Introduction

———

Finding the right accommodation and housing is never easy.

Therefore, clustering neighbourhood may become a helpful tool for people who just arrived.

Input Data

— — —

A dump of all Singapore postal codes retrieved from Onemap.sg API

Out[19]:

	Postcode	Neighborhood	Latitude	Longitude
0	018906	SINGAPORE CHINESE CULTURAL CENTRE	1.27582919242101	103.849576388626
1	018907	TEMPORARY SITE OFFICE	1.27494969350595	103.85166524159
2	018910	NIL	1.27958678750876	103.868955673647
3	018915	TEMPORARY SITE OFFICE	1.27368177267254	103.860075469301
4	018925	CITIBANK TRADE_BRANCH	1.27642434235815	103.85475889956

```
neighborhoods.shape
```

```
(155744, 4)
```

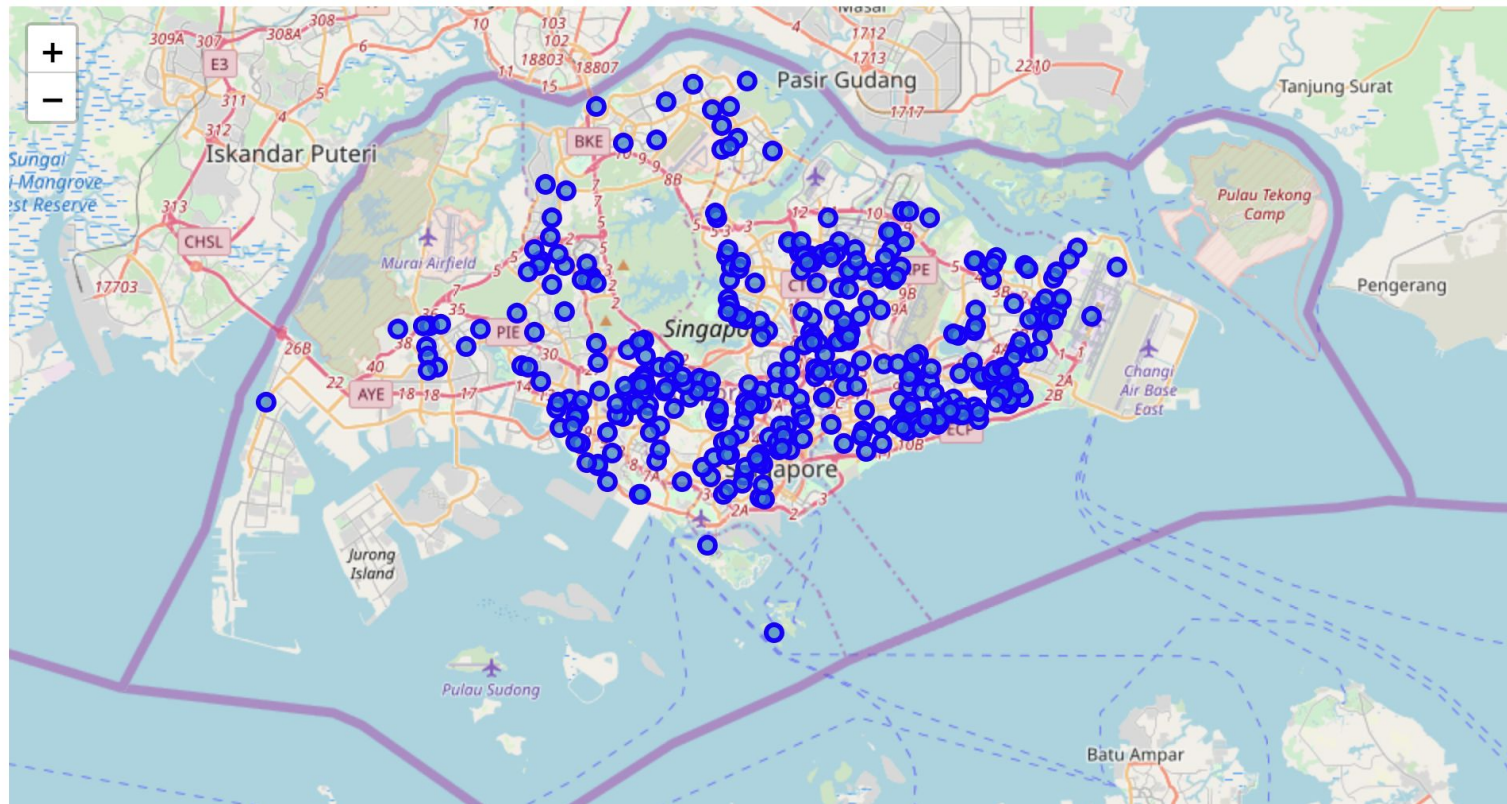# Data Cleaning and Sampling

———

- removing all NIL values
- removing all duplicates
- sampling: randomly select 500 buildings

