

FIT5147 Data Visualization And Exploration

The factors which affect the Melbourne's Housing Price

Student Name:Jiacheng Liu

Student Number:31122434

Lab Number:05

Tutor's Name:Aldrich

Clarence

Introduction

1.1 Background

In today's highly developed society, housing prices have become an important factor affecting people's lives. With limited funds, how to choose a good house as an asset based on various factors is something everyone should understand. The topic of this article is the factors that may affect housing prices, in five selected areas. Focusing on the analysis of location (distance from the city center), room type, and school's impact on housing prices, this article uses D3 for data visualization and adds a relatively large amount of human-computer interaction.

We will discuss the location, type of house, and school. First of all, the first point is location. There is a saying in the real estate industry: "location, location or location". This shows the importance of location to the house. The city center happens to be the most prosperous location in a city. Such a place is generally formed naturally in the long-term development of the city. It is rich in resources, popular, strong in business, complete with supporting facilities, and has formed a complete "ecological chain" with functions such as clothing, food, housing, transportation, and medical education. Work and life are more convenient, and housing prices are natural. It is relatively strong.

The advantage of living in the suburbs is that the environment is quiet and the air is fresh: the suburbs have fresh air, high green rates, good views, and relatively quiet. Ventilation and sunshine are better than the urban area: the suburban houses are mostly low-density residential buildings with large spacing between buildings, and the ventilation and sunshine are better than the urban.

The advantage of living in an apartment is that the apartment is generally equipped with a swimming pool, gym, access control system, security, etc. Australia requires the installation of elevators in apartment buildings above the 3rd floor. Apartment buildings are generally taller, so there are few bugs, good air circulation and smooth ventilation.

The last point is the school. Generally speaking, the price of a house in the best school district will be higher than similar houses in a less popular school district.

At the same time, in general, many wealthy owners live in school districts with more resources. Statistics also show that the richer a community, the higher its test score. Test scores are just one way to measure "good schools", but they are also a highly cited method. There is a self-reinforcing mechanism that may overemphasize the influence of the school itself in the price of these houses.

1.2 Audience

There are two hypothetical audiences in this article. The first is real estate developers who are more inclined to do research to better find land with more investment value. The second is potential buyers who tend to understand the real estate market, So that they can choose the type of house that is more suitable for them with limited funds.

Design

2.1 Design Process

First of all, I will answer according to the original proposed questions. The first point is to brainstorm and select questions with suitable difficulty and new ideas. In this article, I will continue to discuss the first three questions, but what about the first three questions? The answer method has been changed from static pictures to optional pictures for human-computer interaction, which better solves the readability of web pages.

Starting from the introduction of this article, I first compared Melbourne with Sydney and Brisbane, and explained in detail the advantages of Melbourne as a first-time home purchase location. Here is a comparison with a line chart. One of our hypothetical buyers, a home buyer, can clearly see that Melbourne has more investment value than Sydney and Brisbane.

In the first question, we will discuss whether the distance from the city center will affect housing prices, and to what extent. We first need to load the map and display five different areas on the map. In order to be more representative, this article chose the southeast area with a larger population and better facilities. Buyers can also clearly see the distance between the suburbs and the city center. In the initial five design sheets, I prefer to use a line chart to explain it, but it is obvious that the line chart is not intuitive to the map.

As shown in the figure below, the line chart is not suitable for distance comparison. The line chart should be more inclined to indicate increase or decrease, but it is used to obtain information intuitively. It is not as direct as the map.

The distance from CBD and suburb

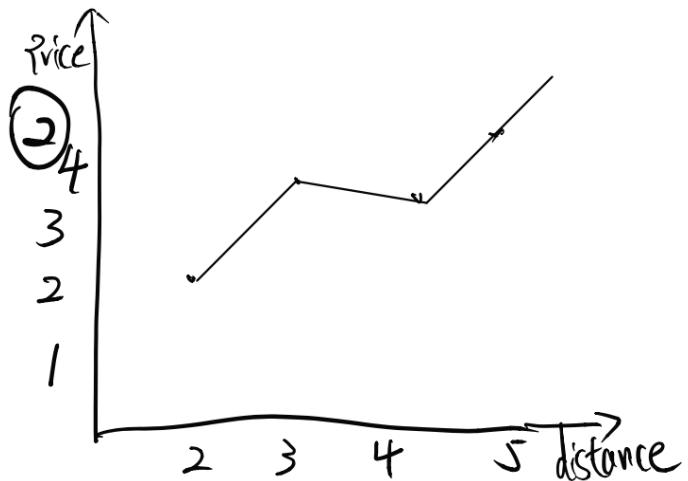


Figure 1. The Line chart in five design sheet

In contrast to the map, home buyers or real estate companies can see the location of the house more clearly, and can also feel the distance more intuitively. In the map, I added a line segment and a label, and the user moved the mouse to the line segment. The user can see the distance calculated according to the latitude and longitude, and the user can better compare with other regions.

