# Data Analyst Portfolio Project

By: Ilham Prawira Hasbi

# About

- This dataset contains Melbourne housing snapshot from 2016 to 2017.

- The dataset was taken from Kaggle.
- This is a snapshot of a dataset created by Tony Pino. It was scraped from publicly available results posted every week from Domain.com.au . The dataset includes Address, Type of Real estate, Suburb, Method of Selling, Rooms, Price, Real Estate Agent, Date of Sale and distance from C.B.D. (Central Business District).

- Notes on Specific Variables

- Rooms: Number of rooms

- Price: Price in 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PI - property passed in; PN - sold prior not disclosed; SN - sold not disclosed; NB - no bid; VB - vendor bid; W - withdrawn prior to auction; SA - sold after auction; SS - sold after auction price not disclosed.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

# About

Date: Date sold

Distance: Distance from CBD

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

BuildingArea: Building Size

CouncilArea: Governing council for the area

# Business Objective

- Comparing real estate prices by region.

- Determine which suburb has the cheapest prices of real estate.

- Find out the average price of real estate and how far it is from the Central Business District (CBD).

- Determine the impact of distance on real estate prices

- Identify the relationship between real estate land size and price

- Identify annual price trends by real estate type

- Find out real estate type with many amenities, affordable prices and close to CBD

- Discover real estate types with larger than average acreage at affordable prices.

# TOOLS

Data Preparation

DBMS, Analysis

Dashboard visualization

# Data Exploration

# General Statistics

- Query:

```sql
SELECT COUNT(*) total_properties,
  ROUND(AVG(price),0) average_price,
  MAX(price) maximum_price,
  MIN(price) minimum_price,
  ROUND(AVG(bedroom2),2) average_bedroom,
  ROUND(AVG(bathroom),2) average_bathroom,
  ROUND(AVG(car),2) average_car
FROM melb_data;
```

| Avg. Average Bathroom | Avg. Average Bedroom | Avg. Average Car | Avg. Average Price | Max. Maximum Price | Min. Minimum Price | Total Properties |
|---|---|---|---|---|---|---|
| 1,53 | 2,91 | 1,61 | 1.075.684,00 | 9.000.000,00 | 85.000,00 | 13.580,00 |

Total properties are 13.580 ; average property price is $1.075.684 ; highest property price is $9.000.000 ; lowest property price is $85.000 ; average number of bedrooms per property is 2,91 ; average number of bathrooms per property is 1,53 ; and average number of carspots per property is 1,61.

# Real estate prices by region

- Query:

```sql
Select Regionname, ROUND(AVG(Price),0) average_price from melb_data
WHERE Regionname Like 'Northern%'
OR Regionname Like 'Southern%'
OR Regionname Like 'Eastern%'
OR Regionname Like 'Western%'
Group by Regionname
ORDER BY average_price;
```

| | Regionname | average_price |
|---|---|---|
| 1 | Western Victoria | 397523 |
| 2 | Northern Victoria | 594829 |
| 3 | Eastern Victoria | 699980 |
| 4 | Western Metropolitan | 866420 |
| 5 | Northern Metropolitan | 898171 |
| 6 | Eastern Metropolitan | 1104079 |
| 7 | Southern Metropolitan | 1372963 |

The region with the cheapest property prices is Western Victoria and the most expensive property is in the Southern Metropolitan region.

