Exploring the surrounding area of schools in Stockholm, Sweden

Agenda

BUSINESS IDEA

DATA

METHODOLOGY

RESULTS

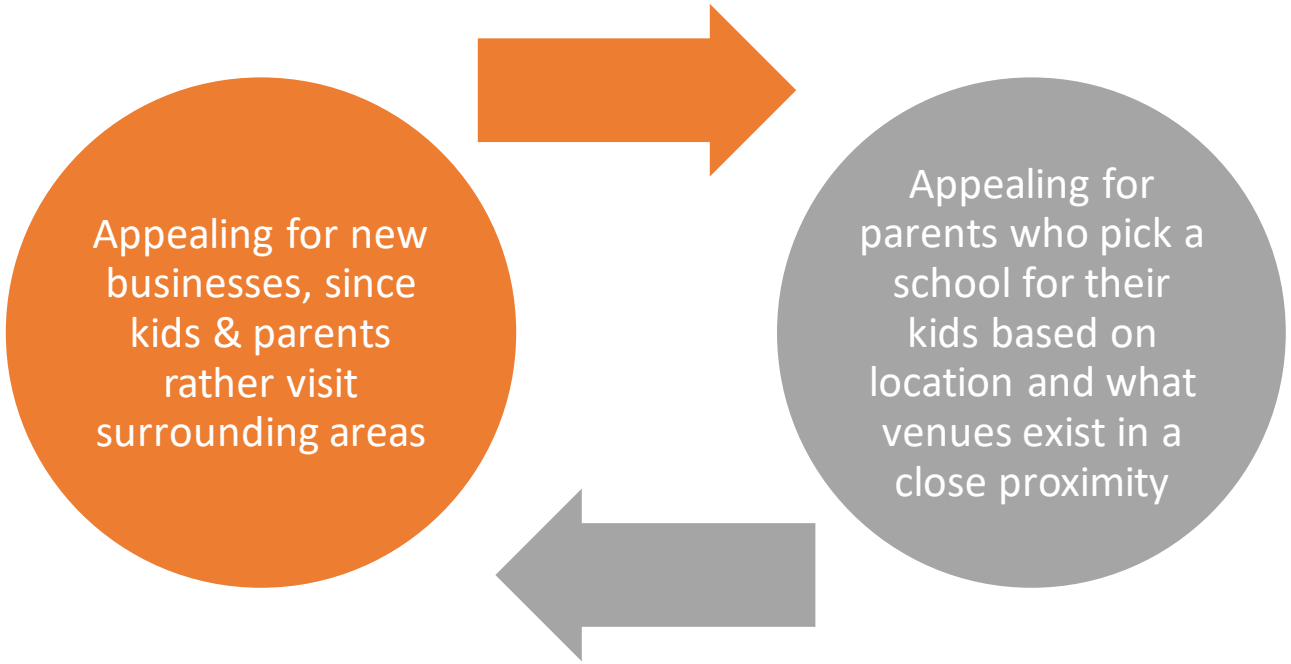
DISCUSSION

TAKE AWAY MESSAGE

Business idea – The goal

- Identify the most common venues next to schools around Stockholm
- Observe if there is any correlation with the number of enrolled students
- Investigate if there is a specific pattern with the most common venues that are somehow related to presence of schools

