WHO NEEDS CUSTOMERS ANYWAY?

Finding the best place to build a cinema:

Coursera IBM Data Science Capstone Project

WHAT & WHY?

- There are many cinemas in the Netherlands.

- There is a hectic business landscape

- If a newly-built cinema fails, it is very costly

- *The question is: which municipalities would be the best to build your new cinema?*

DATA

Three datasets are used to investigate the research question:

- The current number of cinemas per municipality → Foursquare

- The longitude and latitude of each municipality → geopy

- Information for all the municipalities in the Netherlands, such as number of inhabitants and land area → Wikipedia

METHODOLOGY

Different methods were used to investigate the data:

- Spatial explorative analysis

- Correlation explorative analysis using linear regression

- Machine learning: K-means clustering

- Determination of best municipality using histograms

SPATIAL EXPLORATIVE ANALYSIS

Inhabitants per square km
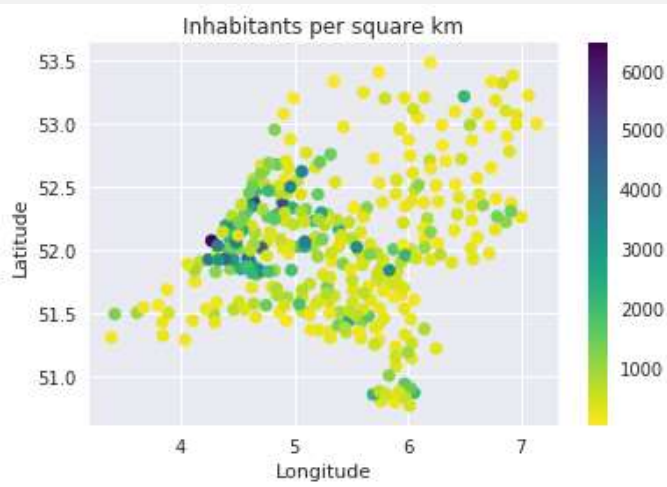
The west and south on average have a higher:

- number of inhabitants per square km

- average household income

- number of cinemas per inhabitant and per square km

The number of potential customers for a new cinemas is more evenly spread.