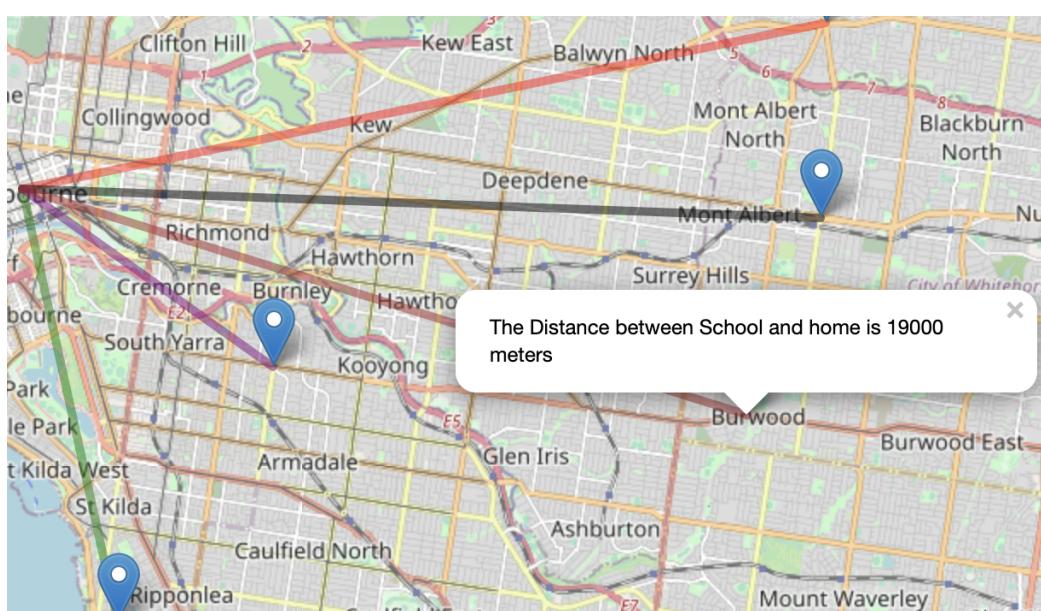


Figure 2. The map

As for the second question, in the influence of house types on housing prices, in the five design sheets, I tend to use pie charts, but in the process of drawing, we can clearly see that pie charts are not intuitive enough, and they are fan-shaped.

The graph does not make a reasonable explanation between the house type and the price, so this article uses a dynamic scatter plot and has three options.

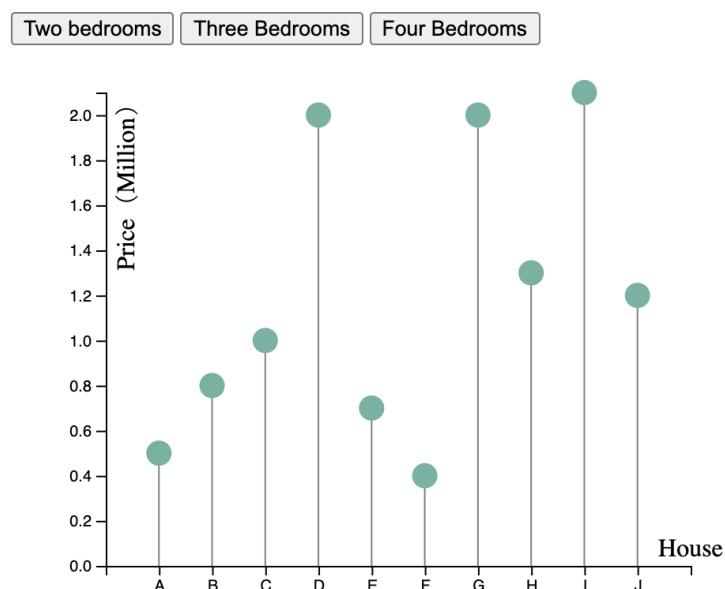


Figure 3. Plot graph

We can make a rough analysis of randomly selected houses based on these three options. The design concept of this map is to randomly select 10 houses and divide them according to the suburbs and house types, so as to get the relationship between house prices and house types.

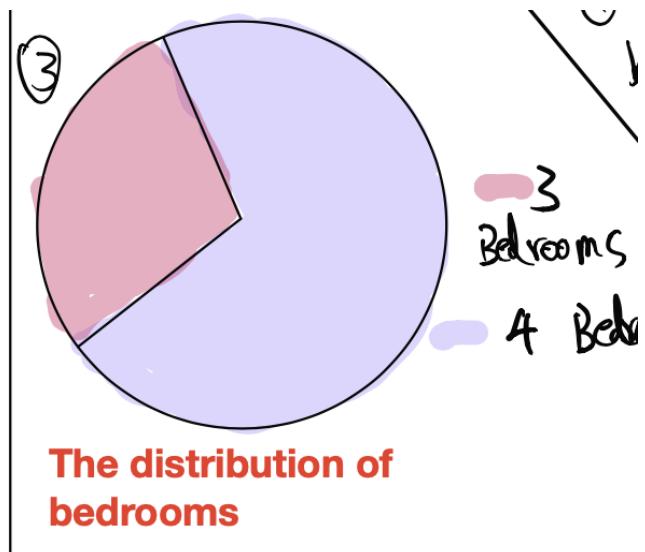


Figure 4 Five design sheet

In the last question, in the initial five design sheets, what I want to discuss next is the relationship between the number of toilets and the housing price, but it is clear that the relationship between the type of housing and the housing price we mentioned in the previous question The relationship already includes this problem. Generally speaking, two-bedrooms have one toilet, and three-bedrooms and four-bedrooms have two or more toilets. So we still put the direction of discussion on the school.

