Australia Real Estate Market Analysis

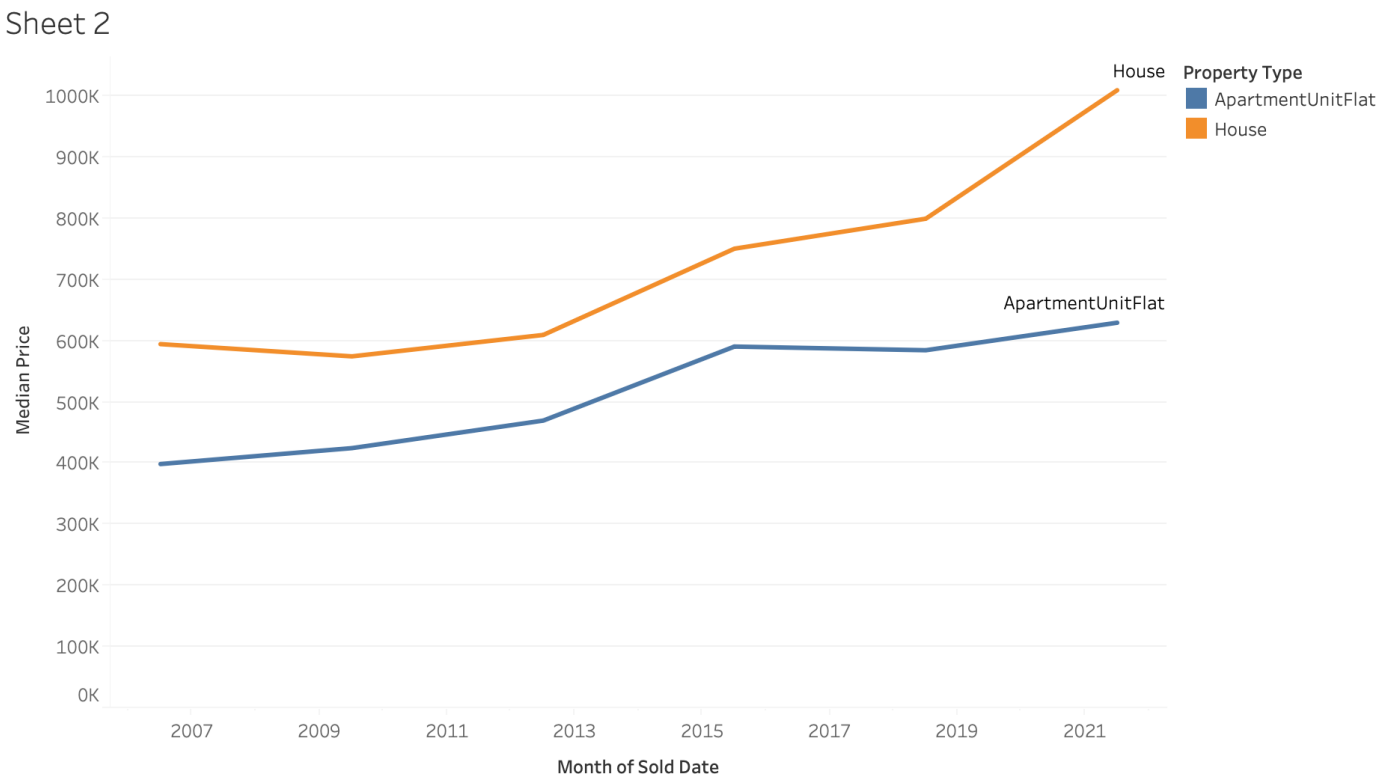
Janie Zhan

1. Problem Statement

1.1 Housing Price Growth in Australia's 3 Main Cities

ANZ economists predicted a 17% rise in house prices across australia in 2021 even with the global pandemic. The housing prices look to be on a rocketing path for almost two decades, attracting investors from all over the world. Strong housing market conditions over the last 15 years have boosted median house values by 412% or $460,000.

