

Capstone for Foursquare

Keith Goh, keith_goh@mymail.sutd.edu.sg

December 13, 2019

1 Introduction

2 Background

Berkeley is the location where many students go to for their undergraduate exchange here in SUTD.

However for STEP students we are expected to source our own housing during our time there.

Hence in this project I would like to create a systematic approach to look at all the neighbourhoods in Berkeley and find the cheapest neighbourhood by median pricing to live in that is most similar to Singapore

2.1 Solution

In this capstone, we hypothesize a situation where a Singaporean kid would have to study a term in Berkeley. However housing is not paid for so the student has to get the cheapest housing that he can find. However, the student is unfamiliar with Berkeley and would like to rent a house as close as possible to Singapore. Hence we are building a simple recommender System for similar neighbourhoods to the one in Singapore and creating a ranking based on price depending on the neighbourhoods. So it is a build up of the previous courses on Coursera.

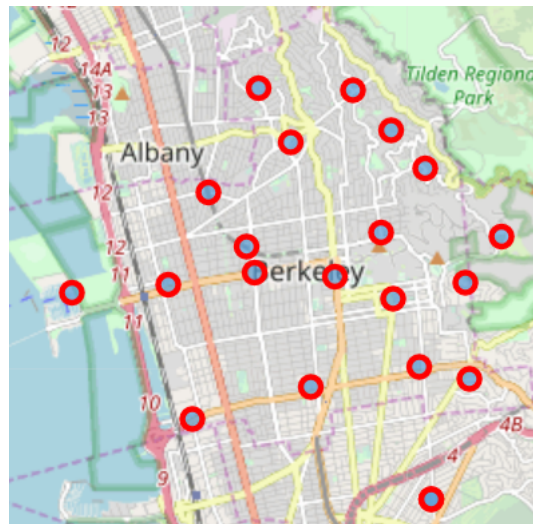
3 Data

3.1 Foursquare

I will then be using foursquare api for trending data and recommending similar to Singapore neighborhoods

3.2 Pricing

I will also be using data from <http://aakashjapi.com/housing-prices-in-berkeley/> for Berkeley housing prices by street.



	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Berkeley Hills	Playground	Park	Intersection	Trail	Dance Studio	Event Space	Electronics Store	Donut Shop	Doctor's Office	Dive Bar
1	Berkeley Marina	Hotel	Park	American Restaurant	Harbor / Marina	Boat or Ferry	Japanese Restaurant	Grocery Store	Restaurant	Gym / Fitness Center	Concert Hall
2	Central Berkeley	Park	Thai Restaurant	Liquor Store	Coffee Shop	Pizza Place	Lounge	Bus Stop	Mexican Restaurant	Electronics Store	Martial Arts Dojo
3	Claremont, Berkeley	American Restaurant	Coffee Shop	Hotel	Bakery	Comic Shop	Nightlife Spot	Resort	Restaurant	Spa	Food Truck
4	Cragmont, Berkeley	Park	Playground	Food	Noodle House	Trail	Dance Studio	Electronics Store	Donut Shop	Doctor's Office	Dive Bar

3.3 Geocoder

Geocoder allows me to get latitude and longitude data for locations to plot on a map and visualize the project.

3.4 Wikipedia

I used wikipedia for a list of neighborhoods in Berkeley

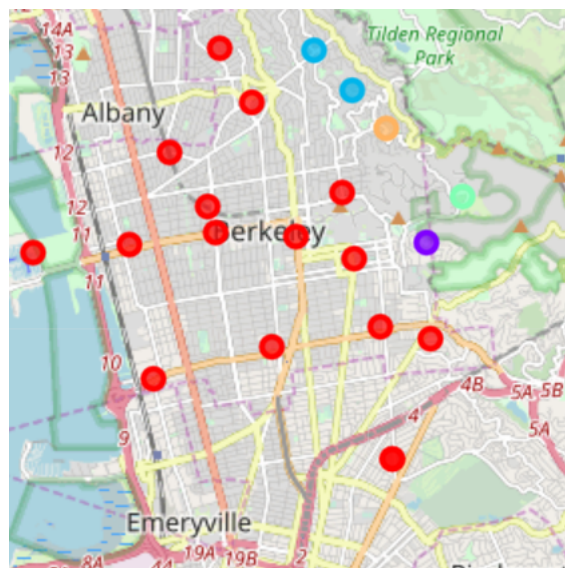
4 Methodology

I first used pandas in order to clean the data. I used folium in order to help visualize the locations at Berkeley and to ensure that the recommended locations were not too far away from the places that i would want to go to.