# Real estate prices by region
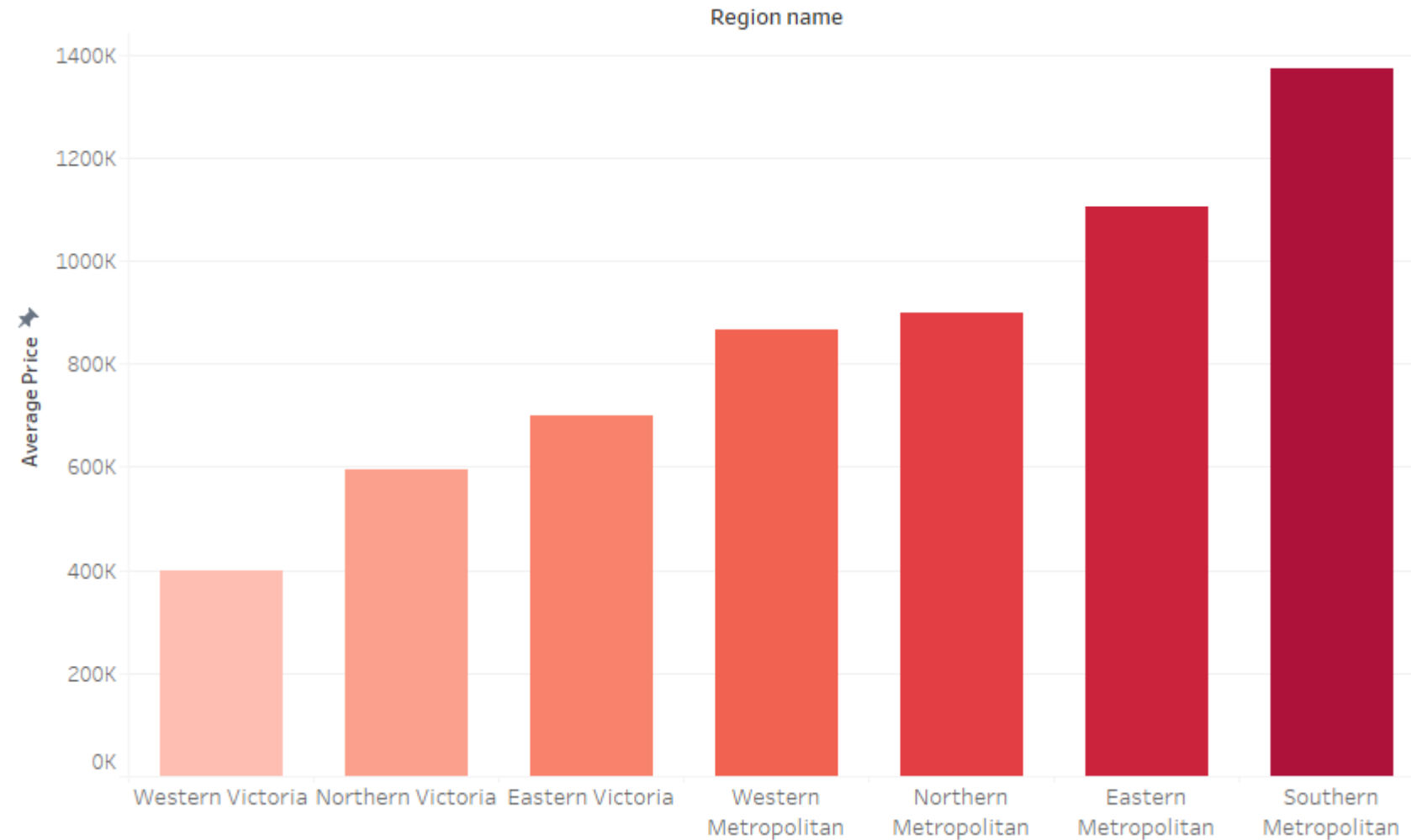
# Suburb that has the cheapest prices of real estate

- Query:

```
select distinct TOP 10 Suburb, MIN(Price) as lowest_price from melb_data
group by Suburb
Order by lowest_price
```

| | Suburb | lowest_price |
|---|---|---|
| 1 | Footscray | 85000 |
| 2 | Caulfield | 131000 |
| 3 | Coburg | 145000 |
| 4 | Albion | 145000 |
| 5 | Hawthorn | 160000 |
| 6 | Brunswick | 170000 |
| 7 | West Footscray | 185000 |
| 8 | Kingsville | 200000 |
| 9 | Melbourne | 210000 |
| 10 | Seddon | 215000 |

The cheapest suburb is Footscray at $85000.

