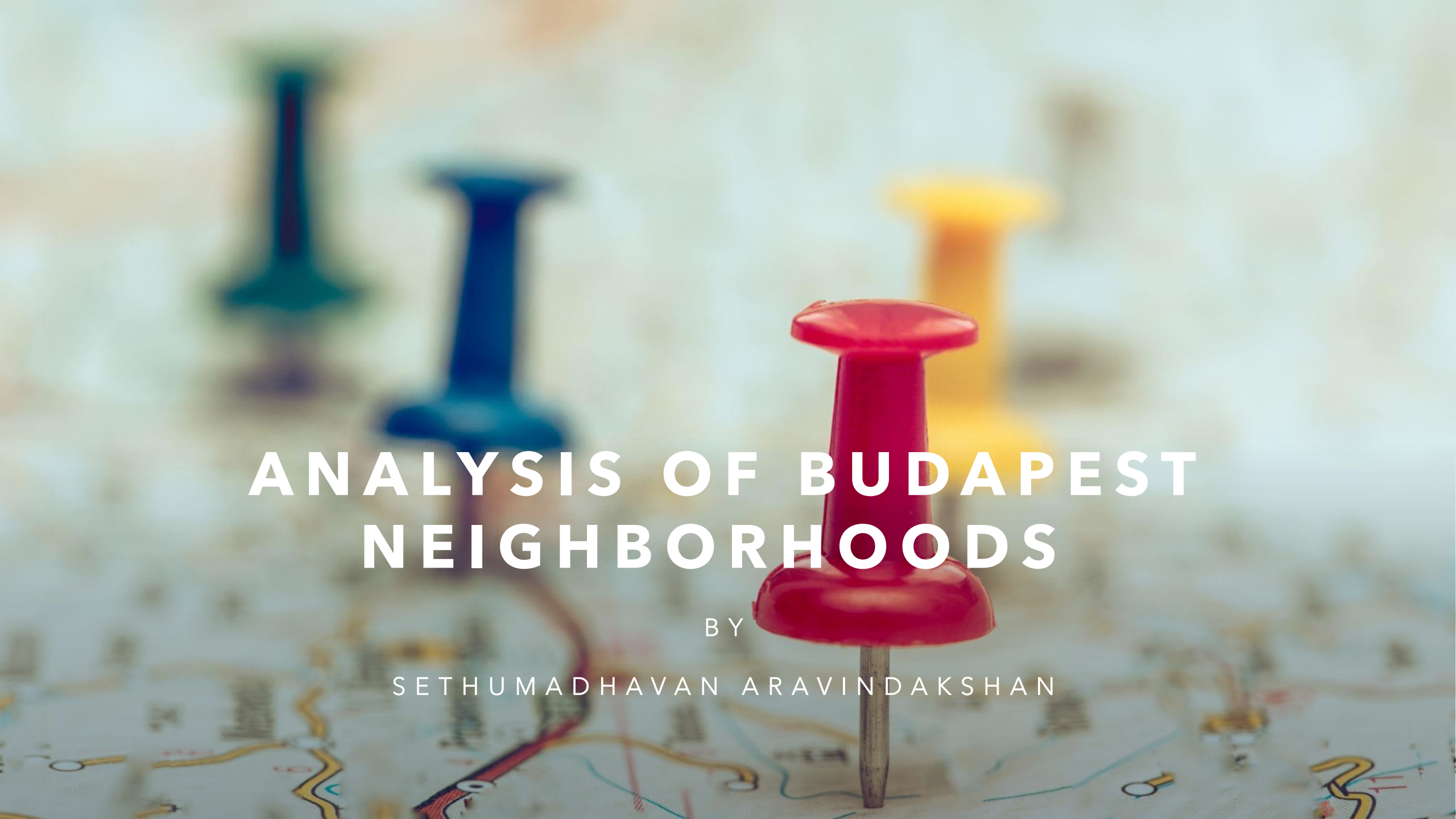


ANALYSIS OF BUDAPEST NEIGHBORHOODS

BY

SETHUMADHAVAN ARAVINDAKSHAN

INTRODUCTION

OBJECTIVE

Our aim is to aid anyone who visits Budapest city for either Employment, Business or Tourism in deciding the place of stay based on what each neighborhood has to offer, and their personal preferences