GEO MAP

Incorporate Foursquare API

`Out[12]:`

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	BEDOKVILLE	1.323681	103.947566	Amy's Laksa	1.323111	103.945957	Noodle House
1	BEDOKVILLE	1.323681	103.947566	Kimly Seafood	1.323110	103.945939	Coffee Shop
2	BEDOKVILLE	1.323681	103.947566	New Changi Eating House	1.323117	103.945954	Asian Restaurant
3	BEDOKVILLE	1.323681	103.947566	Tanah Merah MRT Interchange (EW4)	1.327309	103.946443	Train Station
4	BEDOKVILLE	1.323681	103.947566	Netball Court Quadrangle Bedok South	1.325625	103.950991	Basketball Court
5	BEDOKVILLE	1.323681	103.947566	Changi Naval Base Pick Up Point @ Tanah Merah	1.327365	103.945903	Bus Station
6	BEDOKVILLE	1.323681	103.947566	East Meadows Swimming Pool	1.327519	103.946264	Pool
7	BEDOKVILLE	1.323681	103.947566	Parit Puaka Dalam	1.327666	103.946820	Lake
8	BEDOKVILLE	1.323681	103.947566	Indoor Sports Hall @ Bedok South	1.326071	103.951059	School

Explore Merged Dataset

— — —

- **7371 rows**
- **353 buildings**
- **3815 unique venues**
- **from 320 unique venue categories.**

How many venues in each neighbourhood

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
MAGHAIN ABOTH SYNAGOGUE	100	100	100	100	100	100
PAGODA HOUSE	100	100	100	100	100	100
DBS Resorts World Sentosa Level 1 Visitor Centre	100	100	100	100	100	100
HEAP SENG HOUSE	100	100	100	100	100	100
WATERLOO CENTRE	100	100	100	100	100	100
HERITAGE COURT	100	100	100	100	100	100
ASIA STAR HOTEL	100	100	100	100	100	100
FORT CANNING MRT STATION (DT20)	100	100	100	100	100	100
KIMSIA PARK	100	100	100	100	100	100
68 DUXTON	100	100	100	100	100	100
NANYANG ACADEMY OF FINE ARTS	97	97	97	97	97	97
TEMBELING MANSIONS	88	88	88	88	88	88
THE PATERSON EDGE	77	77	77	77	77	77
UE SQUARE	75	75	75	75	75	75
THOMSON PARK	73	73	73	73	73	73

USE CASE 1: Explore all neighborhoods which is near to the art gallery
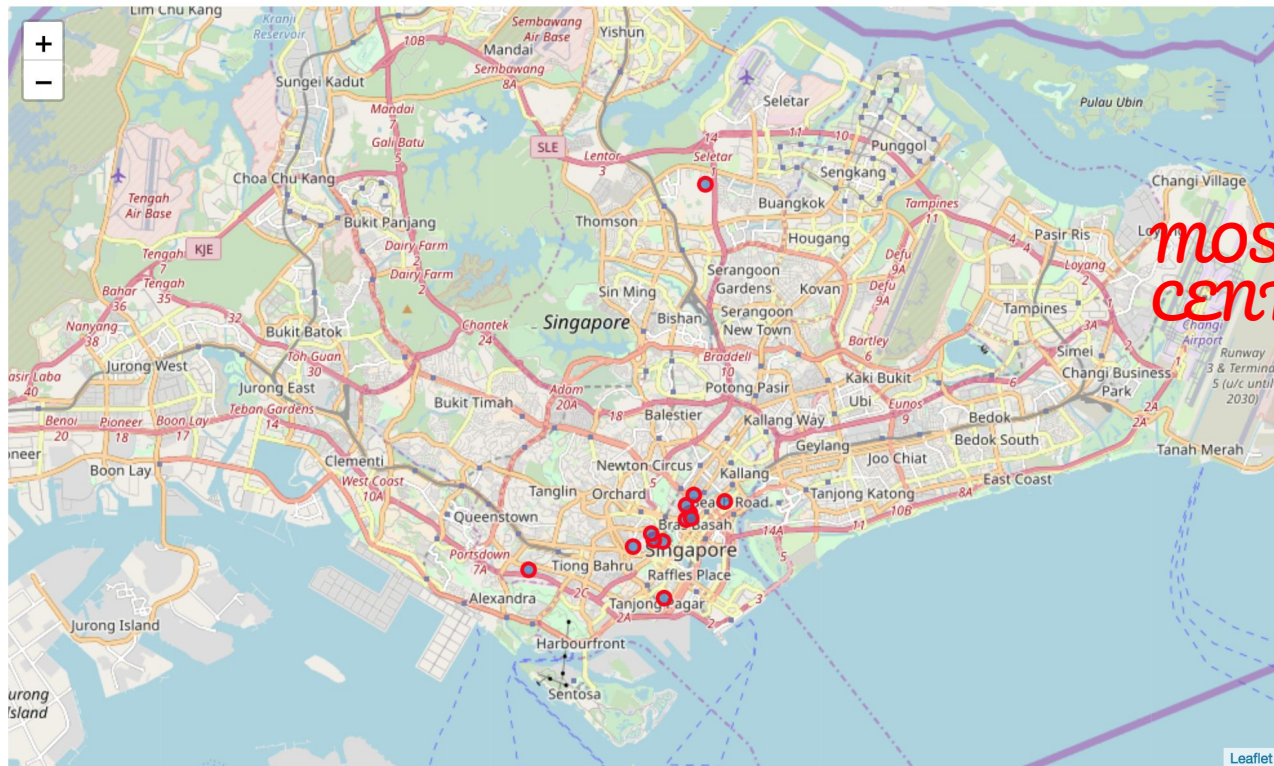
MOST IN THE CENTRAL PART!