This article does not discuss the importance of parking lot to housing prices in accordance with the five design sheets. This is because apart from apartments that require parking lots, other houses do not have strong demand for parking lots and are not universal.

In the five design sheets, I just choose the most basic diagrams, and I can't communicate the information better, so I choose the more advanced diagrams, they can better communicate the information.

VIC Map

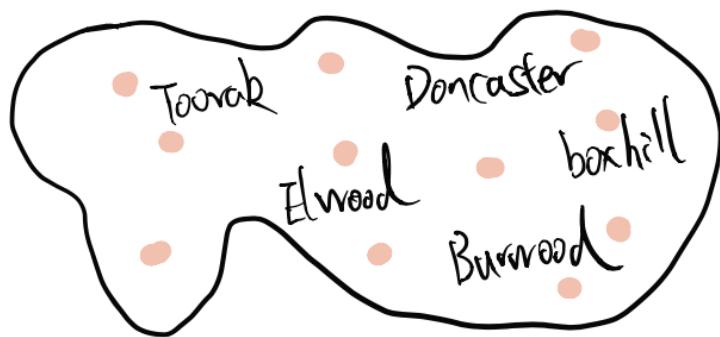


Figure 5 Map about VIC school

Here we will continue to use the map icon to show the distance between the school and the suburban center, and estimate the price based on the number of schools in their area, so that the relationship between the house price and the school is more universal through distance and quantity

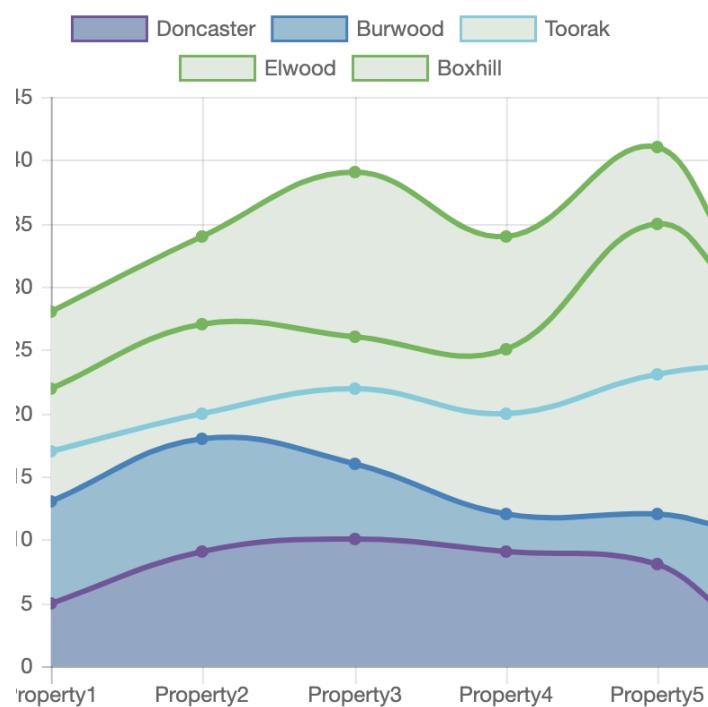


Figure 6 The number of school in each suburb

Obviously, in actual operation, the design of five design sheets is unreasonable. In the process of drawing, I gradually realized that there would be better options to clearly depict the quantitative relationship, so the final interaction selected by the interaction is much better than the method in the five design sheets.

3.Implementation

3.1 High-Level description

The research methods in this article all use D3.js, the first is to select the appropriate data, the data mainly comes from Kaggle and Data In Australia and Open data in Melbourne. Secondly, use ThebeautifulSoup library to crawl data from web pages.

```
from mpl_toolkits.mplot3d import Axes3D
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt # plotting
import numpy as np # linear algebra
import os # accessing directory structure
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
```

Figure 7 All library to use in data wrangling

For the above six libraries, pandas and numpy are the main ones used for data sorting. In D3, I mainly used the following css files for reference, and then in the process of drawing the map, I mainly used Leaflet, but there are also very few parts using the Google Map API.

```
<!-- Google fonts -->
<link href="https://fonts.googleapis.com/css?family=Open+Sans:400,300" rel='stylesheet' type='text/css'>
<link href='https://fonts.googleapis.com/css?family=Raleway' rel='stylesheet' type='text/css'>

<!-- D3.js -->
<script src="https://cdnjs.cloudflare.com/ajax/libs/d3/3.5.6/d3.min.js" charset="utf-8"></script>
```

