The Battle of Neighborhood

Final Report

By

Prabin Shilpakar

Ideal local to open Indoor Play Center in the Greater Melbourne

Introduction

Melbourne metropolitan or the Greater Melbourne is the second most populous city in Australia and the capital of the State of Victoria. The city is extended to 9,993 km^2 consisting of 31 municipalities also known as Local government area (LGA) or city councils. This vibrant and multi-cultural city is the home to 5 million population. It is also a sports capital of Australia. It hosts major annual international sport events such as the Australian Open, the Melbourne Cup, and the Australian Grand Prix. Melbourne is highly rated for entertainment, tourism, and sport. It is also a research and education hub of Australia.

Melbourne is ranked at top 10 world's most livable city for a decade. That's why the city attracts thousands of migrants from around the world annually. So, the city becomes the major economic hub of Australia and an important financial center of the world.

Melbourne is the fastest growing city in Australia, ahead of Sydney due to more affordable housing and cost of living. Growing population bring more business opportunities to fulfill the needs of the citizens. The city major economy includes finance, manufacturing, research, IT, education, transportation and tourism. It attracts many more businesses into the market. With the opportunities there come the competition and risk in the business. So, Market analysis is extremely important to start new business ventures. Data driven decision making practice allows to make smart choice, reduce risk, and develop profitable business.

Business Problem or Idea Description:

When thinking about the business idea in the context of the fastest-growing city like Melbourne, the first thing that strikes in my mind is the rapid change in the landscape and construction work happening everywhere in the city. Exponentially growing traffic consuming a major amount of people's time in commute to the workplace, meaning less time left for kids and family. Especially it is incredibly challenging for parents with the kids to balance between family and work. Physical activities, entertainment, and socialization are important parts of children's growth. So, every parent is keen to take their kids to park or playground and spend some quality time with them. They are looking for a community park and entertainment center which is easily accessible so that they can maximize the utilization of their spare time.

Indoor play centers are becoming popular in the Greater Melbourne communities, especially among the parents who have toddlers and young kids. Since, indoor playground facilities are safer and can be utilized at any weather condition, parents love to bring their children and letting them have fun and be happy. So, for somebody who love kids or passionate about making children happy, starting an indoor play center that can assure safety, affordability, and entertaining, can be an ideal business choice. As with any business to survive and earn profit in the competitive environment, a detail market study is essential.

Besides, basic preparation to start a business, finding strategic location is the most important part to open a successful business. We must analyze various factors before selecting a location that includes:

1. Population distribution in the city
2. Segmentation of LGA
3. Greater Melbourne Demographics
4. Number of Household with young children in each council
5. Availability of indoor playground venues near residential location
6. Location feasibility
7. Competitor in the market, etc.

Target Audience:

The outcome of this project should be beneficial to anyone, who are thinking about starting new business that focus children or wanting to expand the already established indoor play center at new location. As a Data scientist my objective is to locate and recommend an ideal location in the neighborhood of the Greater Melbourne area to start a new indoor Play Center. In this project we will analyze children population density of each city council, availability of indoor playground facilities in the neighborhood, and identify the possible location where the demand is higher for children play center.

Data

In this project we are analyzing the population and demographics of Greater Melbourne. We will use the following datasets to determine the location that has a higher demand for the indoor activities center for young children, who are aged between 0 to 15 years.

1. **Data1: Location data that contains Lat/Long and area of LGA of Melbourne**
2. **Data2: Census data that contains population statistics of Melbourne**
3. **Data3: Neighborhood venues from Foursquare API for market analysis**

Methodology and Data Analysis

Data1:

Greater Melbourne has 31 Local Government Area (LGA) or city councils. So, to understand the distribution of children population and segment the neighborhoods, we essentially need a dataset that contains the 31 LGA, latitude/Longitude coordinates of LGA, and area. Unfortunately, I am unable to find the online dataset that contains Lat/Long coordinates for each LGA of the Greater Melbourne area. Thus, to extract the coordinates first we need to get the boundary map of the LGA of Greater Melbourne, which can be downloaded from the Australian Bureau of Statistics (ABS) website.