The housing market has shown some extraordinary changes over the past fifteen years, 15 years ago, the median apartment/unit value across Australia was just below 400k and houses showed a slightly higher median value, at around 600k. Residential house prices across the capital cities in Australia increased by 7.5 percent from March 2020 to March 2021. Housing affordability in Australia remains a highly political topic with many prospective home buyers feeling priced out of the market.Ratio of prices to income per person is 108.5%. The capital gain over the past 15 years equates to an annual growth rate of 11% for houses.



The trend of median of Price for Sold Date Month. Color shows details about Property Type. The marks are labeled by Property Type. The data is filtered on Sold Date (MY), which keeps 6 of 248 members. The view is filtered on Property Type, which keeps ApartmentUnitFlat and House.

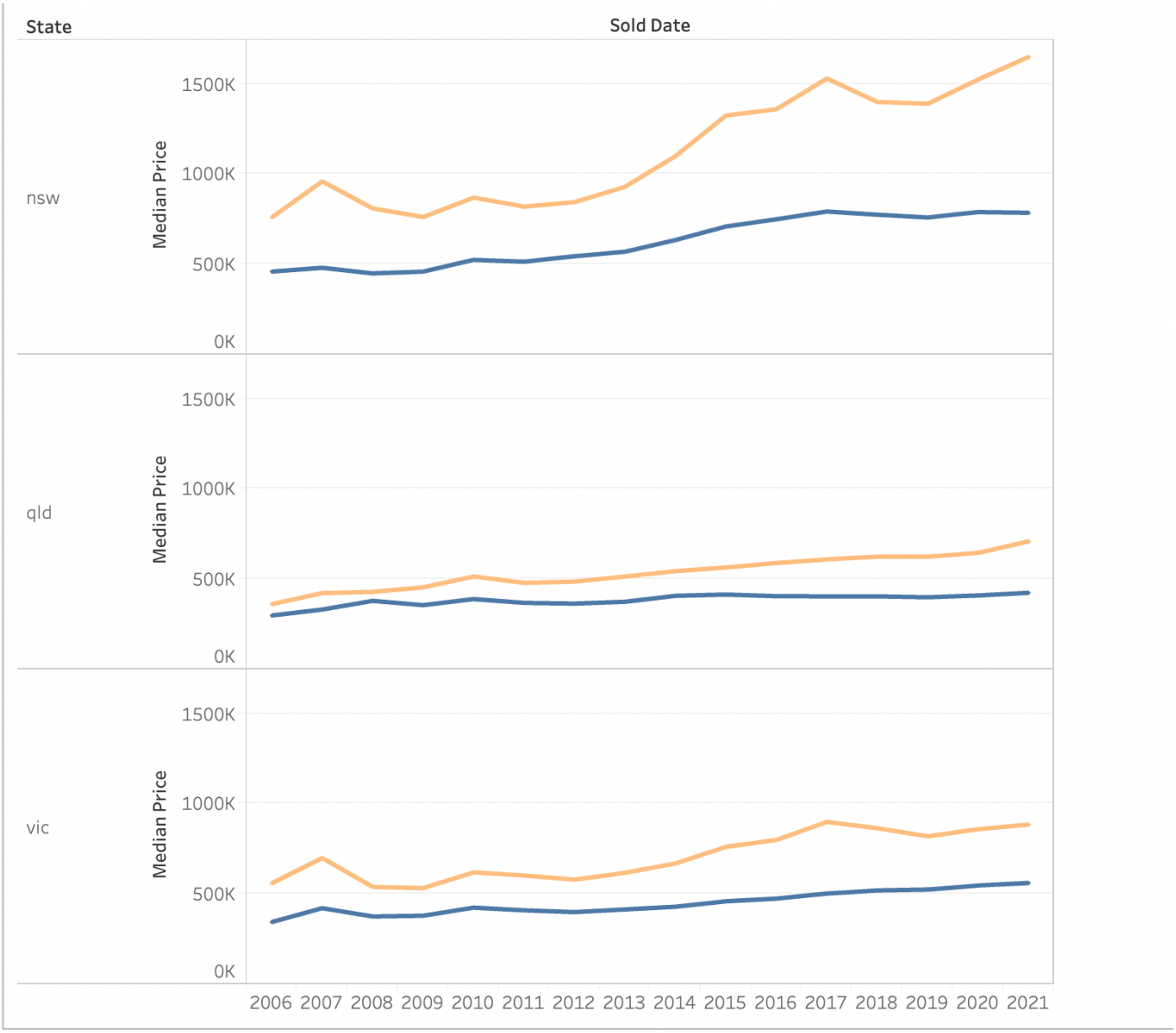
(data retrieved from 3 cities: Sydney, Melbourne, Brisbane)

1.2 Property Investment Criteria

The investment value is considered one of the top factors to be considered for investors and home owners. In this report, we will discuss a list of common selection criteria, develop data models in various dimensions to support decision makings of house buyers.

