

Housing Sales Prices Vs Venues Data Analysis of Sydney

Shanshan Jin | Coursera Capstone Project | 18-07-2020

1 Introduction

1.1 BACKGROUND

Sydney is the most populous city in Australia and Oceania. Most times, when people talking about Sydney, they actually mean Greater Sydney (Greater Capital City Statistical Area). So does this report. As of June 2019, Greater Sydney's estimated metropolitan population was 5,312,163. Based on a total site area of 12,368.2sq km, the current population density of Greater Sydney area is 430 persons per square kilometer. The built urban area is estimated at 4,196sq km which translates to a density of 1,171 persons per square kilometer. Sydney is made up of 658 suburbs, 40 local government areas and 15 contiguous regions.[1] Being a resident of this city, I decided to use Sydney in my project.

1.2 PROBLEM

Although ranking top 10 in the Global Liveability Ranking list [2], the housing price in Sydney is relatively high and growing fast. When people making decision of purchasing a house, they usually want to buy a house with good growing potential through a lower price. At the same time, they may also want to choose the suburb according to the social places' density. Meanwhile, according to the 2016 census, 18.6% of the population were children aged between 0 to 14 years, those with children in the family may also consider education resource in the suburb.

1.3 INTEREST

When we consider all these problems, we can create a map and information chart where the real estate index is placed on Sydney and each suburb is clustered according to the housing price growing rate, venue density and number of schools with OC.

2 Data acquisition and cleaning

2.1 DATA SOURCES

- I got the Sydney property prices from 2000 to 2019 from Kaggle dataset. [3]
- I got the suburb data of Great Sydney from Matthew Proctor's Database of Australian Postcodes [4].
- I got the list of Opportunity Classes from education.nsw.gov.au [5].
- I used Forsquare API to get the most common venues of given Borough of Sydney [6].

2.2 DATA CLEANING

2.3 FEATURE SELECTION

[1] City of Sydney(www.cityofsydney.nsw.gov.au)

[2] Global Liveability Ranking-Wikipedia

[3] Sydney property prices from 2000 to 2019(<https://www.kaggle.com/mihirhalai/sydney-house-prices/data?select=SydneyHousePrices.csv>) cleaned by Mihir Halai.

[4] Database of Australian

Postcodes(https://www.matthewproctor.com/australian_postcodes) by Matthew Proctor.

[5]Opportunity Class list (<https://education.nsw.gov.au/public-schools/selective-high-schools-and-opportunity-classes/year-5/what-are-opportunity-classes/list-of-opportunity-classes>)

[6] [Forsquare API](#)