Link to the website: https://itt.abs.gov.au/itt/r.jsp?ABSMaps

Local Government Areas ASGS ED 2016 Digital Boundaries data was downloaded in ESRI shapefile format. The dataset contains the entire LGA of Australia, but we are only interested in LGA of Greater Melbourne. So, to extract only the LGA of Greater Melbourne I cleaned the dataset using QGIS software. Then, I used the analysis tool in QGIS to calculate the centroid of each LGA polygonal area, which enables us to extract the geographic coordinate information. Finally, the Data1 was exported as a CSV file. The data outcome is the Table 1 shown below. We have 31 LGA geographic coordinates and their area information.

Data 2:

For Greater Melbourne population Statistics that includes the total number of children in each LGA or the number of households with young children in each LGA we use the 2016 census data of Australia available in the Australian Bureau of Statistics (ABS) website. Note that Australian census data are available in every five-year interval since the latest census was acquired in 2016, we must rely on that dataset.

Link to the website: http://stat.data.abs.gov.au/#

The CSV data contains all the Local government area of Australia. It is a large file, so I upload it to IBM Database. We need to clean the dataset that only shows the population information of the Greater Melbourne area. I used Python code in JupyterLab for cleaning the dataset. I also use an excel spreadsheet to facilitate the data cleaning process. The outcome of the population data as shows in Table 2. Then, I joined and Table 1 and Table 2 and calculate population density of each LGA. Direct comparing of population data between cities with variable size is not possible or statistically don't make any sense. So, we need the population density for statistically reasonable comparison. We use simple density function, which is total number of population per unit area for calculation. With the population density in the dataframe, we can do statistical analysis and geospatial analysis of the dataset. The output data looks as below.

	City	Area	Pop_Density	Pop_Den_Age(0-15)
0	Banyule	62.5402	1948.59	371.92
1	Bayside	37.2111	2609.09	528.82
2	Boroondara	60.1774	2778.97	516.77
3	Brimbank	123.3991	1574.72	309.69
4	Cardinia	1282.5678	73.39	17.96

Table 1: LGA Location Data with Latitude and Longitude.

	LGA_CODI	LGA_NAME	State	Area	xcoord	ycoord
1	20660	Banyule (C)	Victoria	62.5402	145.0851409	-37.73044097
2	20910	Bayside (C)	Victoria	37.2111	145.0181539	-37.94183906
3	21110	Boroondara (C)	Victoria	60.1774	145.0630891	-37.81928627
4	21180	Brimbank (C)	Victoria	123.3991	144.8033107	-37.74755294
5	21450	Cardinia (S)	Victoria	1282.568	145.5726853	-38.08301379
6	21610	Casey (C)	Victoria	409.4289	145.3090109	-38.10435017
7	21890	Darebin (C)	Victoria	53.4711	145.0143331	-37.73345861
8	22170	Frankston (C)	Victoria	129.598	145.1737988	-38.1346082
9	22310	Glen Eira (C)	Victoria	38.6906	145.0426025	-37.90120582
10	22670	Greater Dandenong (C)	Victoria	129.5469	145.1900091	-38.00558933
11	23110	Hobsons Bay (C)	Victoria	64.2404	144.8297577	-37.85472062
12	23270	Hume (C)	Victoria	503.8505	144.8245343	-37.58918224
13	23430	Kingston (C) (Vic.)	Victoria	91.3677	145.1050723	-37.98858353
14	23670	Knox (C)	Victoria	113.9115	145.2586697	-37.88898225
15	24210	Manningham (C)	Victoria	113.3478	145.1885608	-37.76159035
16	24330	Maribyrnong (C)	Victoria	31.2254	144.8777723	-37.79526741
17	24410	Maroondah (C)	Victoria	61.4097	145.2652614	-37.80487232
18	24600	Melbourne (C)	Victoria	37.3513	144.9439049	-37.81049809
19	24650	Melton (C)	Victoria	527.5374	144.6259818	-37.69300563
20	24970	Monash (C)	Victoria	81.4829	145.1440335	-37.89714147
21	25060	Moonee Valley (C)	Victoria	43.1386	144.8957617	-37.74949992
22	25250	Moreland (C)	Victoria	50.9519	144.948888	-37.72991856
23	25340	Mornington Peninsula (	Victoria	724.1679	145.0353317	-38.34168734
24	25710	Nillumbik (S)	Victoria	432.3373	145.2368043	-37.62291463
25	25900	Port Phillip (C)	Victoria	20.7092	144.9657972	-37.85181579
26	26350	Stonnington (C)	Victoria	25.6514	145.0336569	-37.85795483
27	26980	Whitehorse (C)	Victoria	64.28	145.1550599	-37.83024851
28	27070	Whittlesea (C)	Victoria	489.6939	145.0776927	-37.54721249
29	27260	Wyndham (C)	Victoria	542.0949	144.6182524	-37.886647
30	27350	Yarra (C)	Victoria	19.5425	144.9985486	-37.80228082
31	27450	Yarra Ranges (S)	Victoria	2468.205	145.6970366	-37.71427637

Table 2: Population data of 31 LGA of Greater Melbourne.