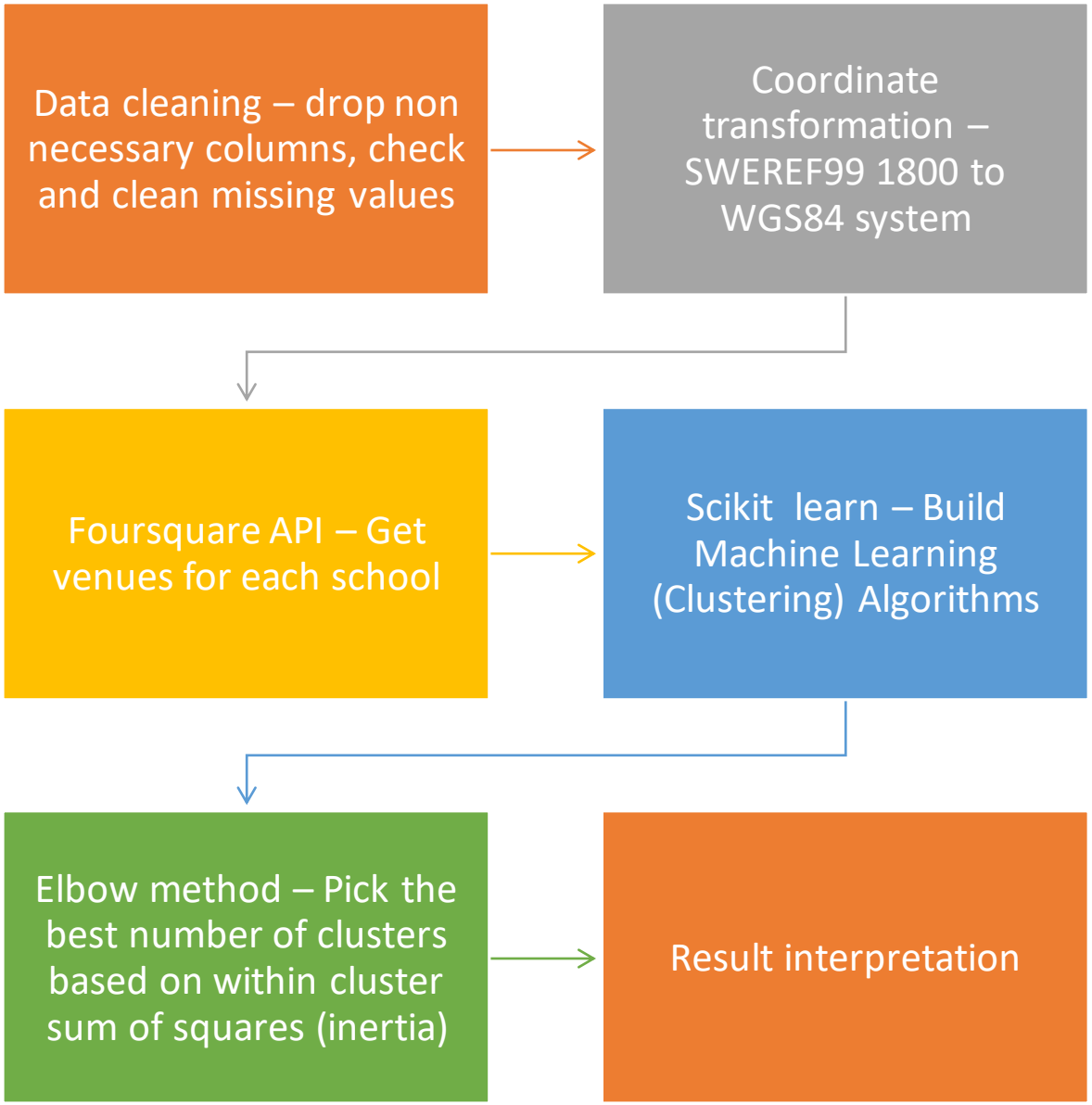
Business idea – Who benefits from this



Data

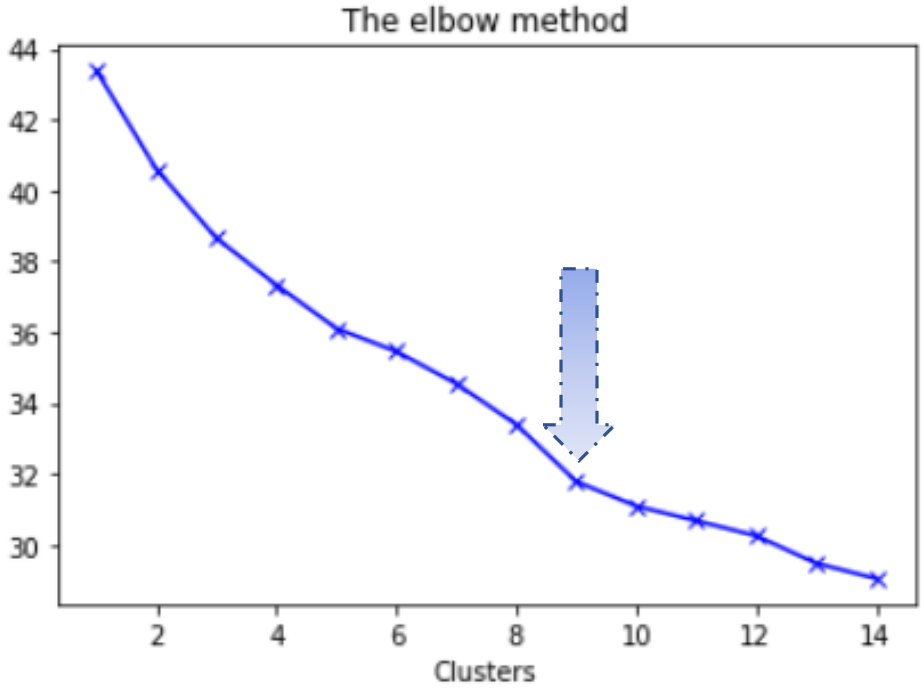
- Stockholm's open source data portal: https://dataportalen.stockholm.se/
- School data for Stockholm region for 2019/20
- 372 schools
- Important features: School name, Region, coordinates, number of enrolled students
- Foursquare API to get venues around Stockholm's schools