Figure 8 Link in D3

In the process of drawing various dynamic pictures, the following css files are mainly used, because these files can better represent pictures and can use more accurate information to transmit.

```
<!DOCTYPE html>
<html lang="en" >
<head>
<meta charset="UTF-8">
<title>The Factor which affect Melbourne Housing Price</title>
<link rel='stylesheet' href='https://ajax.googleapis.com/ajax/libs/jqueryui/1.11.2/themes/smoothness/jquery-ui.css'>
<link rel='stylesheet' href='https://maxcdn.bootstrapcdn.com/bootstrap/3.2.0/css/bootstrap.min.css'>
<link rel='stylesheet' href='https://startbootstrap.com/templates/agency/font-awesome-4.1.0/css/font-awesome.min.css'>
```

Figure 9 CSS in D3

4.User Guide

4.1 Instruction

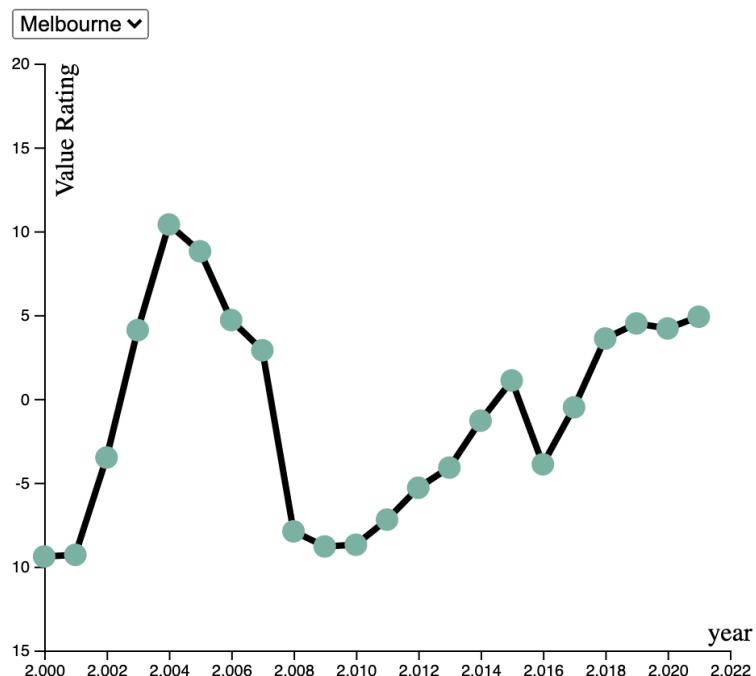


Figure 10 Housing Price Compare

The main purpose of this chart is to compare the housing price growth trends of Sydney, Brisbane, and Melbourne. Starting from 2000, housing prices are about the same. We can clearly see the forecast from 2000 to 2022. The data for 2022 is mainly It uses Bayesian algorithm to predict.

This article collects data when housing prices increase, such as elections, policies, and nearby high-tech zones, as well as

population proportions. For example, in 2000, Monash University's impact on Clayton's housing prices was not too high, and Clayton's housing prices were also at an average value. However, in 2010, Monash University's large enrollment of international students caused Clayton housing prices to rise sharply. In 2022, the gradual disappearance of Covid-19 and the further liberalization of transit in Australia will lead to a kind of retaliatory consumption and a surge in housing prices.

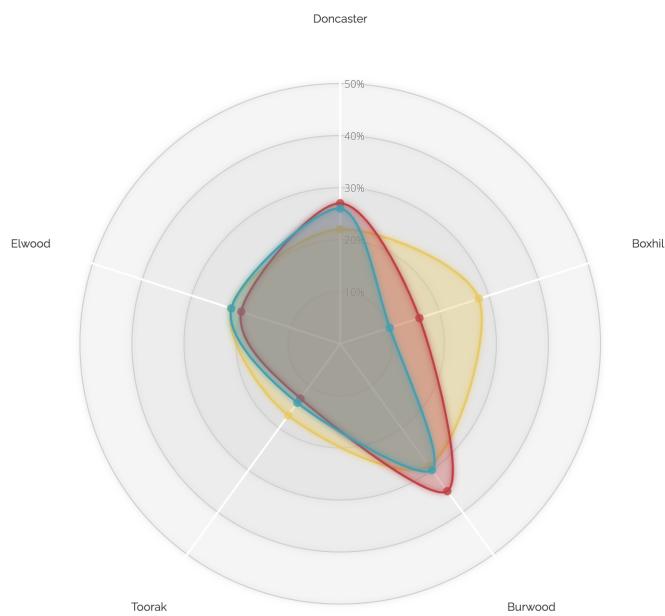


Figure 11 Conclusion

This graph shows the percentage of the distance from the city center to the sales price of five houses

Here is a distinction based on the room type of the selected house. When a buyer wants to check the price of a two-bedroom house, he can click the button and check the price distribution. We have assigned two houses to each suburb, for example, A And B represents the location of the house in Doncaster.