To obtain that results, First the neighborhoods have to be grouped as follows :

- Fully developed neighborhood
- Moderately developed neighborhood
- Developing neighborhood
- Underdeveloped neighborhood

DATA SOURCES

List of all districts and postal codes of Budapest city

Geographical co-ordinates of each neighborhood is extracted using the Geocoder from Geopy python library.

Venue data from Foursquare API

DATA PREPARATION AND EDA

Data Preparation :

Data Normalization – Min Max scaler

	Arts & Entertainment	College & University	Event	Food	Nightlife Spot	Outdoors & Recreation	Professional & Other Places	Residence	Shop & Service	Travel & Transport
0	0.813793	0.602151	0.75	0.995833	0.744856	0.912088	0.957055	0.630435	0.879581	0.905882
1	0.027586	0.010753	0.00	0.054167	0.000000	0.142857	0.110429	0.021739	0.109948	0.047059
2	0.027586	0.096774	0.00	0.170833	0.024691	0.214286	0.300613	0.239130	0.251309	0.129412
3	0.062069	0.279570	0.00	0.395833	0.094650	0.307692	0.435583	0.260870	0.413613	0.341176
4	0.965517	1.000000	1.00	1.000000	1.000000	1.000000	1.000000	0.913043	0.963351	0.976471

EDA :

Descriptive statistics -

	Arts & Entertainment	College & University	Event	Food	Nightlife Spot	Outdoors & Recreation	Professional & Other Places	Residence	Shop & Service	Travel & Transport
count	23.000000	23.000000	23.000000	23.000000	23.000000	23.000000	23.000000	23.000000	23.000000	23.000000
mean	0.222050	0.294792	0.260870	0.340761	0.201467	0.314399	0.383234	0.291925	0.407407	0.387999
std	0.354476	0.324360	0.380477	0.373558	0.334975	0.321622	0.329554	0.320181	0.332037	0.351402
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.014286	0.027473	0.000000	0.062500	0.006173	0.069892	0.134731	0.059524	0.129630	0.096491
50%	0.035714	0.186813	0.000000	0.170833	0.024691	0.198925	0.257485	0.142857	0.328042	0.263158
75%	0.171429	0.434066	0.500000	0.433333	0.104938	0.467742	0.631737	0.309524	0.677249	0.649123
max	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