LGA_NAME	LGA_CODE	State	Total Population	Age(0-15)	Age(16-24)	Age(25-40)	Age(41-65)	Age(66 above)
Banyule	20660	Victoria	121865	23260	12990	26676	39036	19903
Bayside	20910	Victoria	97087	19678	9836	14479	35518	17576
Boroondara	21110	Victoria	167231	31098	22956	34362	53597	25218
Brimbank	21180	Victoria	194319	38216	25072	47319	59321	24391
Cardinia	21450	Victoria	94128	23032	10868	21536	28289	10403
Casey	21610	Victoria	299301	71957	37517	71794	89633	28400
Darebin	21890	Victoria	146719	24611	17110	42581	42551	19866
Frankston	22170	Victoria	134143	26587	14963	29316	44065	19212
Glen Eira	22310	Victoria	140875	27310	16104	33881	43639	19941
Greater Dandenong	22670	Victoria	152050	29116	19435	40141	42821	20537
Hobsons Bay	23110	Victoria	88778	17247	8936	21114	29183	12298
Hume	23270	Victoria	197376	46697	25823	47872	58190	18794
Kingston	23430	Victoria	151389	28366	15417	32919	50138	24549
Knox	23670	Victoria	154110	28882	18284	32578	52193	22173
Manningham	24210	Victoria	116255	20178	13324	20902	38398	23453
Maribyrnong	24330	Victoria	82288	13849	10238	28112	22650	7439
Maroondah	24410	Victoria	110376	21916	11996	24508	35424	16532
Melbourne	24600	Victoria	135959	9203	39877	55653	23107	8119
Melton	24650	Victoria	135443	35123	15735	34822	39417	10346
Monash	24970	Victoria	182618	30832	28321	40929	52851	29685
Moonee Valley	25060	Victoria	116671	20782	13114	28568	36585	17622
Moreland	25250	Victoria	162558	27397	18697	51968	43335	21161
Mornington Peninsula	25340	Victoria	154999	28663	14527	23127	52694	35988
Nillumbik	25710	Victoria	61273	12994	7643	9860	23432	7344
Port Phillip	25900	Victoria	100863	12320	8895	37900	31028	10720
Stonnington	26350	Victoria	103832	13704	13407	33632	28235	14854
Whitehorse	26980	Victoria	162078	29185	21522	35421	49140	26810
Whittlesea	27070	Victoria	197491	44345	22972	52017	56553	21604
Wyndham	27260	Victoria	217122	56452	23473	64815	57467	14915
Yarra	27350	Victoria	86657	10131	9304	35914	22865	8443
Yarra Ranges	27450	Victoria	149537	30370	16771	29045	51805	21546

Popolation Data analysis:

The bar chart shown in figure 2 shows that some city such as Yarra Ranges has largest area, but least population density, whereas city like Glen Eira has small area but higher population density. Similarly, Port Phillip has highest population density when compare the population of all age. But, we are only interested in the city which has higher concentration of young children. So, I run another statistical analysis. As shown in figure 3 Glen Eira has the highest population density of people age between 0-15 years old, follows by 'Port Phillip', 'Moreland', 'Stonnington', and 'Bayside'. Similarly, when we check the choropleth map (figure 4), all the major cities with higher population density of children location at the center of Greater Melbourne around city center. Now, we have Top 10 candidates based on population analysis. We need to do further investigation to narrow down the best location for Indoor play center.