Methodology

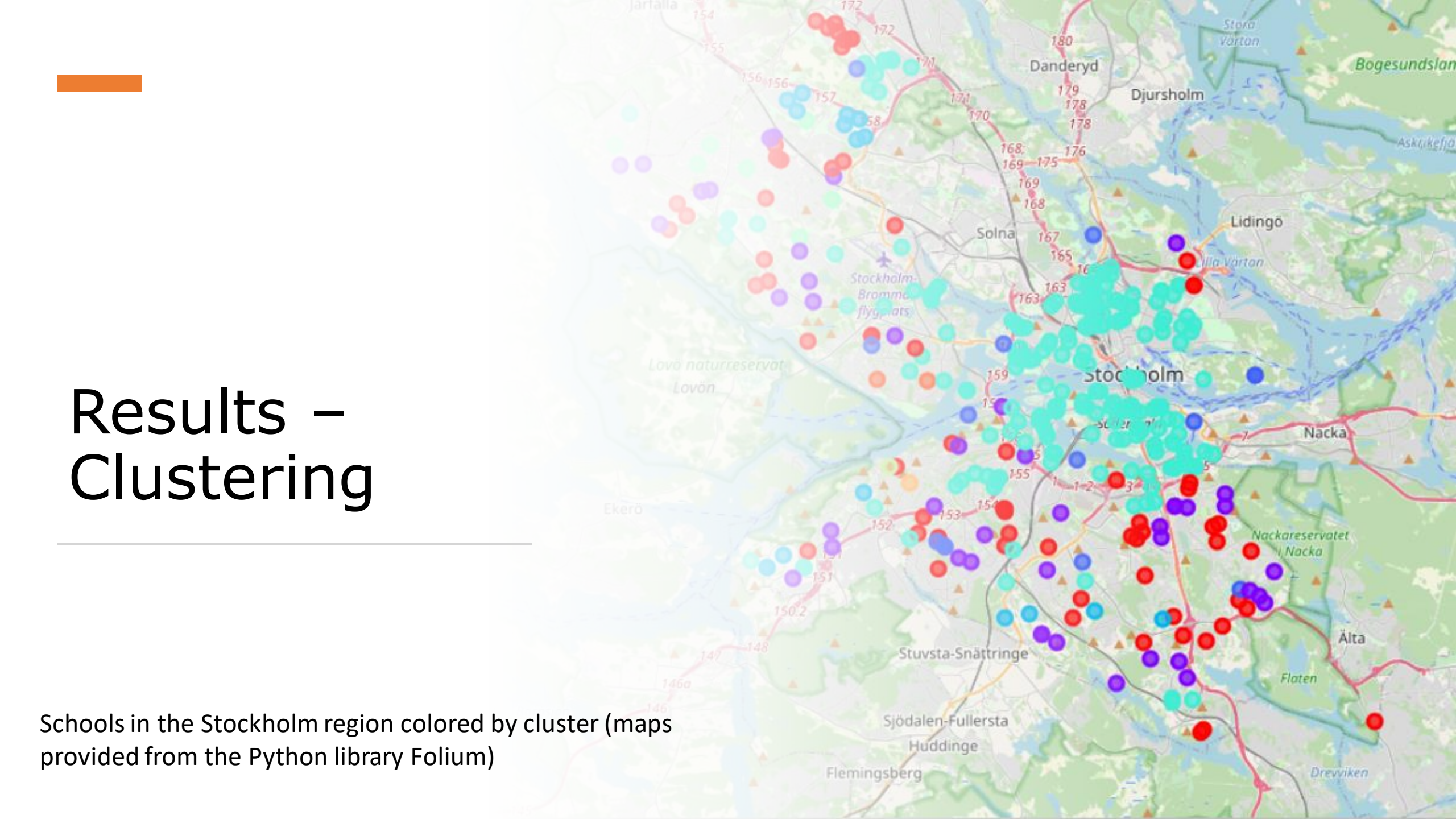


Results – Clustering

The elbow method indicated 9 clusters as the optimal number of clusters for the school dataset



Results – Clustering



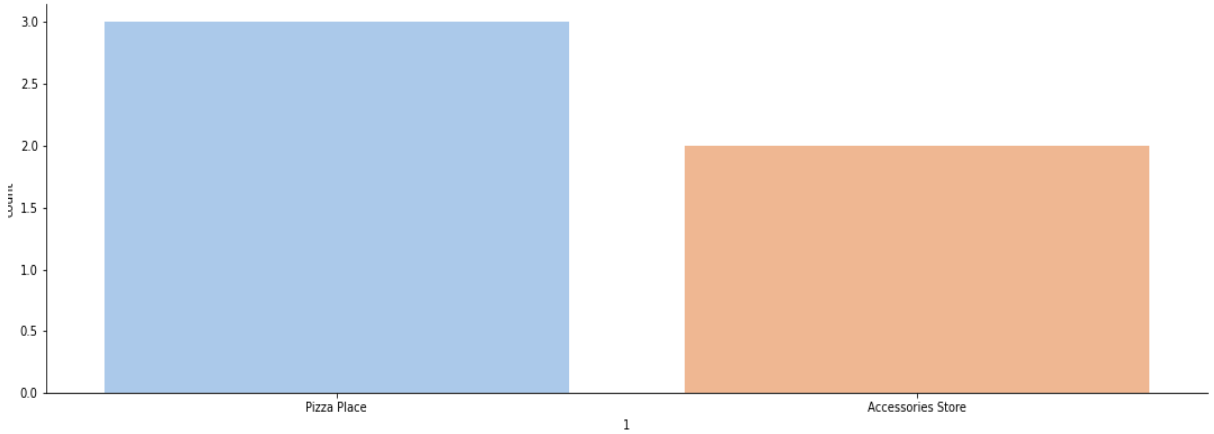
Schools in the Stockholm region colored by cluster (maps provided from the Python library Folium)