MODELLING

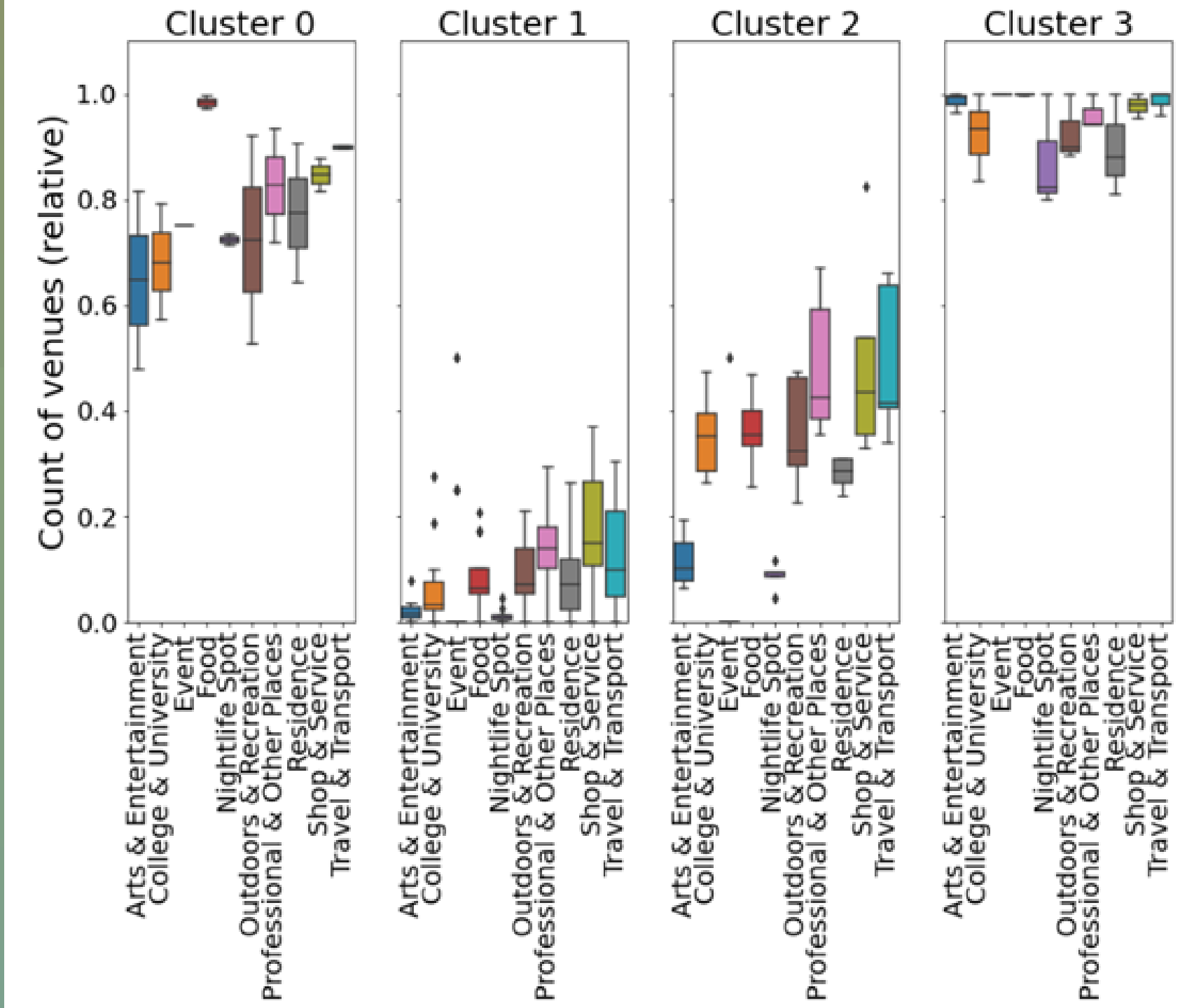
MODEL BUILDING

Unsupervised Machine Learning

Clustering Technique

K-means clustering

CLUSTERS


Cluster 0
Cluster 1
Cluster 2
Cluster 3
Count of venues (relative)
0.0
0.2
0.4
0.6
0.8
1.0
Arts & Entertainment
College & University
Event
Food
Nightlife Spot
Outdoors & Recreation
Professional & Other Places
Residence
Shop & Service
Travel & Transport