Two bedrooms Three Bedrooms Four Bedrooms

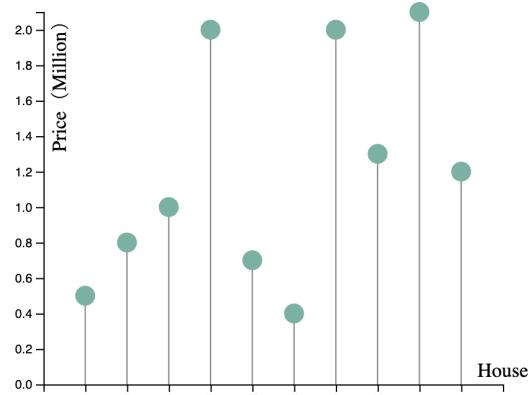


Figure 12 House Typing

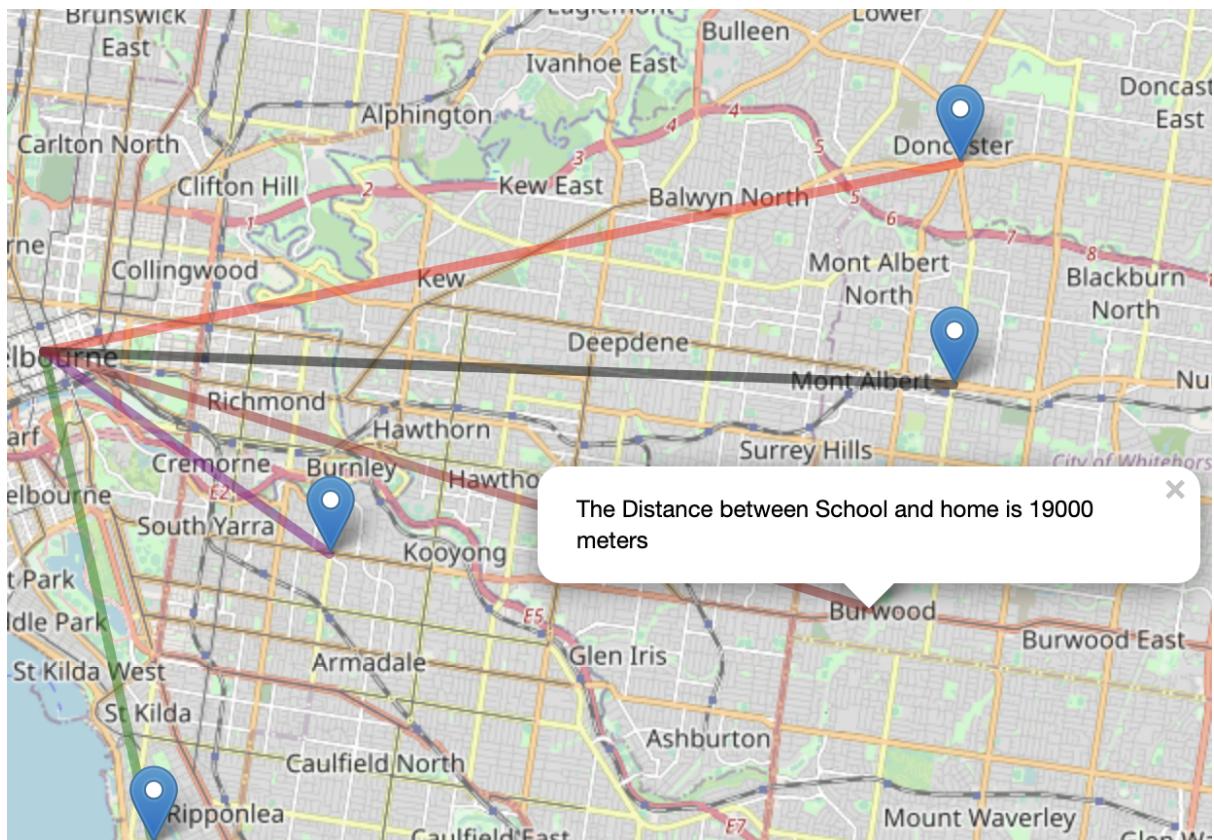


Figure 13 Distance From CBD

The purpose of this picture is to let buyers understand the distance between each suburb and the city center more intuitively. When the mouse is moved to the line representing the distance, the distance can be checked.

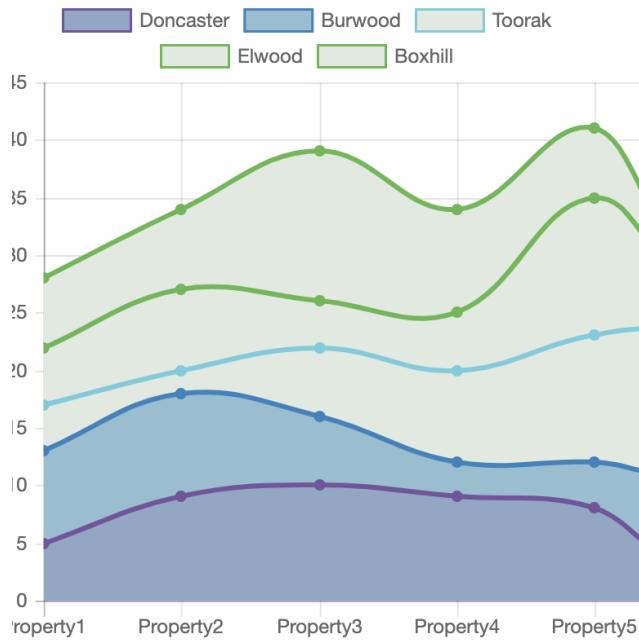


Figure 14 The number of school every suburb has

The purpose of this picture is to point out the 5 houses selected in each area and the school resources owned by each. When the mouse is placed on it, the specific values can be better seen