# Suburb that has the cheapest prices of real estate
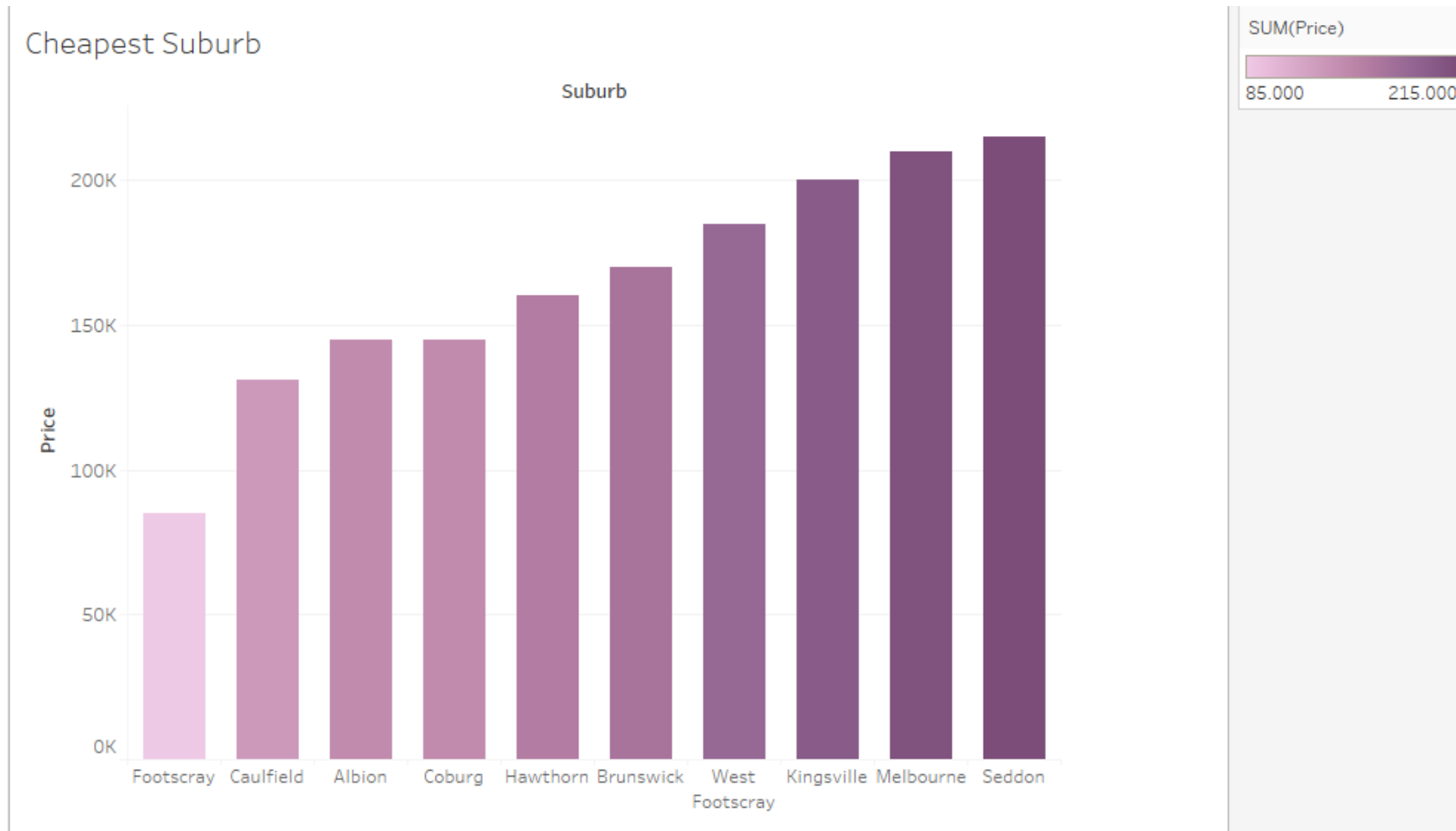
# Average price of distance of Properties to CBD (Central Business District)

- Query:

```sql
SELECT distance_range, ROUND(AVG(Price),0) average_price from(
    SELECT CASE
     WHEN distance < 5 THEN '0-5 km'
     WHEN distance < 10 THEN '5-10 km'
     ELSE '10+ km'
   END AS distance_range, Price
FROM melb_data) melb_data
GROUP BY distance_range
ORDER BY average_price ASC;
```

| | distance_range | average_price |
|---|---|---|
| 1 | 10+ km | 1000017 |
| 2 | 0-5 km | 1123638 |
| 3 | 5-10 km | 1155632 |

The average property price is not affected by the distance to the CBD. Properties 5-10 km away are more expensive than those 0-5 km away.