Cluster Neighbourhood

```
map_clusters
```

Out[56]:
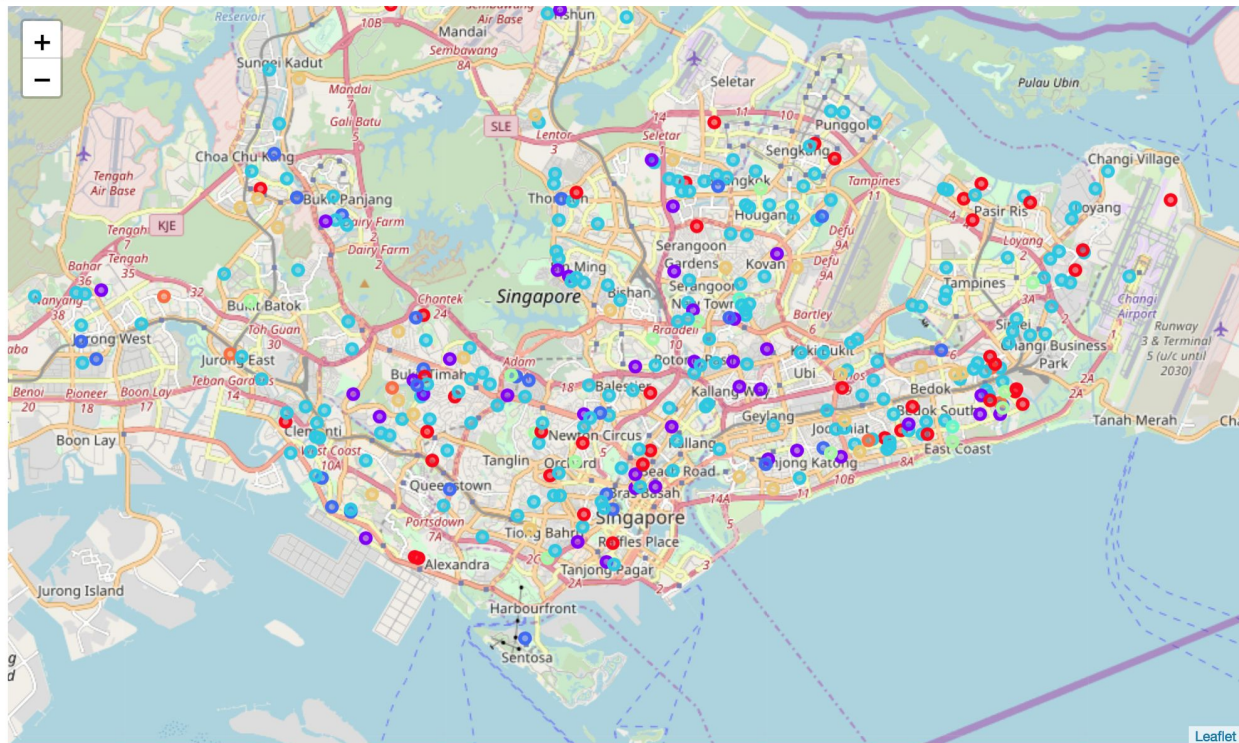
Cluster Neighbourhood: Cluster's size

