How to select neighbourhoods for your first rental property in Sydney

Introduction

Description & Discussion of the Background

Sydney, capital of New South Wales and one of Australia's largest cities covers 12,367.7 sq km land and is made up of 35 local councils. As of June 2019, Sydney has a population of 5.73 million and spread into 658 suburbs. The metropolitan area of Sydney sprawl about 70 km to the west, 40 km to the north, and 60 to the south. Every year we have new migrater from countries all over the world. Over the last 8 years, Sydney have a very strong population growth rate which reach's it peaks in 2013 of 10.88% with a total 1.7 million growth of population since year 2011 (4.029 Million). [1] The strong population growth causes the high demand of residential property needs. According to afr.com financial review, rental vacancies in Sydney dropped to 2.9 percent which is lowest we have seen in the last 2 years. [2]

I have been living in Sydney almost 13 years now, with all the rental and investment experience I have gain over the years. I'm now consider myself quite familiar of the most popular neighbourhood in Sydney and know which suburb is might be a good choice to live at different life stages. Every now and then, I have some new friends decide to move to Sydney from China. Most of the time, they come to consult me on where should they rent when they first move to this country. Well, this might sound like a simple question, however each new comer actually always has different needs and preference on their dream property and the recommendation needs to vary depend on what feature they care the most.

With all the "consulting experience" I had; I can almost summarise these features in below category.

1. **Price**. (This directly link to the future tenant's budget depend on their financial situation)
2. **Location**. (If they come to Australia to study at Sydney Uni, of course they would prefer rent some place close or convenience to their daily routine)
3. **Quality or feature of the property** (Such as: is this property has a dish washer, security intercom, parking spots etc)
4. **Community** (This will normally link to tenants' life style. They will be asking question such as is this area secured? Any amenities available around the area is what I usually go? (Restaurants, shops, grocery stores etc. If they have kids, they normally will be looking for neighbourhood with good school as well)

For this paper, we will be mainly analysis 3 factors: **Rental Price**, **neighbourhood security and available amenities in the area as well** to see how we can use data science to help our new comer to select their dream property when they first move to Sydney. I will also use one of the real-life examples to take you through whole process + share you all the data and python code I have complete along the way.

Data Description:

For this analysis, I have collected quite a few data via the internet which most of them are publicly available for download. Detail of these dataset and samples are listed below:

Rental Price Data:

NSW Communities & Justice provide Rent and Sales Report which is consider the sole authoritative source of data on NSW rent movements. This report is published quarterly and has been published on a regular basis since 1987 for rental movement in NSW.

For this analysis, I have downloaded the most recent rent table as of JUN-2019. It is an xlsx file and link to the table is available below. This file contains weekly rent and bond by Quartile and NSW postcode and property type. It also gives you the last Quarter or year rent movement as well. This will be the data we use to identify the rental price for each suburb in Sydney. Please notice the data is on postcode level, not neighbourhood level, hence we will use a proxy in our analysis to assume the rental price for each neighbourhood in one suburb will be around the same. [3]

https://www.facs.nsw.gov.au/resources/statistics/rent-and-sales/dashboard

	Postcode	Dwelling_Types	Bedroom_Numbers	First_Quartile_Weekly_Rent	Median_Quartile_Weekly_Rent	Third_Quartile_Weekly_Rent	New_Bonds_Lodged	Total_Bonds_Held	Annual_change_in_Median_Weekly_Rent
20	2000	Total	Total	610	720	900	1253	7946	0.036
21	2000	Total	Bedsitter	450	493	580	38	223	-0.0248
22	2000	Total	1 Bedroom	575	650	700	687	3522	0.0833
23	2000	Total	Not Specified	540	645	820	67	750	-0.0373
24	2000	Total	2 Bedrooms	830	915	1050	408	2985	-0.0368
25	2000	Total	3 Bedrooms	1195	1390	1700	48	420	0.0692
27	2000	House	Total	813	1000	1225	s	114	0.0526
34	2000	Other	Total	543	605	665	243	1430	0.2124
36	2000	Other	1 Bedroom	543	605	665	225	889	0.2124
47	2000	Flat/Unit	Total	650	780	920	989	6327	-0.025