# Average price of distance of Properties to CBD (Central Business District)
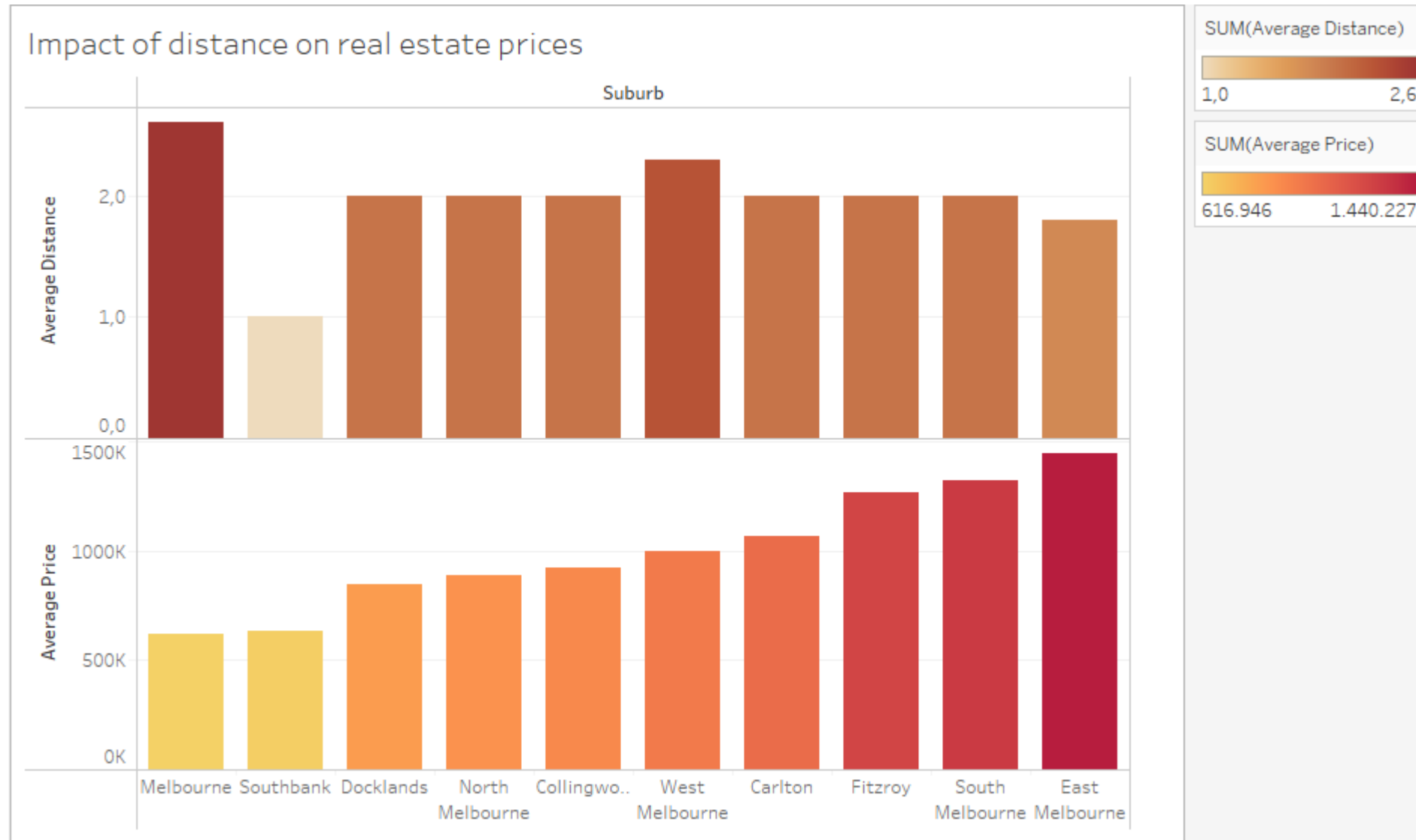
# Impact of distance on real estate prices

- Query:

```
Select TOP 10 Suburb, ROUND(AVG(Distance),1) average_distance, ROUND(AVG(Price),0)average_price
from melb_data
GROUP BY Suburb
ORDER BY average_distance ASC;
```

| | Suburb | average_distance | average_price |
|---|---|---|---|
| 1 | Southbank | 1 | 634012 |
| 2 | East Melbourne | 1,8 | 1440227 |
| 3 | Fitzroy | 2 | 1263186 |
| 4 | Collingwood | 2 | 924192 |
| 5 | North Melbourne | 2 | 884638 |
| 6 | Docklands | 2 | 843333 |
| 7 | Carlton | 2 | 1066948 |
| 8 | South Melbourne | 2 | 1317917 |
| 9 | West Melbourne | 2,3 | 994516 |
| 10 | Melbourne | 2,6 | 616945 |

The average price of a property in a suburb is not affected by its distance from the CBD. Prices in more distant Melbourne suburbs are cheaper than in East Melbourne, which is closer to the CBD.