- 1.Affordability of apartment in budget 500,000, 800,000 and 1000,000 in 2021
- 2.Affordability of house in budget 800,000, 1500,000 and 2000,000 in 2021
- 3.Historical capital growth in each suburb since 2006
- 4.Highest and lowest growth value in the last three years
- 5.Numbers and percentages of Houses/apartments in different suburbs

1.3 Impact of COVID-19 on Australian Property Market



In [ ]:

```
%%sql
DELETE from properties_cleaned where sold_date >='32';
```

In [1]:

```
%load_ext sql
```

In [2]:

```
%sql postgresql://postgres:password@this_postgres/postgres
```

In [3]:

```
%%sql
Drop table properties_affordability_apart;
create table properties_affordability_apart as
WITH recent_apartment_sales AS (
    SELECT
          suburb
        , price
        , CASE WHEN price <= 500000 THEN 1 ELSE 0 END AS p_0_500k
        , CASE WHEN price <= 800000 THEN 1 ELSE 0 END AS p_0_800k
        , CASE WHEN price <= 1000000 THEN 1 ELSE 0 END AS p_0_1m
      FROM properties_cleaned
     WHERE sold_date like '%2021%' and property_type iLIKE 'Apartment%'
), grouped AS (
    SELECT suburb
        , COUNT(1) total_sales
        , SUM(p_0_500k) count_0_500k
        , SUM(p_0_800k) count_0_800k
        , SUM(p_0_1m) count_0_1m
        , SUM(p_0_500k)*100.0/count(1) as affordability_500k
        ,SUM(p_0_800k)*100.0/count(1) as affordability_800k
        ,SUM(p_0_1m)*100.0/count(1) as affordability_1m
     from recent_apartment_sales
     group by suburb
     having count(1)>=30
)
SELECT * FROM grouped;
```

...

In [4]:

```
%%sql
select * from properties_affordability_apart limit 10;
```

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

Out[4]:

suburb	total_sales	count_0_500k	count_0_800k	count_0_1m	affordability_500k	
taringa-qld-4068	53	43	50	51	81.1320754716981132	94.3
hawthorn-east-vic-3123	41	11	37	41	26.8292682926829268	90.2
guildford-nsw-2161	31	26	31	31	83.8709677419354839	100.0
chippendale-nsw-2008	35	3	21	27	8.5714285714285714	60.0
murrumbeena-vic-3163	38	18	31	35	47.3684210526315789	81.5
pascoe-vale-vic-3044	41	10	36	40	24.3902439024390244	87.8
bentleigh-east-vic-3165	30	5	20	26	16.6666666666666667	66.6
woollahra-nsw-2025	39	0	2	8	0E-20	5.1
maroubra-nsw-2035	70	1	19	43	1.4285714285714286	27.1
brunswick-vic-3056	125	49	115	123	39.2000000000000000	92.0

In [5]:

```
%%sql
Drop table properties_affordability_house;
create table properties_affordability_house as
WITH recent_house_sales AS (
    SELECT
          suburb
        , price
        , CASE WHEN price <= 800000 THEN 1 ELSE 0 END AS p_0_800k
        , CASE WHEN price <= 1500000 THEN 1 ELSE 0 END AS p_0_1500k
        , CASE WHEN price <= 2000000 THEN 1 ELSE 0 END AS p_0_2m
      FROM properties_cleaned
     WHERE sold_date like '%2021%' and property_type iLIKE 'House%'
), grouped AS (
    SELECT suburb
        , COUNT(1) total_sales
        , SUM(p_0_800k) count_0_800k
        , SUM(p_0_1500k) count_0_1500k
        , SUM(p_0_2m) count_0_2m
        , SUM(p_0_800k)*100.0/count(1) as affordability_800k
        ,SUM(p_0_1500k)*100.0/count(1) as affordability_1500k
        ,SUM(p_0_2m)*100.0/count(1) as affordability_2m
     from recent_house_sales
     group by suburb
     having count(1)>=30
)
SELECT * FROM grouped;
```

...