5. Conclusion

5.1 Findings

We can learn from this project that the location, school and house type are all but not a single factor affecting the housing price. Only when they are combined can they affect the housing price. For example, living in a two-bedroom apartment and located in For people in the city center, there are not too many schools nearby, but the housing prices remain high. On the contrary, for people living in box hill four-bedroom houses, there are also not many schools around them, but the housing prices are also very high. This is because People who are willing to pay high prices don't care about commuting costs, but a set is located in Elwood, and is close to the city center. Nearby educational resources are also good, but the house is sold at a very low price, because the crime rate nearby is too high. From this we can conclude that if we want to consider the influencing factors of housing prices, we still have to start from many aspects.

In the writing process of this article, the most worthy part is to sort out a large amount of data, and clear out most of the useless, invalid, blank, and abnormal values through the Filter and Select methods. And compared to the use of D3, the interface is first of all exquisite, users can select the part they want to watch through the menu, and a large number of dynamic graphs and interaction graphs are used to convey information. And every dynamic graph is the most appropriate one of the methods that can convey information.

6.Bibliography

1. (2013). D3.

https://www.d3-graph-gallery.com/graph/lollipop_button_data_csv.html

2. Radar. (2015). D3 Gallery.

<http://bl.ocks.org/nbremer/21746a9668ffdf6d8242>

3.D3 Gallery. (2014). 3. D3.

https://www.d3-graph-gallery.com/circular_barplot.html

4. (2015). D3. <https://www.d3-graph-gallery.com/connectedscatter.html>

5.*Leaflet — an open-source JavaScript library for interactive maps.*

(2018). Leaflet. <https://leafletjs.com/>

6.D. (2021, June 6). *Density contours with linear binning*. D3.

<https://observablehq.com/@d3/contourdensity-linear-binning>

7.*LeafletJS - (Markers) _LeafletJS|WIKI*. (2012). Door.

https://iowiki.com/leafletjs/leafletjs_markers.html

8.*LeafletJS - (Vector Layers) _LeafletJS|WIKI*. (2014). Door.

https://iowiki.com/leafletjs/leafletjs_vector_layers.html

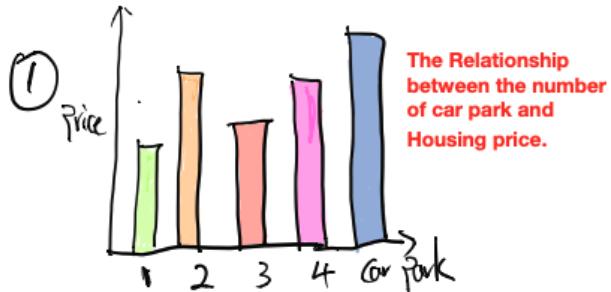
9.W. (2018). *woxingxiao/BubbleSeekBar*. GitHub.

<https://github.com/woxingxiao/BubbleSeekBar>

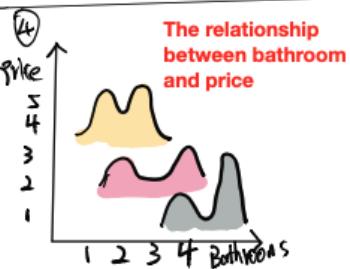
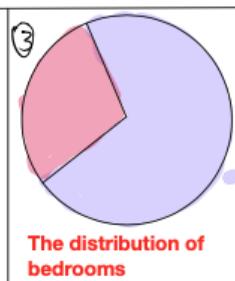
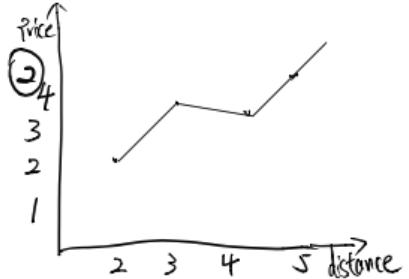
10.*Attention Required!* | Cloudflare. (2020). Celine.

<https://codepen.io/celincky/pen/zdqso>

Generate Ideas



The distance from CBD and suburb



Filter the idea
keep all of them,

Categorize the sketches

① and ② show the distribution of car park and distance.