RESULTS

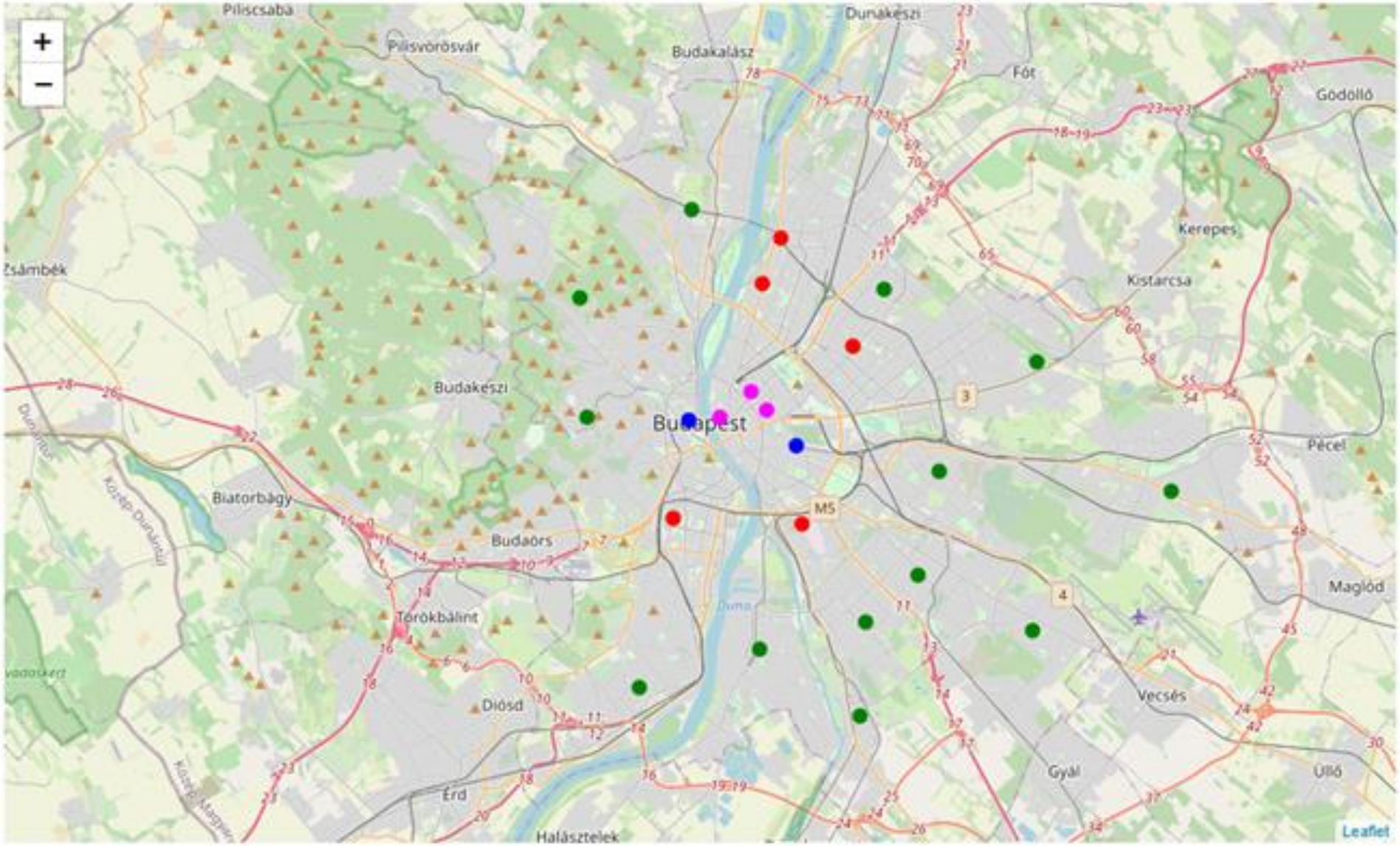
	Arts & Entertainment	College & University	Event	Food	Nightlife Spot	...	Professional & Other Places	Residence	Shop & Service	Travel & Transport	Cluster
0	0.813793	0.602151	0.75	0.995833	0.744856	...	0.957055	0.630435	0.879581	0.905882	3
1	0.027586	0.010753	0.00	0.054167	0.000000	...	0.110429	0.021739	0.109948	0.047059	0
2	0.027586	0.096774	0.00	0.170833	0.024691	...	0.300613	0.239130	0.251309	0.129412	0
3	0.062069	0.279570	0.00	0.395833	0.094650	...	0.435583	0.260870	0.413613	0.341176	2
4	0.965517	1.000000	1.00	1.000000	1.000000	...	1.000000	0.913043	0.963351	0.976471	1
5	1.000000	0.763441	1.00	0.995833	0.794239	...	0.944785	0.847826	1.000000	0.947059	1
6	0.951724	0.946237	1.00	1.000000	0.823045	...	0.987730	1.000000	0.989529	1.000000	1
7	0.462069	0.784946	0.75	0.979167	0.711934	...	0.766871	0.869565	0.790576	0.864706	3
8	0.186207	0.451613	0.00	0.354167	0.086420	...	0.411043	0.239130	0.319372	0.405882	2
9	0.013793	0.021505	0.00	0.066667	0.004115	...	0.288344	0.065217	0.151832	0.211765	0