Cluster Labels	Neighborhood	Latitude	Longitude	1st Most Common Venue	2n Co Ve
0	44	44	44	44	
1	41	41	41	41	
2	26	26	26	26	
3	195	195	195	195	
4	2	2	2	2	
5	13	13	13	13	
6	25	25	25	25	
7	7	7	7	7	

venues uniformly distributed, therefore one big cluster

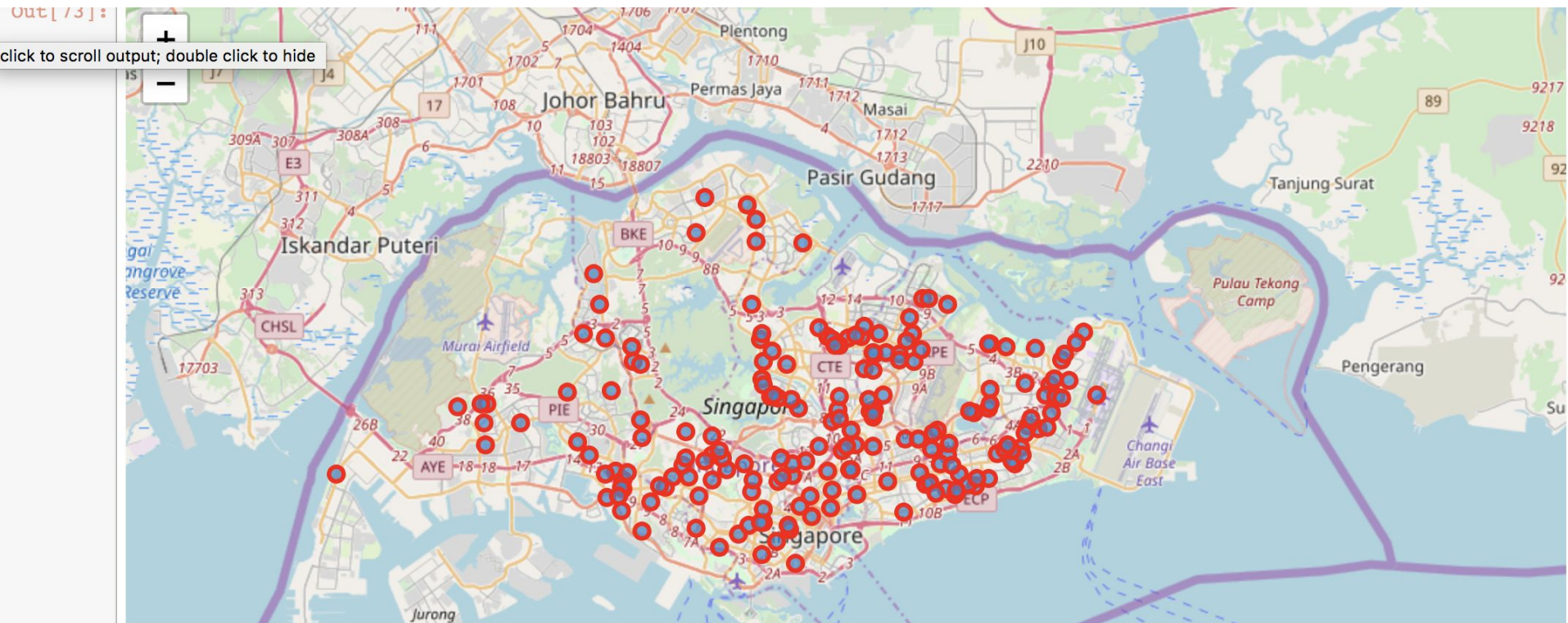# USE CASE 2: THE PATERSON EDGE

— — —

The Business Problem:

We have a customer who recently relocated to Singapore. The first month he was staying in the Paterson Edge Condo. Unfortunately, his monthes fees are very high in this area, but the neirborhood is very lovely. Therefore, our customer is looking for a new house in a less expencive area but with a nice community around him.

SOLUTION:

———

Using our clusters, we are going to find to which cluster The Paterson Edge is belong. Then we will highlight all members of this cluster on the SG MAP, and analise it.

RESULTS

click to scroll output; double click to hide

The Paterson Edge belongs to the biggest cluster. This is good, because our customer can select from different options. For instance, we can propose to our customer consider the SEA VIEW PARK condo in

THANK YOU!