In [6]:

```
%%sql
select * from properties_affordability_house limit 10;
```

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

Out[6]:

suburb	total_sales	count_0_800k	count_0_1500k	count_0_2m	affordability_800k	a
eltham-north-vic-3095	39	2	35	39	5.1282051282051282	89.7
cherrybrook-nsw-2126	74	1	5	41	1.3513513513513514	6.7
roxburgh-park-vic-3064	146	135	146	146	92.4657534246575342	100.(
daisy-hill-qld-4127	41	29	40	41	70.7317073170731707	97.5
frenchs-forest-nsw-2086	61	0	1	24	0E-20	1.(
guildford-nsw-2161	119	23	114	118	19.3277310924369748	95.7
heritage-park-qld-4118	38	38	38	38	100.0000000000000000	100.(
caroline-springs-vic-3023	152	109	151	151	71.7105263157894737	99.3
newport-nsw-2106	30	0	0	1	0E-20	
lidcombe-nsw-2141	43	1	24	37	2.3255813953488372	55.8

In [7]:

```
%%sql
 select sold_date
        ,  suburb
        ,PERCENTILE_CONT(0.5) WITHIN GROUP(ORDER BY price) as price
    from properties_cleaned
    where sold_date >'2006'
    group by sold_date,suburb
    order by sold_date desc
    limit 10
```

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

Out[7]:

sold_date	suburb	price
31 Oct 2020	albany-creek-qld-4035	622000.0
31 Oct 2020	alexandria-nsw-2015	1400000.0
31 Oct 2020	allambie-heights-nsw-2100	1830000.0
31 Oct 2020	altona-vic-3018	648000.0
31 Oct 2020	annandale-nsw-2038	1614500.0
31 Oct 2020	annangrove-nsw-2156	2200000.0
31 Oct 2020	annerley-qld-4103	900000.0
31 Oct 2020	ascot-vale-vic-3032	1497500.0
31 Oct 2020	ashwood-vic-3147	1250000.0
31 Oct 2020	aspendale-gardens-vic-3195	871000.0

In [8]:

```
%%sql
--Drop table history_capital_growth;
Create table history_capital_growth as
With  history_sales as (
    select date_part('year',sold_date::timestamp) as sold_year
        ,  suburb,property_type
        ,PERCENTILE_CONT(0.5) WITHIN GROUP(ORDER BY price) as price
    from properties_cleaned
    where date_part('year',sold_date::timestamp) >=2006 and property_type ilike 'Hou
    group by date_part('year',sold_date::timestamp),suburb,property_type
),
windowed as (
    select *
         ,lag(price) over (order by sold_year)as p1
         ,lag(price,3) over (order by sold_year) as p3
         ,lag(price,5) over (order by sold_year) as p5
         ,lag(price,8) over (order by sold_year) as p8
         ,lag(price,10) over (order by sold_year) as p10
         ,lag(price,12) over (order by sold_year) as p12
         ,lag(price,15) over (order by sold_year) as p15
    from history_sales
),
final as (
    select *
         ,(price-p1)/p1 as g1
         , (price-p3)/p3 as g3
         , (price-p5)/p5 as g5
         ,(price-p8)/p8 as g8
         ,(price-p10)/p10 as g10
         ,(price-p12)/p12 as g12
         ,(price-p15)/p15 as g15
    from windowed
)
select * from final limit 10;
```