# Impact of distance on real estate prices

# Relationship between real estate land size and price

- Query:

```
select TOP 10 Suburb, ROUND(AVG(Price),0) average_price, ROUND(AVG(Landsize),2) average_landsize
from melb_data
group by Suburb
order by average_landsize DESC;
```

| | Suburb | average_price | average_landsize |
|---|---|---|---|
| 1 | Silvan | 1077500 | 57245 |
| 2 | New Gisborne | 1355000 | 44500 |
| 3 | Gisborne | 648625 | 8437,88 |
| 4 | Fitzroy | 1263186 | 7965,51 |
| 5 | Bullengarook | 535000 | 4000 |
| 6 | North Warrandyte | 1226000 | 2839 |
| 7 | Warrandyte | 1219200 | 2142,8 |
| 8 | Travancore | 1010450 | 1854,6 |
| 9 | Mount Evelyn | 567500 | 1650 |
| 10 | Monbulk | 720000 | 1611 |

The price of a property in the suburbs is not affected by the land size of the property. Prices in Monbulk suburb are more expensive than in Gisborne.

# Relationship between real estate land size and price
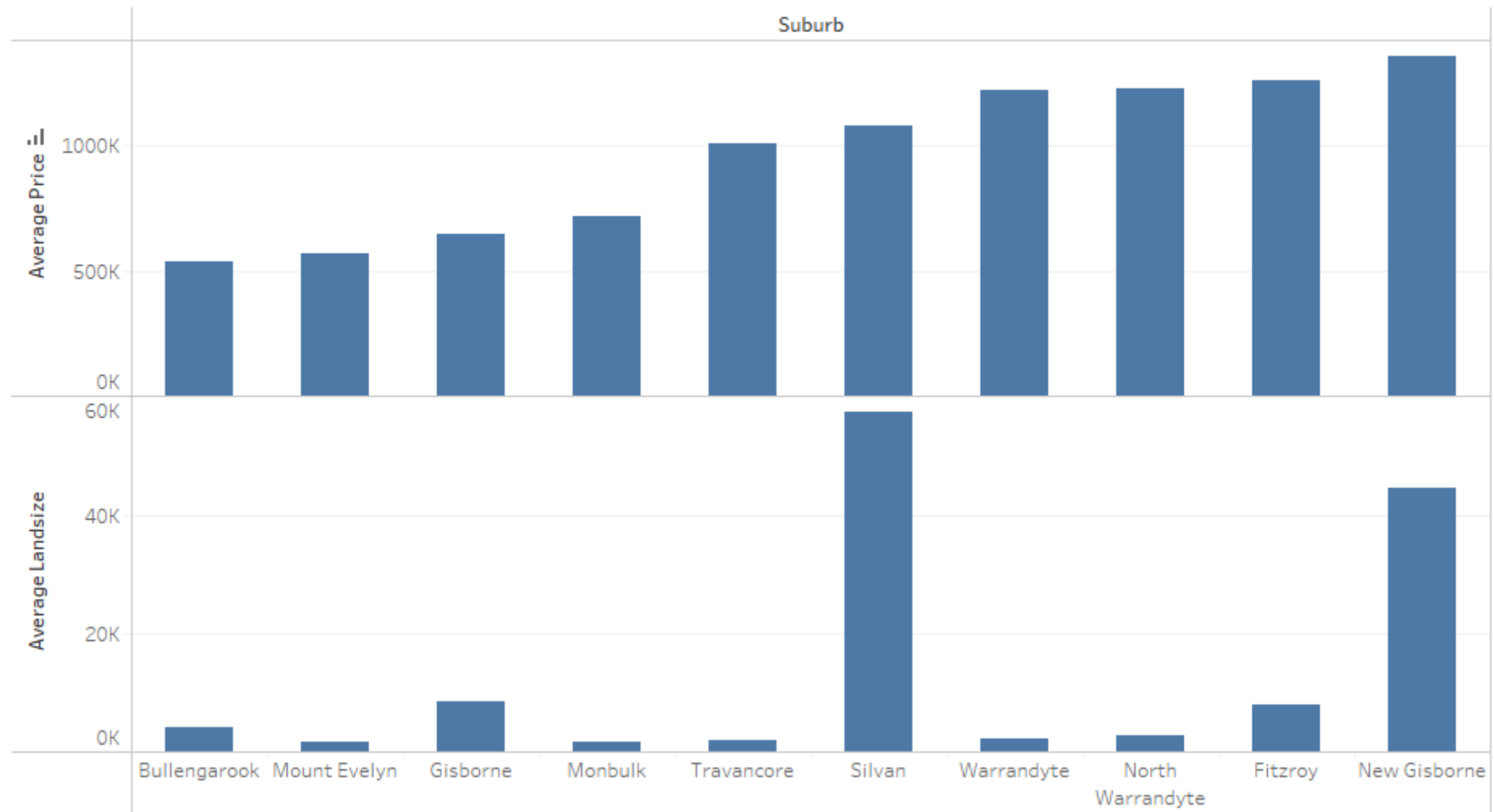
# Annual price trends by real estate type