③ show the relationship of the type of room and number of bedrooms

④ shows the relationship between bathroom and price

Combine and Refine

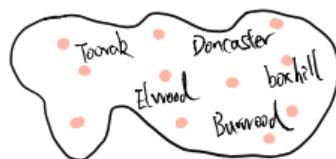
question and solution

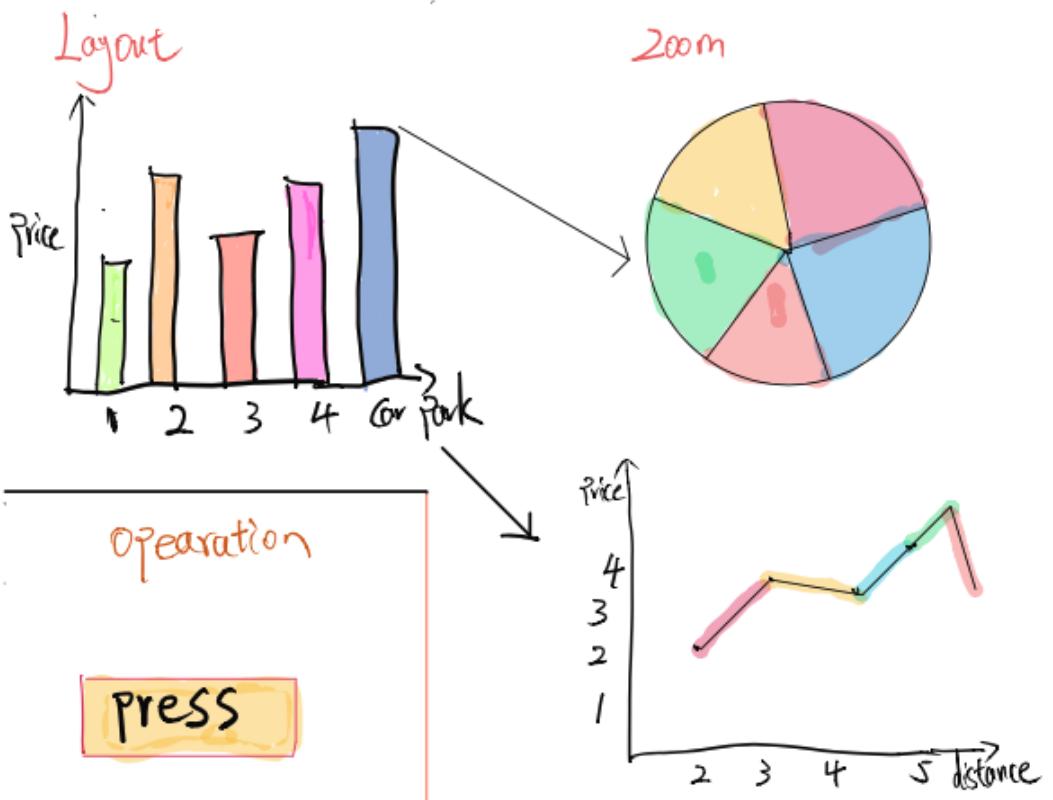
Does the idea solve
the problems?

Yes

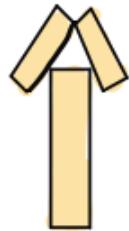
① and ② to answer the question 1, to explain the relationship between car park and distance. ③ to answer the question 2, which type of house is the most popular. ④ to answer to question about the bathroom's price to affect the price.

VIC Map





when you click it,

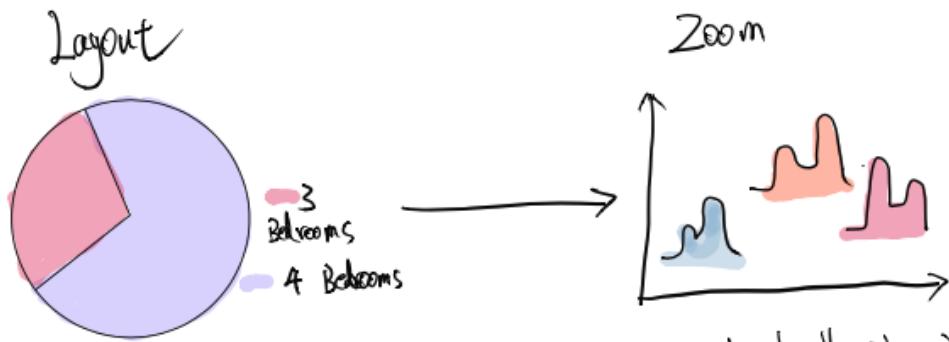


Discussion :

advantage: Using this map can enhance the user experience and make it easier for users to find data quickly

disadvantage:

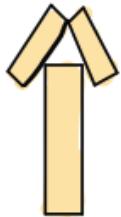
The image is too simple and cannot contain too many elements, which will result in the need for users to switch modes continuously



when you choose to zoom, it will show a average value of all price in melbourne,
It will provide a reference to check the file.

Operation

Display



when you click it, you can
check the graph transfering.