Figure 2: Comparison of Population Density of Greater Melbourne**.**

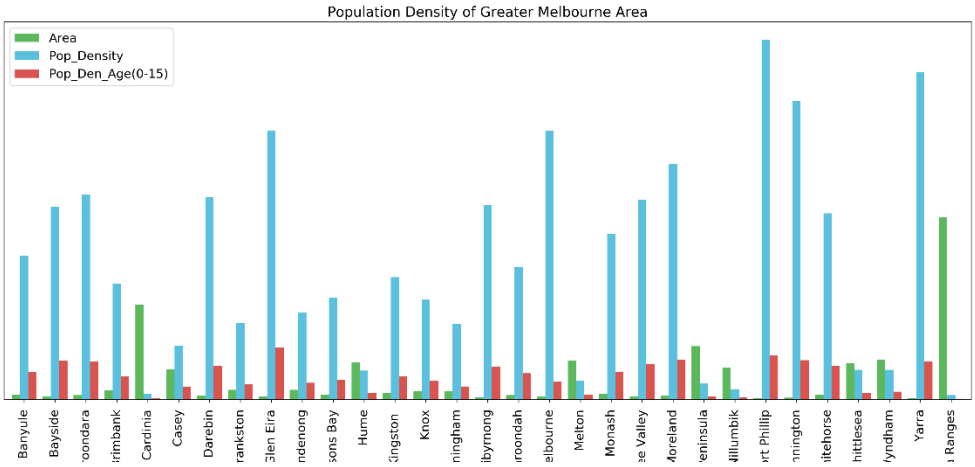


Figure 3: Population density plot of Young Children in Greater Melbourne LGA

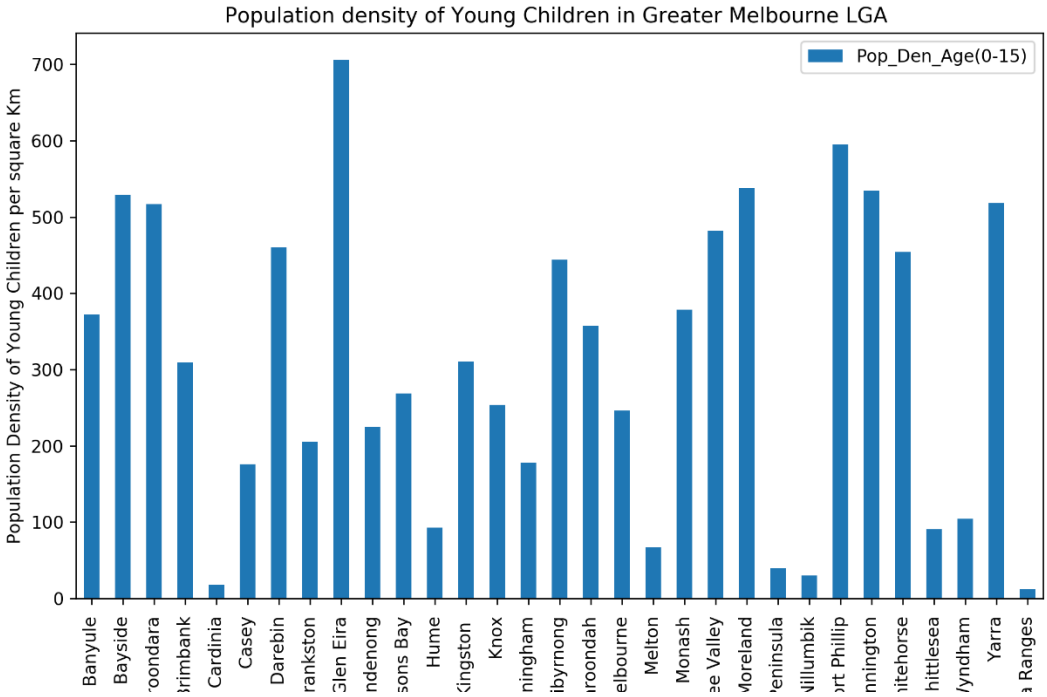
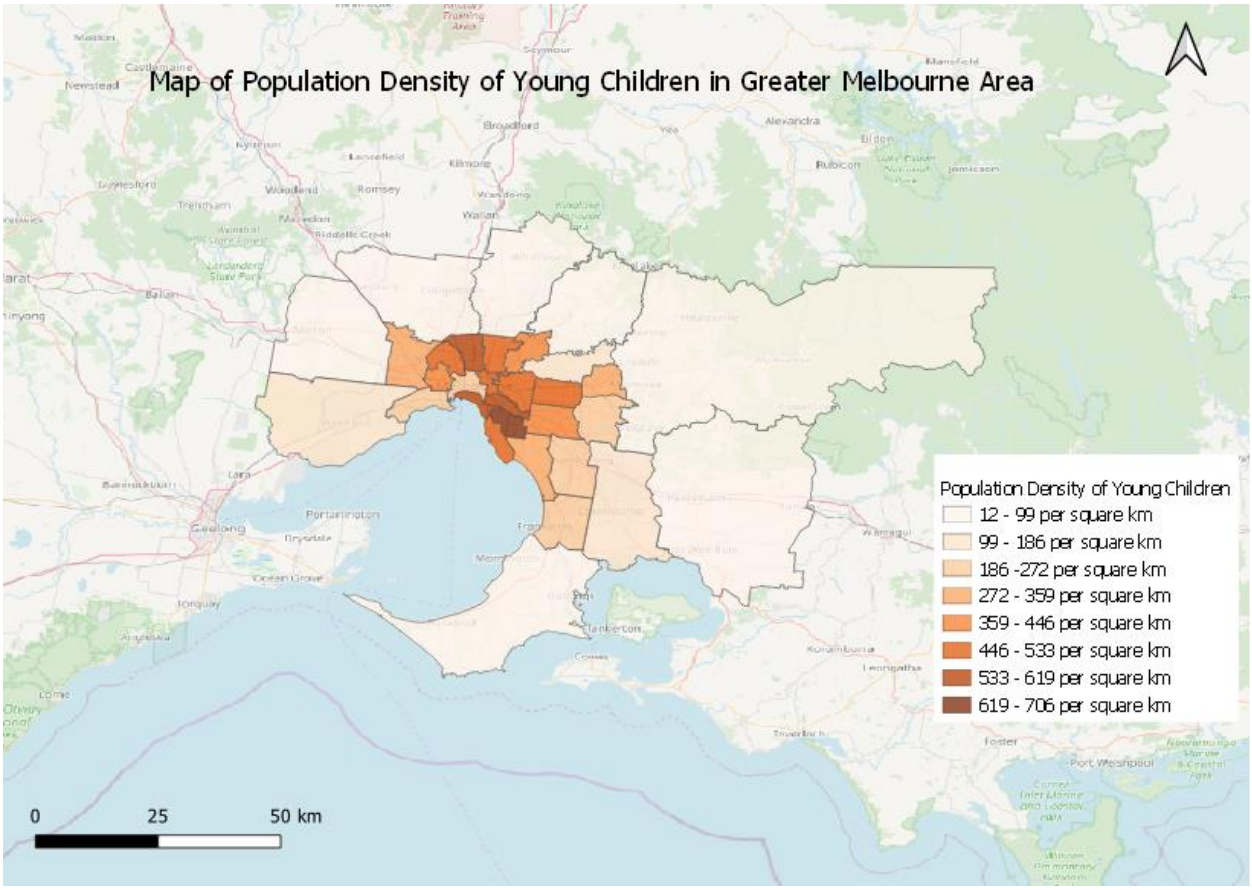


Figure 4: Choropleth map showing population density of young children in Greater Melbourne.



We find top 10 cities where the density of young children population is higher, which means these city councils have a higher demand for Kids Activities Center. Now we need to analyze neighborhood venues in Greater Melbourne so that we can track down the competition in the market and the factor that can influence the decision. We will use the Foursquare API for the analysis.

Data 3:

First, I calculated the geographic coordinate of the center point of Greater, which is obtained using QGIS coordinate extraction tool is boundary shape file. Then, I created a dataframe that only contain Top 10 cities than fall under our selection. The following table 3 shows the top 10 cities. The geographical coordinates data of top 10 LGA was obtain from 'Data 1', which I used as input for the Foursquare API, which will help to get venue information for each neighborhood. We I used the Foursquare API to explore neighborhoods in the Melbourne area that can influence the location for an indoor play center. For example, exploring shopping malls in the neighborhood, where the flow of family hangout is higher during the weekend. Similarly, explore the number of indoor play centers that are already established in the neighborhood. Then, plot violin diagram, which is a great way to visualize frequency distribution datasets. This analysis helps to decide the location that shows high demand for the indoor children play center.

Table 3: Top 10 Cities.