- Query:

```sql
select Type, YEAR(Date) AS year_properties, ROUND(AVG(Price),0) average_price
from melb_data
GROUP BY Type, YEAR(Date)
ORDER BY Type, YEAR(Date) ASC;
```

| | Type | year_properties | average_price |
|---|---|---|---|
| 1 | h | 2016 | 1285995 |
| 2 | h | 2017 | 1208451 |
| 3 | t | 2016 | 900155 |
| 4 | t | 2017 | 964874 |
| 5 | u | 2016 | 585187 |
| 6 | u | 2017 | 628592 |

Prices for property types h (house, cottage, villa, semi, terrace) decreased from 2016 to 2017, while prices for property types t (townhouse) and u (unit, duplex) increased.

# Annual price trends by real estate type
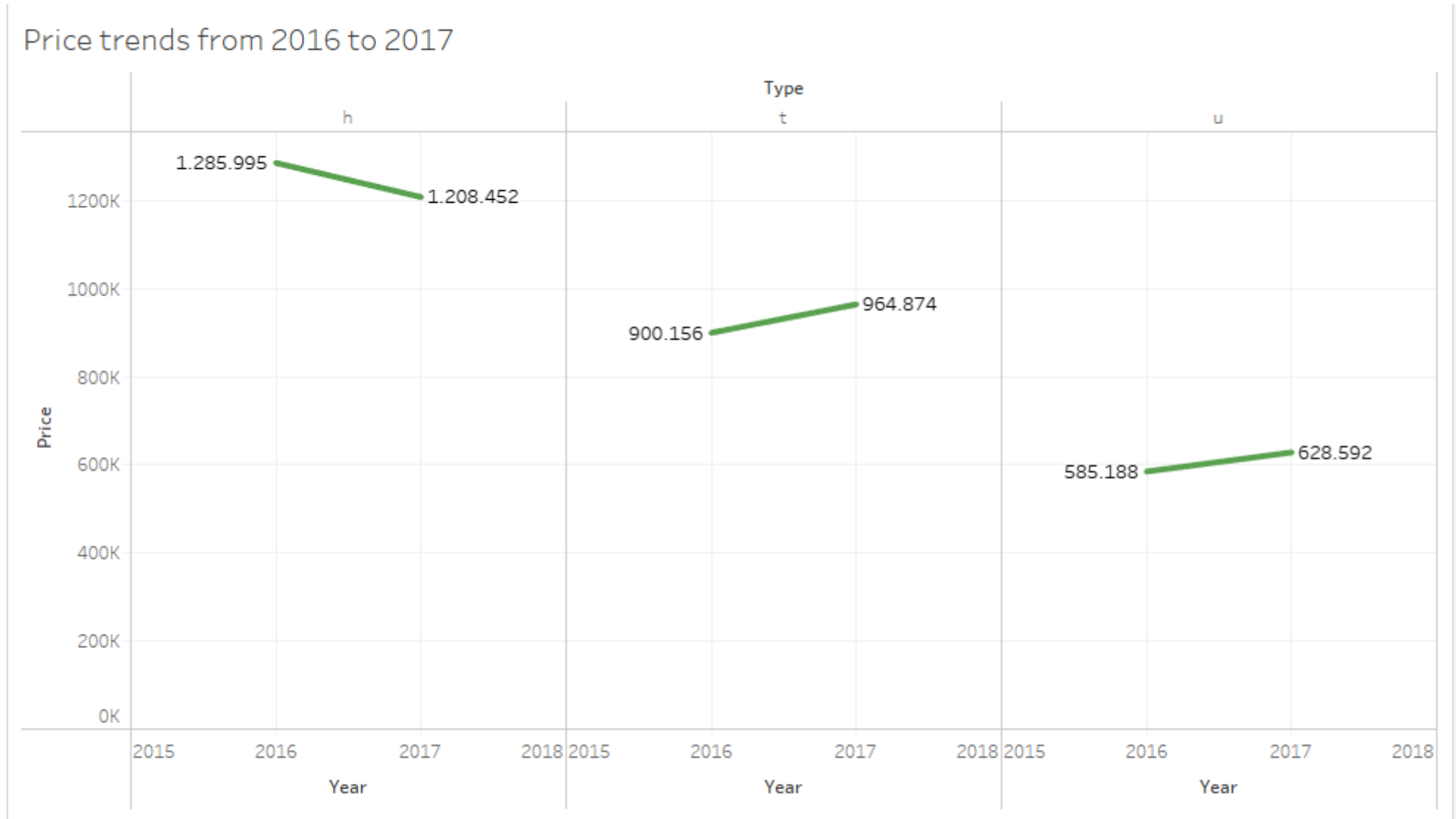
# Real estate type with many amenities, affordable prices and close to CBD