The First Analysis was done using Folium Map. I ensured that the data points were all in and around the University of Berkeley and removed data points which were far away.

I then used Foursquare in order to get trending data to cluster the points by.

	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
Berkeley Hills	37.889657	-122.258631	2	Playground	Park	Intersection	Trail	Dance Studio	Event Space	Electronics Store	Donut Shop	Doctor's Office
Central Berkeley	37.870495	-122.281958	0	Park	Thai Restaurant	Liquor Store	Coffee Shop	Pizza Place	Lounge	Bus Stop	Mexican Restaurant	Electronics Store
Claremont,Berkeley	37.856090	-122.245200	0	American Restaurant	Coffee Shop	Hotel	Bakery	Comic Shop	Nightlife Spot	Resort	Restaurant	Spa
Downtown Berkeley	37.870101	-122.268139	0	Coffee Shop	Japanese Restaurant	Ice Cream Shop	Yoga Studio	Theater	Movie Theater	Music Venue	Pizza Place	Café
Elmwood,Berkeley	37.857785	-122.253869	0	Chinese Restaurant	Italian Restaurant	Theater	Dance Studio	Coffee Shop	Café	Grocery Store	Cosmetics Shop	American Restaurant



5 Results

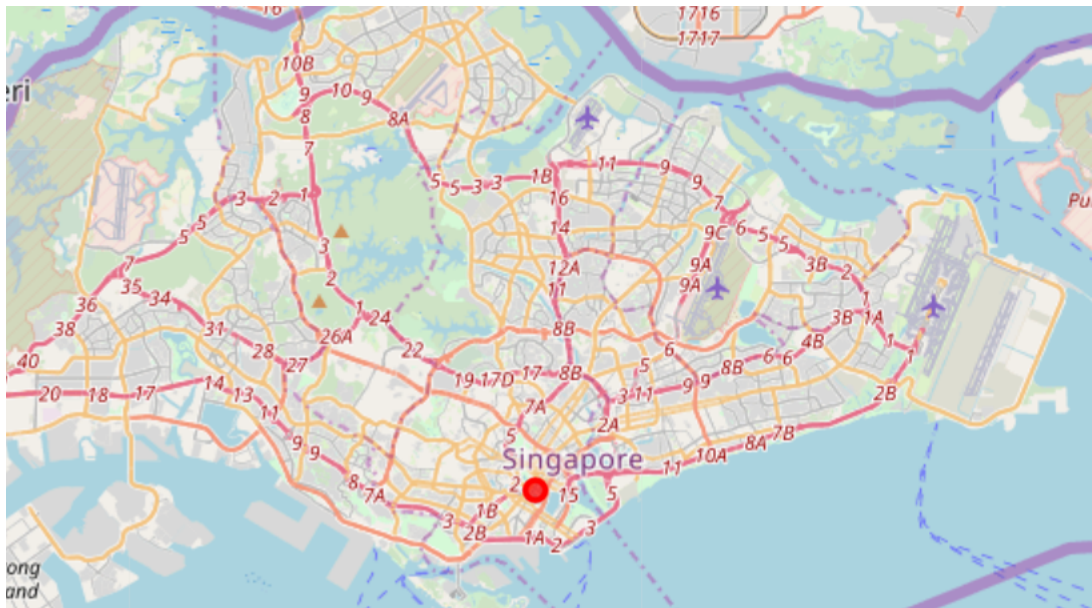
From my analysis it seemed that there were quite a number of neighborhoods that were similar to Singapore hence I had a good variety of choices.