City	Area	Pop_Density	Pop_Den_Age(0-15)
Glen Eira	38.6906	3641.07	705.86
Port Phillip	20.7092	4870.44	594.90
Moreland	50.9519	3190.42	537.70
Stonnington	25.6514	4047.81	534.24
Bayside	37.2111	2609.09	528.82
Yarra	19.5425	4434.28	518.41
Boroondara	60.1774	2778.97	516.77
Moonee Valley	43.1386	2704.56	481.75
Darebin	53.4711	2743.89	460.27
Whitehorse	64.2800	2521.44	454.03

Following steps where used to obtain the Data 3.

1. Explore the number one city in the 'top10' dataframe to become familiar with the data.
2. Get the top 100 venues that are in the city Glen Eira within a radius of 1000 meters
3. Create a nearby venues function for all top 10 cities in Greater Melbourne. The outcome looks as follows.

	City	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Glen Eira	-37.901206	145.042602	Gauge Espresso	-37.903569	145.039953	Café
1	Glen Eira	-37.901206	145.042602	On Top Bar	-37.903473	145.039006	Bar
2	Glen Eira	-37.901206	145.042602	Mountain Of Bears	-37.904180	145.039860	Coffee Shop
3	Glen Eira	-37.901206	145.042602	Jimmy Grants	-37.904230	145.040280	Greek Restaurant
4	Glen Eira	-37.901206	145.042602	Remezzo	-37.904220	145.040200	Tapas Restaurant
...	...	...	...	...	...	...	...
95	Stonnington	-37.857955	145.033657	Prickly Pear	-37.863110	145.027990	Café
96	Stonnington	-37.857955	145.033657	Neon Tiger	-37.852540	145.036295	Café
97	Stonnington	-37.857955	145.033657	Malvern Hotel	-37.851592	145.029939	Beer Garden
98	Stonnington	-37.857955	145.033657	Cooper and Milla's	-37.855934	145.023722	Breakfast Spot
99	Stonnington	-37.857955	145.033657	Gardiner And Field	-37.858667	145.022942	Café

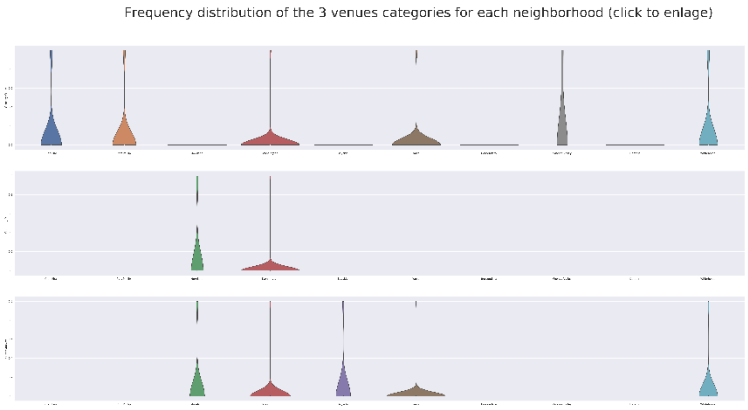
4. Check how many venues were returned for each neighborhood.100 venues limit is imposed by the free Foursquare account.
5. Find out how many unique categories can be curated from all the returned venues. The result shows 102.
6. Create the dataframe that contains the mean of the frequency of occurrence of each category for each city.
7. Obtained the Top 10 venues list.

	City	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	Bayside	Café	Sports Bar	Basketball Court	Convenience Store	Playground	Field	Sports Club	Supermarket	Park	Gas Station
1	Boroondara	Café	Park	Bar	Train Station	Convenience Store	Pizza Place	Dumpling Restaurant	Seafood Restaurant	Italian Restaurant	Light Rail Station
2	Darebin	Café	Light Rail Station	Food Truck	Pizza Place	Bus Line	Board Shop	Japanese Restaurant	Fast Food Restaurant	Indian Restaurant	Martial Arts Dojo
3	Glen Eira	Café	Burger Joint	Grocery Store	Sandwich Place	Japanese Restaurant	Breakfast Spot	Lighting Store	Indian Restaurant	Food & Drink Shop	Coffee Shop
4	Moonee Valley	Grocery Store	Coffee Shop	Park	Café	Bus Line	Food & Drink Shop	Wine Bar	Fish & Chips Shop	Electronics Store	Falafel Restaurant
5	Moreland	Café	Shopping Mall	Bakery	Supermarket	Chinese Restaurant	Soccer Field	Paper / Office Supplies Store	Liquor Store	Fish & Chips Shop	Park

Now, we have obtained the venues information which are location within one-kilometer radius from the center of each Top 10 cities. We can check which city has greater concentration of **Shopping Mall**, **Grocery Store** and **Supermarket** where people with kids visited frequently. I have picked these three venues based on my personal experience and market observation.