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

Out[8]:

sold_year	suburb	property_type	price	p1	p3	p5	p8	p10
2000.0	burwood-nsw-2134	ApartmentUnitFlat	502000.0	None	None	None	None	None
2000.0	coogee-nsw-2034	ApartmentUnitFlat	730000.0	502000.0	None	None	None	None
2000.0	croydon-nsw-2132	ApartmentUnitFlat	302000.0	730000.0	None	None	None	None
2000.0	strathfield-nsw-2135	ApartmentUnitFlat	800000.0	302000.0	502000.0	None	None	None
2000.0	surrey-hills-vic-3127	ApartmentUnitFlat	147000.0	800000.0	730000.0	None	None	None
2001.0	ashfield-nsw-2131	ApartmentUnitFlat	192000.0	147000.0	302000.0	502000.0	None	None
2001.0	burwood-nsw-2134	ApartmentUnitFlat	920000.0	192000.0	800000.0	730000.0	None	None

	2001.0	concord-nsw-2137	ApartmentUnitFlat	415500.0	920000.0	147000.0	302000.0	None	None
	2001.0	cremorne-nsw-2090	ApartmentUnitFlat	375000.0	415500.0	192000.0	800000.0	502000.0	None
	2001.0	enfield-	ApartmentUnitFlat	230000.0	375000.0	920000.0	147000.0	730000.0	None

In [9]:

```
le house_apartment_percentage;
able house_apartment_percentage as
perty_number AS(
ct suburb
 , CASE WHEN property_type ilike 'House%' THEN 1 ELSE 0 END AS house_sales
 , CASE WHEN property_type ilike 'Apartment%' THEN 1 ELSE 0 END AS apartment_sales
OM properties_cleaned
s (
CT suburb
 ,count(1) as total_sales
 ,sum(house_sales)count_house_sales
 ,sum(apartment_sales)count_apartment_sales
 ,sum(house_sales)*100.0/count(1) as percentage_house_total
 ,sum(apartment_sales)*100.0/count(1) as percentage_apartment_total
 property_number
 by suburb
 as (
CT suburb, count_house_sales,count_apartment_sales,percentage_house_total,percentage_
 when count_apartment_sales > 0 THEN count_house_sales*100.0/count_apartment_sales EI
 sum

 from final;
```

...

In [10]:

```
%%sql
Select * from house_apartment_percentage limit 5;
```

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

Out[10]:

suburb	count_house_sales	count_apartment_sales	percentage_house_total	percentage_apartm
northgate-qld-4013	477	102	81.8181818181818182	17.49571183
roxburgh-park-vic-3064	2006	19	98.9151873767258383	0.936883629191
taringa-qld-4068	395	713	35.2363960749330955	63.60392506
hawthorn-east-vic-3123	775	642	54.1201117318435754	44.83240223
daisy-hill-qld-4127	693	7	98.2978723404255319	0.992907801418

2. Data Source

This report's data is retrieved from "www.domain.com.au" (http://www.domain.com.au"), analysis performed on initial json scripts.

Data models:

- 1013 Australian suburbs, 3 states(QLD,NSW,VIC), across postcode 2000-2200 from ,3000-3200, 4000 - 4200, over 860,000 rows of data.
- General information includes : Suburb,postcode,price,property type, sold date, sold type,numbers of beds, numbers of bathroom, numbers of parkings, land size, address, longtitude and latitude
- libraies : pandas

3. Assumptions

This report assumes that:

- All data retrieved from www.domain.com.au (http://www.domain.com.au) is correct, and excluding from input error and manual error
- The sample size is statistically significant
- The observation of data is closly related [statistic models]

4. Summary

In this report, the input data is a list records of property sales, the purpose of this report is enable the audience make better decisions when choosing the suburbs to invest in out of all suburbs.

5. Conclusions/recommendations

This report aims at enabling target audience to make decisions based on historical facts, in the data analysis world, we called those facts `features` , we have 5 features. Different investors will focus on different features. For example, investors who prefer high historical capital growth would probably look into those suburbs:

In [11]:

```
%%sql
Select suburb from history_capital_growth limit 10;
```

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

Out[11]:

suburb
burwood-nsw-2134
coogee-nsw-2034
croydon-nsw-2132
strathfield-nsw-2135
surrey-hills-vic-3127
ashfield-nsw-2131
burwood-nsw-2134
concord-nsw-2137
cremorne-nsw-2090
enfield-nsw-2136