NSW Neighbourhoods Geo Coordinate Data:

Corra website has also supply the full list of Australia neighbourhoods with its longitude and latitude coordinates. Link provided below. Please notice, in NSW, we can have multiple neighbourhoods linked to 1 postcode. For this analysis, we will go down to neighbourhoods' level given some suburbs

have a larger size compare to others and neighbourhoods will generally give you a more even split of interest area of Sydney. This will be the Geo data we will feed into Foursquare API to get venue data for each neighbourhood in Sydney. [4]

http://www.corra.com.au/australian-postcode-location-data/

	postcode	suburb	state	dc	type	lat	lon
0	200	AUSTRALIAN NATIONAL UNIVERSITY	ACT	AUSTRALIAN NATIONAL UNI LPO	Post Office Boxes	-35.277272	149.117136
1	221	BARTON	ACT	NaN	LVR	-35.201372	149.095065
2	800	DARWIN	NT	DARWIN DELIVERY CENTRE	Delivery Area	-12.801028	130.955789
3	801	DARWIN	NT	DARWIN DELIVERY CENTRE	Post Office Boxes	-12.801028	130.955789
4	804	PARAP	NT	PARAP	Post Office Boxes	-12.432181	130.843310

Sydney Postcode and Suburb Mapping:

Prospectshop.com website powered by Equifax provided a list of postcodes used by Sydney Metro area only. For our analysis, we will be mainly focus on Sydney not remote area in NSW, as the rental price for rural area can be significantly vary compare to metropolitan area. This list of postcodes will be mainly used for us to filter the data to narrow to Sydney only. [5]

https://www.prospectshop.com.au/resources.aspx

	Suburb	Postcode	Area
0	Sydney City	2000	Central & Inner Metropolitan
1	Ultimo	2007	Central & Inner Metropolitan
2	Chippendale	2008	Central & Inner Metropolitan
3	Pyrmont	2009	Central & Inner Metropolitan
4	Surry Hills	2010	Central & Inner Metropolitan
5	Kings Cross	2011	Central & Inner Metropolitan
6	Alexandria	2015	Central & Inner Metropolitan
7	Redfern	2016	Central & Inner Metropolitan
8	Waterloo	2017	Central & Inner Metropolitan
9	Rosebery	2018	Central & Inner Metropolitan

Monthly Criminal Incidents for NSW:

NSW Bureau of Crime Statistics and Research has collected and provided significant crime, court and custody datasets which updated on annually or quarterly. These data are intended for use by people who want to perform their own analysis of the information. I have downloaded a copy of Monthly data on all criminal incidents recorded by police from 1995 to Sept 2019 from link provided below. It will you details on incident count by postcode by crime type by month for NSW. This data will be mainly used to identify the neighbourhood security in Sydney. Again, this data is on postcode level, not neighbourhood level, hence we will use a proxy in our analysis to assume the neighbourhood security in one suburb will be around the same. [6]