10	0.096552	0.236559	0.00	0.329167	0.115226	...	0.619632	0.217391	0.801047	0.652941	2
11	0.013793	0.182796	0.25	0.066667	0.000000	...	0.226994	0.130435	0.062827	0.200000	0
12	0.144828	0.387097	0.50	0.445833	0.086420	...	0.711656	0.282609	0.539267	0.417647	2
13	0.075862	0.344086	0.00	0.262500	0.045267	...	0.380368	0.282609	0.287958	0.641176	2
14	0.027586	0.053763	0.00	0.187500	0.008230	...	0.226994	0.130435	0.329843	0.264706	0
15	0.034483	0.268817	0.00	0.037500	0.012346	...	0.141104	0.043478	0.115183	0.094118	0
16	0.006897	0.000000	0.00	0.012500	0.004115	...	0.073620	0.000000	0.104712	0.005882	0
17	0.006897	0.021505	0.00	0.062500	0.000000	...	0.018405	0.108696	0.083770	0.088235	0
18	0.006897	0.075269	0.50	0.204167	0.045267	...	0.196319	0.108696	0.335079	0.305882	0
19	0.006897	0.010753	0.00	0.058333	0.012346	...	0.134969	0.065217	0.151832	0.252941	0
20	0.000000	0.064516	0.00	0.050000	0.004115	...	0.153374	0.043478	0.230366	0.100000	0
21	0.075862	0.032258	0.00	0.100000	0.008230	...	0.104294	0.021739	0.287958	0.023529	0
22	0.013793	0.021505	0.25	0.000000	0.008230	...	0.000000	0.000000	0.000000	0.000000	0