As seen here Singapore is a red point which is similar to many of the different points in Berkeley, Hence it is in the same cluster as them and hence since the majority of points in Berkeley are red, it seems that Singapore is a close match to Berkeley neighborhoods. However the Lowest median price was at Westbrae,Berkeley . This kinda made sense as it was at the outskirts of the town.

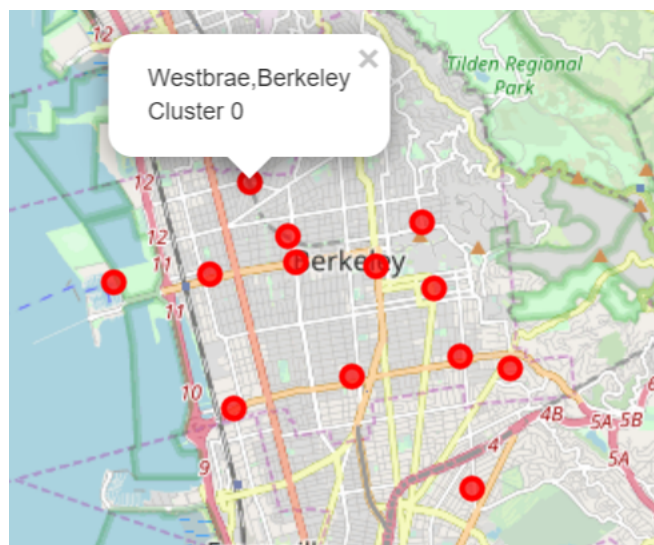
6 Discussion

I then sorted the prices by neighborhood and took the median of the prices This was because it would be unlikely that i would be able to get the cheapest prices for that neighborhood and this is an indicator of how much i would spend.

However, I am not really that sure about the housing market in Berkeley and would need



	latitude	longitude	address	median	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Central Berkeley	37.867078	-122.286791	2171 Allston Way	1235.250000	37.870495	-122.281958	0	Park	Thai Restaurant	Liquor Store	Coffee Shop	P P
Rockridge,Berkeley	34.161532	-118.505347	1912 Addison St	1356.333333	37.839955	-122.251821	0	Italian Restaurant	Café	Bar	Thai Restaurant	Books
Downtown Berkeley	37.872122	-122.270531	1949 Milvia St	1397.500000	37.870101	-122.268139	0	Coffee Shop	Japanese Restaurant	Ice Cream Shop	Yoga Studio	The
Claremont,Berkeley	36.148852	-115.123023	1906 Bonita Ave	1421.000000	37.856090	-122.245200	0	American Restaurant	Coffee Shop	Hotel	Bakery	Cc S
Southside,Berkeley	37.864419	-122.264696	2223 Dwight Way	1467.142857	37.866954	-122.258415	0	Accessories Store	Coffee Shop	Pool	Pizza Place	Japar Restau
Thousand Oaks,Berkeley	49.271359	-123.044902	1922 Grant St	1318.000000	37.895323	-122.281410	0	Gym / Fitness Center	Japanese Restaurant	Coffee Shop	Thai Restaurant	Gi Restau
Golden Gate,Berkeley	37.730689	-122.450376	2225 Hearst Ave	1225.071429	37.850775	-122.292499	0	Gym	Sandwich Place	Café	Camera Store	Cc S
Northside,Berkeley	37.874599	-122.264297	2283 Hearst	1624.000000	37.876052	-122.260311	0	Café	Sandwich Place	Pizza Place	Coffee Shoo	In Restau



to spend more time researching it.

The k-clustering might have given us too many results and we may want to narrow it down

We may have outdated data regarding the prices

Foursquare trending is constantly changing hence I would need to check back again before going to Berkeley

7 Conclusion

Though I managed to get a result from this experiment.

I would like to emphasize that the trending data is constantly changing hence I would probably need to update my results before going to Berkeley.

I would also need to find prices of rent more up to date before going to Berkeley.

I would also need to find out more about the neighborhoods regarding safety before deciding about the neighborhood.

However this is a good start for me to start exploring the locations to stay.