https://www.bocsar.nsw.gov.au/Pages/bocsar_datasets/Datasets-.aspx

	Postcode	Offence category	Subcategory	2017-01-01 00:00:00	2017-02-01 00:00:00	2017-03-01 00:00:00	2017-04-01 00:00:00	2017-05-01 00:00:00	2017-06-01 00:00:00	2017-07-01 00:00:00	...	2018-03-01 00:00:00	2018-04-01 00:00:00	2018-05-01 00:00:00	2018-06-01 00:00:00	2018-07-01 00:00:00	2018-08-01 00:00:00	2018-09-01 00:00:00	2018-10-01 00:00:00	2018-11-01 00:00:00	2018-12-01 00:00:00
0	2000	Homicide	Murder *	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
1	2000	Homicide	Attempted murder	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
2	2000	Homicide	Murder accessory, conspiracy	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
3	2000	Homicide	Manslaughter *	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
4	2000	Assault	Domestic violence related assault	20	18	18	15	28	16	25	...	18	16	13	18	19	27	19	13	18	28
5	2000	Assault	Non-domestic violence related assault	116	111	123	156	125	93	119	...	140	95	119	97	109	81	111	112	118	127
6	2000	Assault	Assault Police	7	6	6	5	13	1	4	...	11	8	7	1	8	6	8	4	8	11
7	2000	Sexual offences	Sexual assault	11	2	7	6	13	9	3	...	11	3	4	8	3	9	4	7	11	9
8	2000	Sexual offences	Indecent assault, act of indecency and other s...	13	9	19	19	16	11	46	...	14	10	8	25	28	9	15	7	16	22

Foursquare API Venues data:

Foursquare API is a social location service provided by Foursquare which allows users to explore the venue by given location details (Typically GEO location coordinates). For this analysis, we will be using the NSW neighbourhoods' geo data as a base to search all venues around the list of neighbourhoods given in the file. Then grouping them into different category to explore how they distribution around Sydney suburbs.

To use Foursquare API, we will require account to be registered with Foursquare before we can call below request (with registered credential) to retrieve the information we need. (Details how to do that will be shared in the Python code supplied)

https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}

: {'meta': {'code': 200, 'requestId': '5de9bb086001fe001c2ce01a'},
  'response': {'suggestedFilters': {'header': 'Tap to show:',
    'filters': [{'name': 'Open now', 'key': 'openNow'}]},
   'headerLocation': 'Central Business District',
   'headerFullLocation': 'Central Business District, Sydney',
   'headerLocationGranularity': 'neighborhood',
   'totalResults': 188,
   'suggestedBounds': {'ne': {'lat': -33.84660099099999,
     'lng': 151.21903734877398},
    'sw': {'lat': -33.86460100900001, 'lng': 151.19740265122599}},
   'groups': [{'type': 'Recommended Places',
    'name': 'recommended',
    'items': [{'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4d97e60e2bd6f04ddd795c50',
       'name': 'Harbour Bridge Pylon Lookout',
       'location': {'address': 'Bradfield Hwy.',

Foursquare API Venues Group:

This is a dataset compiled by myself manually based on the Venue type returned by Foursquare API. Currently the venue type provided by Foursquare API have hundreds of different types. I have provided this mapping file as part of feature engineering process to better group these venues to common type recognized by most of audience such as supermarket, shopping mall, restaurant etc.

	type	group
0	CafÃ©	cafe
1	Park	park
2	Pub	bar
3	Pizza Place	food_other
4	Thai Restaurant	restaurant
5	Japanese Restaurant	restaurant
6	Italian Restaurant	restaurant
7	Coffee Shop	food_other
8	Bakery	food_other
9	Bar	bar

References:

[1] "Greater Sydney"

https://www.cityofsydney.nsw.gov.au/learn/research-and-statistics/the-city-at-a-glance/greater-sydney

[2] "Sydney rental vacancy rate drop". Su-Lin Tan Reporter

https://www.afr.com/property/residential/sydney-rental-vacancy-rate-drops-20191011-p52ztk"

[3] "NSW Rent and Sales Report". NSW Communities & Justice

https://www.facs.nsw.gov.au/resources/statistics/rent-and-sales/dashboard

[4] NSW Postal Geo Data

http://www.corra.com.au/australian-postcode-location-data/

[5] "Sydney metropolitan Postcode". Equifax

https://www.prospectshop.com.au/resources.aspx

[6] "NSW all criminal incidents recorded". NSW Bureau of Crime Statistics and Research

https://www.bocsar.nsw.gov.au/Pages/bocsar_datasets/Datasets-.aspx