Results – Most Popular venues per cluster

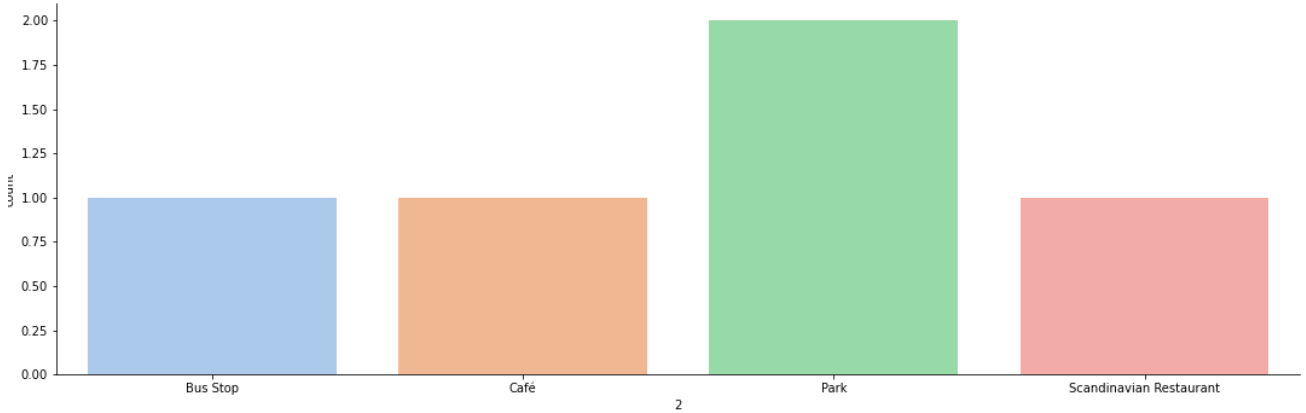
Cluster	MostFreqVen1	MostFreqVen2	MostFreqVen3	MostFreqVen4	MostFreqVen5	MeanNumStudents	ClusterSize
0	Bakery	Grocery Store	Gym / Fitness Center	Park	Accessories Store	362.582090	67
1	Pizza Place	Pizza Place	Pizza Place	Accessories Store	Accessories Store	429.727273	44
2	Bus Stop	Café	Park	Park	Scandinavian Restaurant	238.733333	15
3	Bus Stop	Convenience Store	Grocery Store	Metro Station	Accessories Store	318.312500	16
4	Café	Café	Bakery	Bakery	Gym / Fitness Center	328.556561	221
5	Bus Stop	Bus Stop	Accessories Store	Accessories Store	Advertising Agency	497.714286	7
6	Harbor / Marina	Accessories Store	Advertising Agency	American Restaurant	Amphitheater	153.000000	1
7	Deli / Bodega	Accessories Store	Advertising Agency	American Restaurant	Amphitheater	1204.000000	1
8	Tram Station	Accessories Store	Advertising Agency	Advertising Agency	Amphitheater	398.500000	2

