	0.372	0.064	0.133	0.018	

Now 8 suburbs are left for the investors from 905 suburbs in the top 5 cities.

We can see 7 of 8 suburbs are in Victoria and the 1 left is in Queensland. But Lutwyche in Queensland has the highest growth with our criteria.

4. Result Table

In [12]:

```
%%sql
select c.suburb_id, s.suburb, s.state, s.population, s.median_income
    , apt_5yr_growth, apt_2017_growth, apt_2018_growth, apt_2019_growth, apt
    , ttl_apt_tier1, apt_affordability_tier1, distance_km
from combined_suburbs c
left join suburbs s on c.suburb_id = s.suburb_id
where apt_affordability_tier1 >= 0.3 and ttl_apt_tier1 >=20 and distance_km
order by apt_5yr_growth desc
limit 5
```

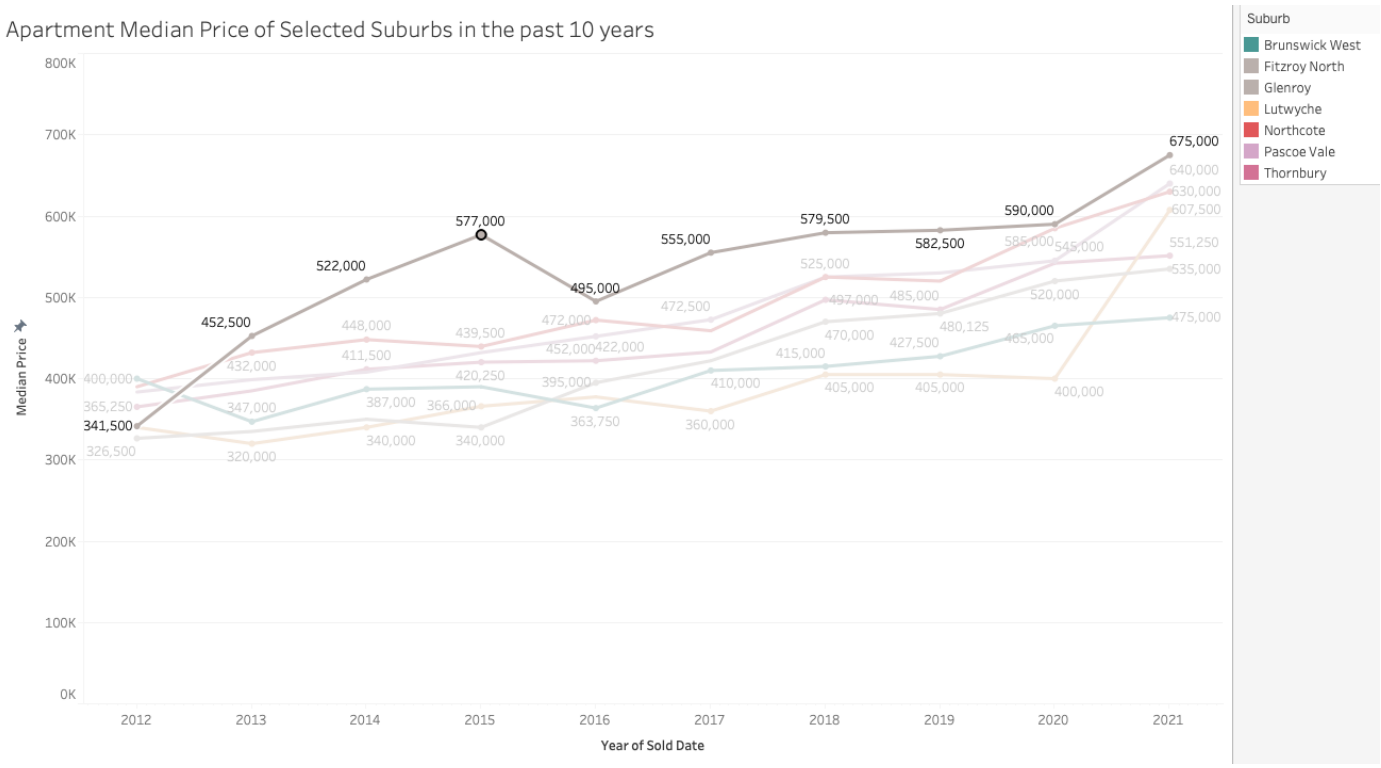
```
 * postgresql://postgres:***@this_postgres
5 rows affected.
```

Out[12]:

suburb_id	suburb	state	population	median_income	apt_5yr_growth	apt_2017_gro
fitzroy-north-vic-3068	Fitzroy North	VIC	12339	52052	0.446	0.
lutwyche-qld-4030	Lutwyche	QLD	3454	47476	0.446	-0.
reservoir-vic-3073	Reservoir	VIC	50474	28132	0.395	0.
pascoe-vale-vic-3044	Pascoe Vale	VIC	17051	38012	0.394	0.
glenroy-vic-3046	Glenroy	VIC	22245	28028	0.372	0.

Source: Domain, World Bank, Calculation

From the selected suburbs above, we take Fitzroy North's apartment median price line as an example. Though the 5-year growth of median price reached 44.6%, there was a big drop of -39.6% in 2016. For defensive investors, they try to avoid high volatility and may not invest in this suburb, especially in the gloomy apartment market at the moment. But for aggressive investors, the highest 5 year growth may attract them most.



Source: Domain, Calculation

Conclusion

This project gets data from data.worldbank.org and domain.com.au transform the data and calculates features including:

- Population
- Median Income
- Median Price 2017 -2021
- 5 year growth
- Annual growth 2017-2021
- Distance to CBD

We can have as many features as nessessary to support a particular decision-making process. The features we calculated above are just examples.

Disclaimer

The sole purpose of this research is to provide as many features as possible about a suburb, so that people can find the most appropriate suburb to invest in according to his own condition.

[^1] Source: Geonames (https://www.geonames.org/)

[^2] Source: ibisworld (https://www.ibisworld.com/au/industry/residential-real-estate-advertising/5512/)