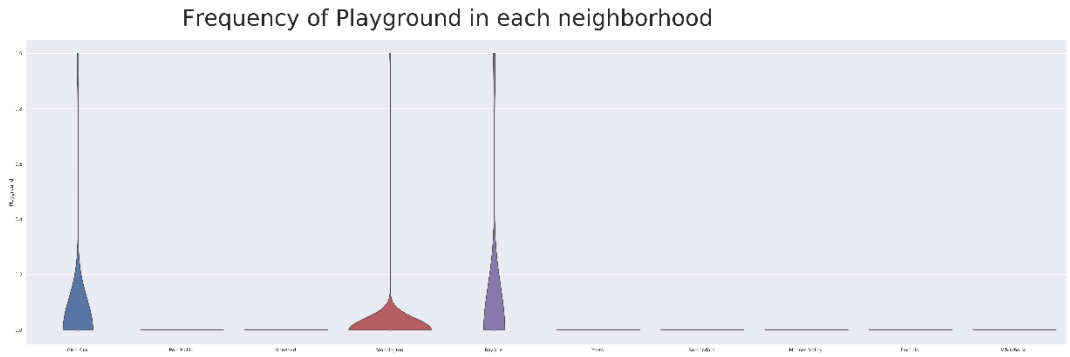
The statistical data was plotted in Violin diagram.

Figure 5: Showing frequency of 3 venues in top 10 cities



The violin plot shows that City of Stonnington has higher frequency of 3 venues follows by Moreland, Yarra, Glen Eira, and Port Phillip. Second, step is to see if there are any Playground or Kids Activities Center in these cities. So, the next violin plot was created.

Figure 6: Frequency of Playground in each neighborhood.



Results

1. Analysis of Population density shows that the higher demand for Indoor Play Center is in the city of Glen Eira, follows by 'Port Phillip', 'Moreland', 'Stonnington', and 'Bayside'
2. The violin plot shows that City of Stonnington has higher frequency of 3 venues follows by Moreland, Yarra, Glen Eira, and Port Phillip.