Results – Most Popular venues- examples

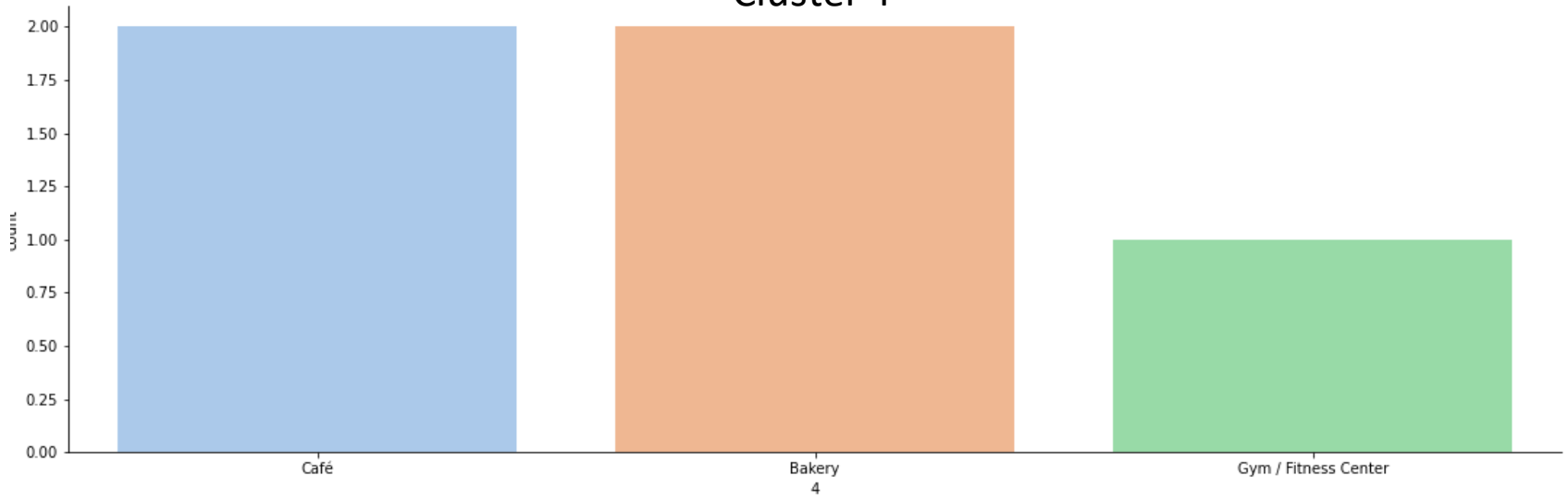
Cluster 1



Cluster 2



Cluster 4



Results – Correlation matrix

As the results show, the mean number of students for each cluster is not particularly affected by the type of most frequent venues.

```
ven.corr()
```

	MostFreqVen1	MostFreqVen2	MostFreqVen3	MostFreqVen4	MostFreqVen5	MeanNumStudents	ClusterSize
MostFreqVen1	1.000000	-0.278375	-0.110390	-0.659893	0.051873	0.085029	-0.201720
MostFreqVen2	-0.278375	1.000000	0.842451	0.226066	-0.547965	-0.276911	0.296557
MostFreqVen3	-0.110390	0.842451	1.000000	0.366223	-0.160335	-0.288443	0.123720
MostFreqVen4	-0.659893	0.226066	0.366223	1.000000	0.215778	-0.259541	0.207402
MostFreqVen5	0.051873	-0.547965	-0.160335	0.215778	1.000000	-0.047688	0.193308
MeanNumStudents	0.085029	-0.276911	-0.288443	-0.259541	-0.047688	1.000000	-0.191926
ClusterSize	-0.201720	0.296557	0.123720	0.207402	0.193308	-0.191926	1.000000

Discussion

Stakeholders

- Build cafeterias or a gym next to a school in Cluster 1, since Pizza place and Accessories stores are the most common venues
- Build a restaurant in Cluster 4, since it is not included into the list of the common venues
- **As a stakeholder you aim to pick an area where the venue you want to build is not highly present**

Parents

- Pick a school from Cluster 2, since it seems probable to have available a park to play, cafes and restaurants as well as a bus stop.
- **As a parent you will probably pick a school that is surrounded by many venues where you can eat, relax, play or train for you and/or your kids. At the same time having a bus stop, metro station etc. can help to get home easier if you live far away from the city center.**

Take away message

Picking a location to expand business can be tricky, but open source data along with APIs, such as Foursquare can make the job easier

As a parent, it is important to pick a school after careful research, thus this kind of project can help to decide fast without having to visit all the candidate schools

Similar projects can be conducted in other fields such as: building a new park, a new block of flats, or even for conserving nature based on what exists in the surrounding area and how we can preserve an ecosystem in the best possible manner

Thank you for your time