INTERPRETTING RESULTS

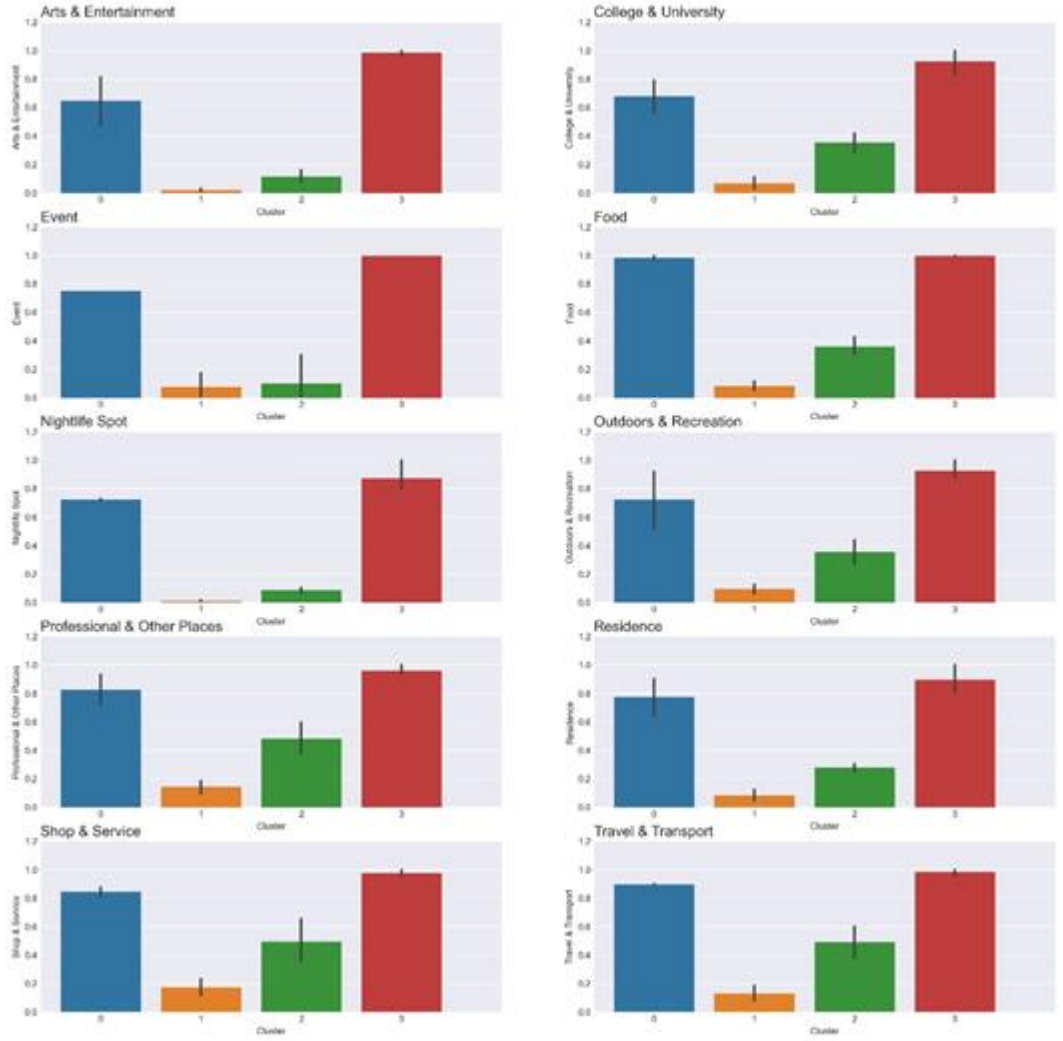
The obtained clusters can be interpreted and visualized on the map as below:

- Cluster 0 (Blue) - is the moderately developed districts of the city.
- Cluster 1 (Green) - has low frequencies for all venue categories. They appear to be underdeveloped neighborhoods of the city.
- Cluster 2 (Red) - has average scores with more professional places and Transport services being the most popular hence falling under the developing neighborhoods bucket. These are mostly residential suburbs.
- Cluster 3 (Pink) - has consistently high frequencies for all venue categories. This is the most diversely developed part of city.

Cluster Markers added to the Budapest Map



CLUSTER WISE FREQUENCIES FOR EACH CATEGORY



DISCUSSIONS

ADDITIONAL KEY FINDINGS

As per our analysis we can see that few neighborhoods of Budapest city can be classified as fully developed, moderately developed, developing and underdeveloped Neighborhoods based on the data from Foursquare API. But with Budapest city as one of the fast-growing tech hubs of Eastern Europe, we can expect the urban footprint of the Budapest city will expand as new growth areas keeps establishing every year.

Districts IV, IX, XI, XIII & XIV emerge as the major growing areas and might even see a transformation from Developing neighborhood to downtown/ uptown as its land usage and the number of tech companies in these districts is increasing in last few years.

FUTURE SCOPE

It would be interesting to further study how we need to consider rental prices and the population data to Predict a growth pattern in these neighborhoods which will also lead us to identify early business and service opportunities in currently underdeveloped areas.

CONCLUSION

Our aim is to aid anyone who visits Budapest city for either Employment, Business or Tourism in deciding the place of stay based on what each neighborhood has to offer, and their personal preferences and our analysis was able to successfully serve the purpose.

THANK YOU!!!