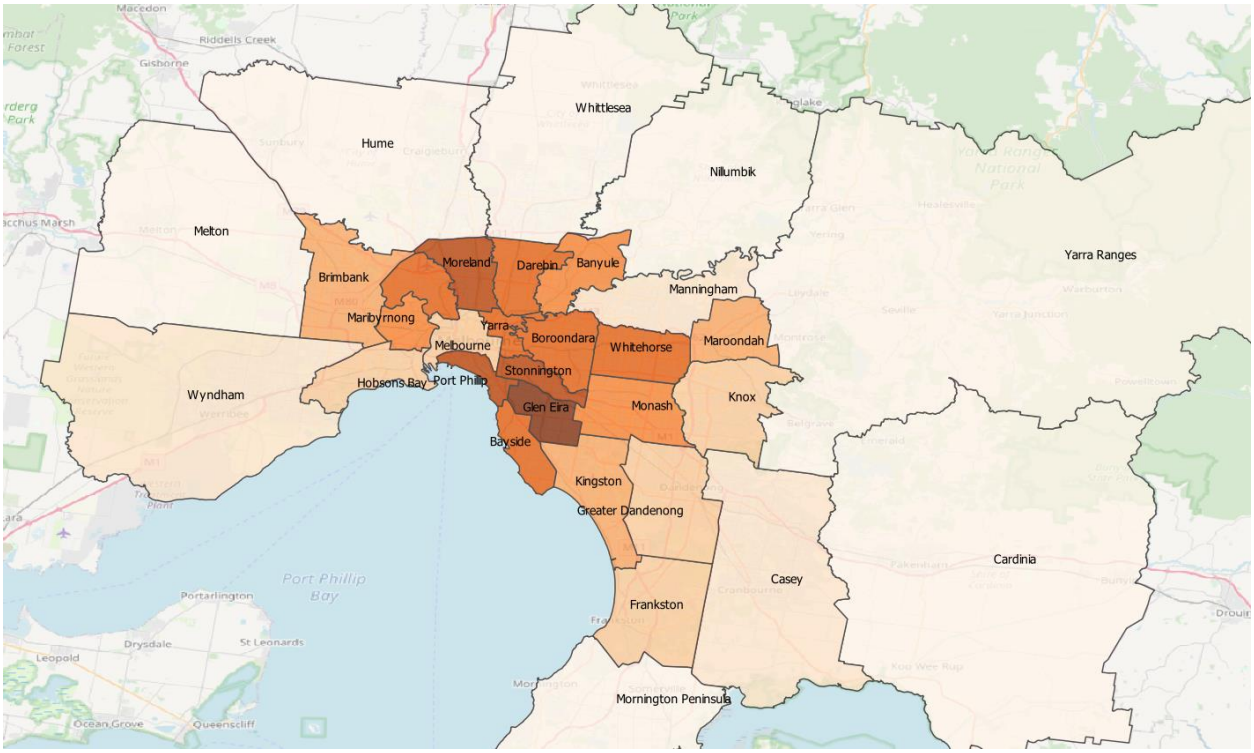
Discussion

The objective of this data analysis project is to locate and recommend an ideal location in the neighborhood of the Greater Melbourne area to start a new indoor Play Center. For this we segmented the data into 31 Local Government Area (LGA) or City Councils that constitute the Greater Melbourne Area. The first result of the population data analysis shows that 'Glen Eira' is the city with the highest population density of young children in the Greater Melbourne area, follows by 'Port Phillip', 'Moreland', 'Stonnington', and 'Bayside'. The folium map plot(figure 4) shows that all these city councils are the neighboring city of central Melbourne. So, it is no surprise that the population density is higher in these LGA. This means these are the location where the demand for Indoor Play Center is higher. Now, to narrow down the ideal location we did further analysis on the venues in this location that may influence the feasibility and profitability of the play center business.

As our major target is the working parents with younger kids who are looking for a place like a community park, entertainment center which is easily accessible so that they can maximize the utilization of their spare time with their kids. So, time and distance to the venue play an important role in this analysis. Therefore, we decided to pick the center point of each LGA to explore for venues using the Foursquare API. Based on the market observation and my personal experience usually families with children are mostly seen near Grocery stores, supermarkets, and Shopping malls. So, we analyzed the frequency distribution of these venues around 1000-meter radius from the center point of each LGA(Figure 5 and 6). The violin plot shows that 'Stonnington' has the highest concentration of all the 3

venues that can influence the business location. Besides Stonnington, all top 5 cities have some frequency of 3 venues. Similarly, 'Moreland' can be picked as our second-best location, where the flow of people is higher because it has a higher frequency of Shopping malls and Supermarkets. 'Glen Eira' fall to the third-best choice in our analysis because of the lack of Shopping Mall and Supermarket. Even though Glen Eira has the highest density of young children, it is the neighboring city that shares most of its boundary with Stonnington(Figure 7), so our selection of Stonnington is still valid.

Figure 7: 31 LGA of Greater Melbourne.



Conclusion

Physical activities and socialization are important parts of children's growth. Every parent wants their kids to enjoy there life and grow up in the best possible environment as possible. They want to give as much time as possible to their children so that they can provide proper guidance and support. In the busy city life balancing time between family and work is very tough. So, city parents are looking for a place like a community park, entertainment center which is easily accessible so that they can maximize the utilization of their spare time with their kids. One of the options can be an indoor play center, especially for the parents who have toddlers and young kids.

For someone who decided to open a new indoor play center should know where is the concentration of his target clients so that he can earn as much profit as possible. Based on our analysis Stonnington city council is an ideal location to open the Indoor Play Center. It has the top 5 population density of young children among all Local Government Area of Greater

Melbourne. It is one of the busiest locations with shopping malls, grocery stores, and supermarkets. Our second best choice is Moreland, follows by Glen Eira, Port Phillip, and Bayside. The accuracy of this project analysis results is limited to the Foursquare API database and selection of the parameters. The selection may change with the availability of an accurate dataset. But, for the statistical and geospatial analysis we are quite confident with the outcome.

References

https://en.wikipedia.org/wiki/Melbourne

Thank you