- Query:

```sql
Select Type, Suburb, Address, COUNT(*) total_properties from melb_data
where Bedroom2 > 1
AND Bathroom > 1
AND Car > 1
AND Landsize > 1000
AND Distance <= 4
AND Price < 700000
Group By Type, Suburb, Address
ORDER BY total_properties;
```

|   | Type | Suburb | Address | total_properties |
|---|------|--------|---------|------------------|
| 1 | u | Collingwood | 226/3 Hoddle St | 1 |
| 2 | u | Maribyrnong | 2/6 Horizon Dr | 1 |
| 3 | u | Richmond | 607/2 McGoun St | 1 |

Properties with many amenities such as more than 1 bedroom, more than 1 bathroom, land area of more than 1000, distance to CBD less than 4 and with affordable prices are found in Collingwood, Maribyrnong and Richmond suburbs with type U (unit, duplex).

# Real estate types with larger than average acreage at affordable prices.

- Query:

```sql
SELECT Type, suburb, address, landsize
FROM melb_data
WHERE landsize > (SELECT ROUND(AVG(landsize),1) FROM melb_data)
AND Price <= 270000
ORDER BY landsize DESC;
```

| | Type | suburb | address | landsize |
|---|---|---|---|---|
| 1 | u | Malvern East | 7/696 Waverley Rd | 2410 |
| 2 | u | Albion | 10/117 Anderson Rd | 2347 |
| 3 | u | Glenroy | 14/854 Pascoe Vale Rd | 1996 |
| 4 | u | Brunswick | 4/831 Park St | 1250 |
| 5 | u | Albion | 1/121 Anderson Rd | 1175 |
| 6 | u | Glen Huntly | 8/131 Grange Rd | 1162 |
| 7 | u | West Footscray | 4/38 Hampton Pde | 1059 |
| 8 | u | Footscray | 9/3 Gordon St | 1015 |
| 9 | u | Pascoe Vale | 4/566 Pascoe Vale Rd | 995 |
| 10 | u | Preston | 5/49 Austral Av | 631 |

Real estate with larger than average land size and affordable prices can be found in the suburb of Malvern East at 7/696 Waverley Rd.

# Real estate types with larger than average acreage at affordable prices.

Real estate types with larger than average land size

# Conclusions

- The real estate distance factor to the CBD is not affected by price.

- The land size factor of a real estate is not affected by price.

- The determining factor of high and low real estate prices is the location, whether it's a suburb or a region.

- Despite the price increase, the lowest priced property is the U type (unit, duplex).