Discussion :

advantage: This image provides an approximate average value for reference, which will facilitate consumers to determine whether the conclusion is universal for housing prices in different regions

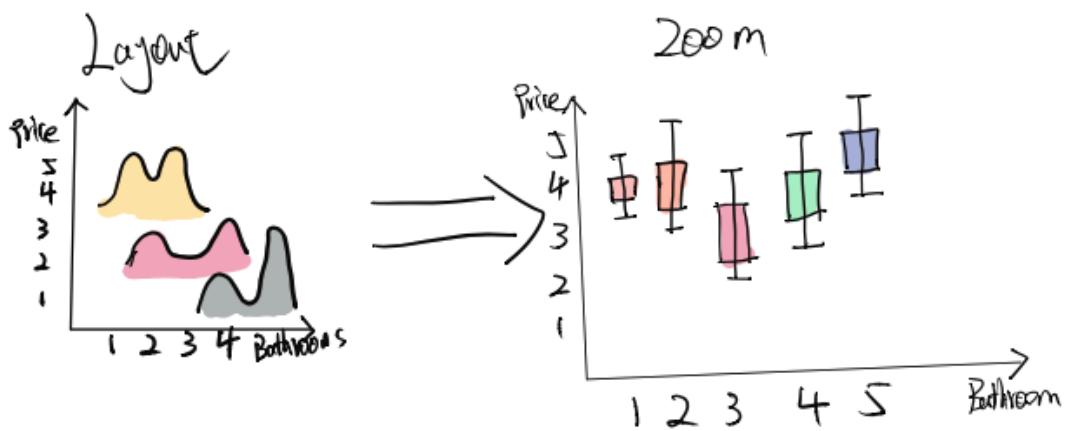
disadvantage: It's good to provide an average value, but we have to consider that the reference items sometimes mislead users to analyze, which will seriously affect the user's experience.

Meta-info :

Title: the relationship between the house type
and price

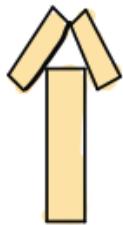
sheet: 3

Author: JIACHENG LIN



Operation

Display



Discussion :

advantage.

Using this graph, we can try to find outliers and filter the outliers, because we only consider whether the number of toilets has an impact on the house price

disadvantage

The disadvantage is that the toilet is a kind of the number of rooms, and the above problems may cause the data to overlap, and the conclusions drawn are not accurate enough.

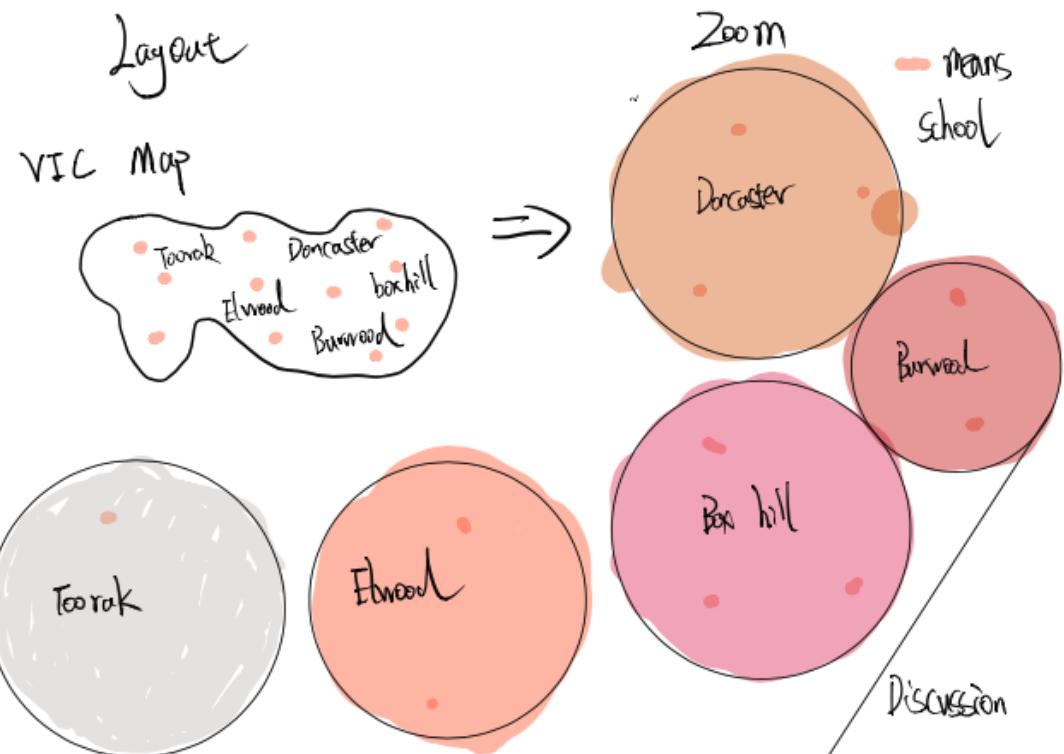
Meta-info :

Title: How the number of bathrooms affect Price

and price

sheet: 4

Author: JIACHENG LIN



Discussion :

advantage

The advantage of this map is that the entire map is divided into five, so that it is easy to find the number of schools in each suburb

disadvantage

The disadvantage is that it is difficult to find schools jointly owned by neighboring suburbs, which leads to duplication of calculations.

Meta-Info :

Author : JIACHENG LIU
date: 16-May-2021

sheet : 5

In this project, I will mainly use D3 to build a dynamic visualization map, such as Q1. We can highlight the values to be compared and display their corresponding values. In the last question, we can display each area. The path from the center store to the school is highlighted, the house price is marked on the label, and then the comparison is made.