Number of cinemas per inhabitant

Average household income

Number of potential customers for a new cinema

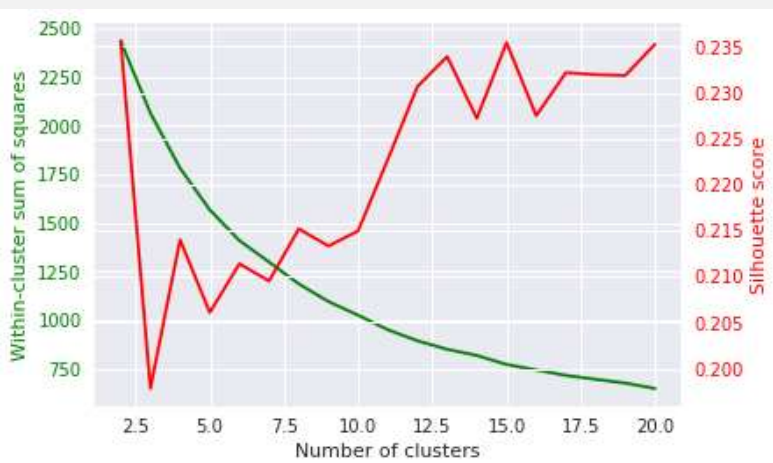Number of cinemas per square km
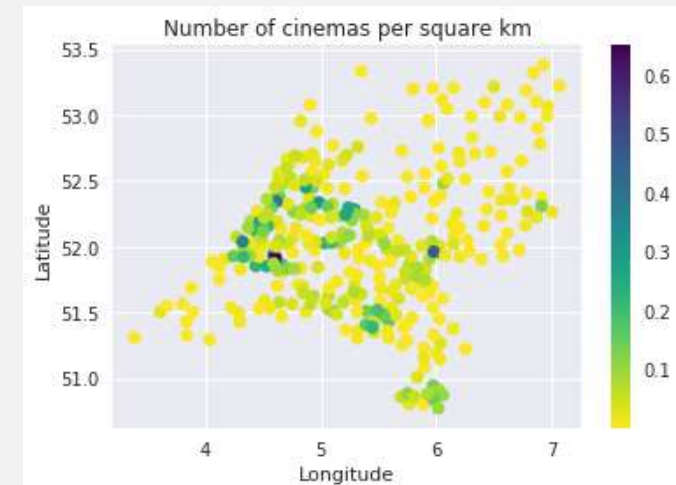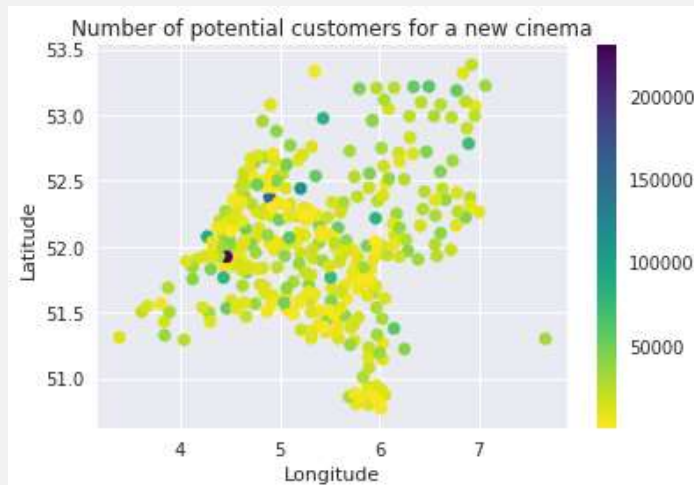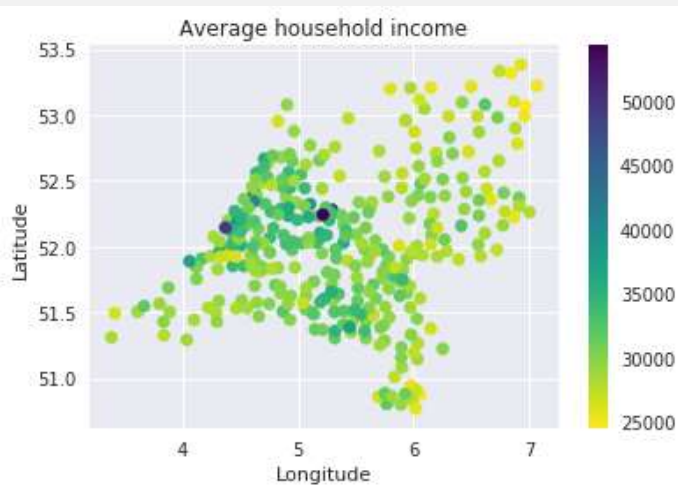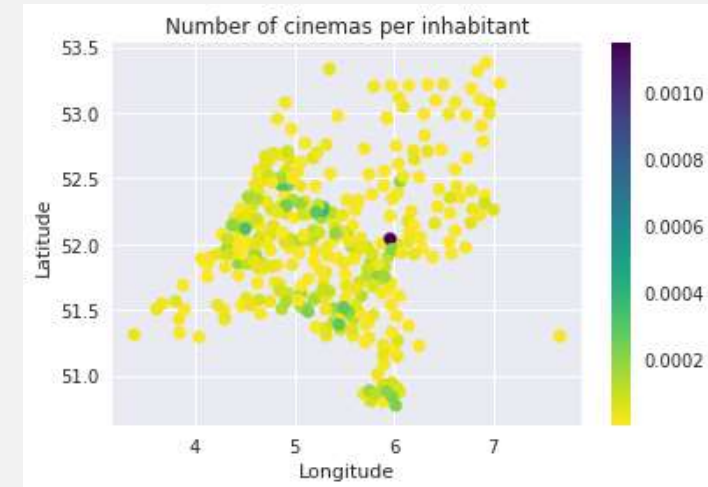
CORRELATION EXPLORATIVE ANALYSIS USING LINEAR REGRESSION

There is very poor correlation in a municipality between:

- The number of inhabitants (per square km)
- The number of cinemas (per inhabitant or per square km)
- The average household income
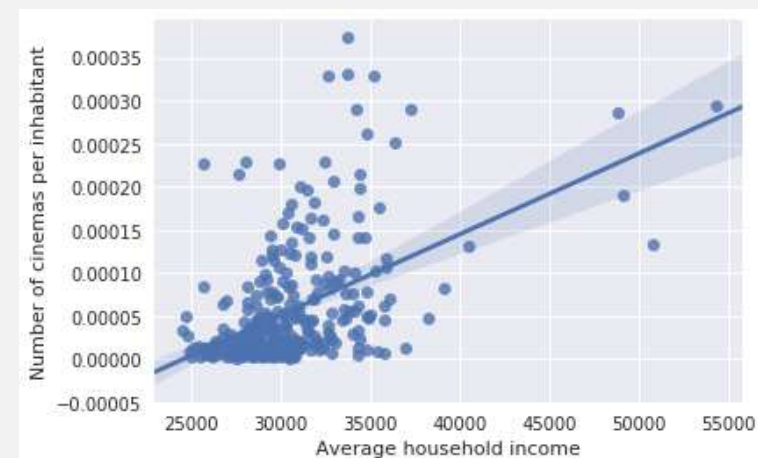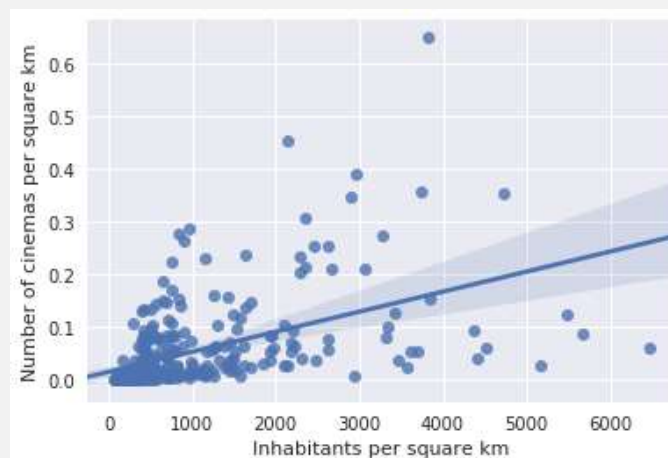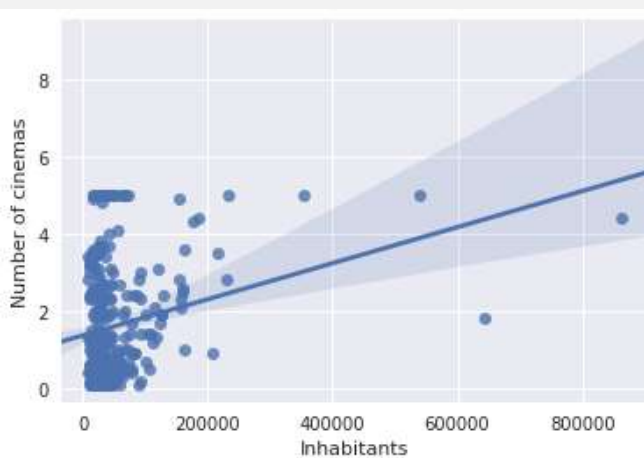
MACHINE LEARNING: K-MEANS CLUSTERING

The silhouette method gave 8 as the optimal number of clusters.

The following clusters were found:

Label	Characteristics
0	North, few cinemas, lower average household income
1	Randstad, some cinemas, lower average household income
2	East, few cinemas, lower average household income
3	Scattered, many cinemas, high average household income
4	Randstad, few cinemas, lower average household income
5	Big 4 cities, many cinemas, lower average household income
6	Randstad, many cinemas, above average average household income
7	Center/south/east, many cinemas, lower average household income
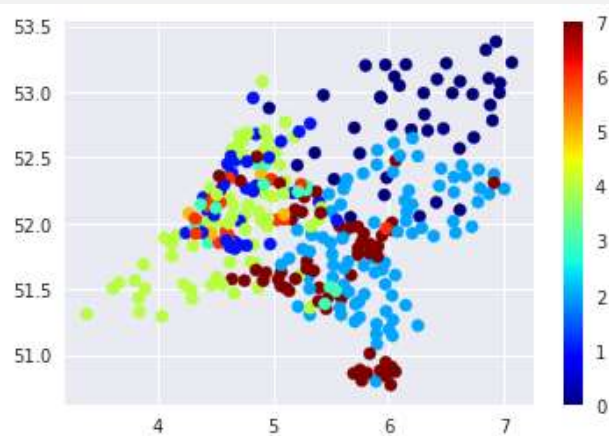
DETERMINATION OF METRIC FOR BEST MUNICIPALITY
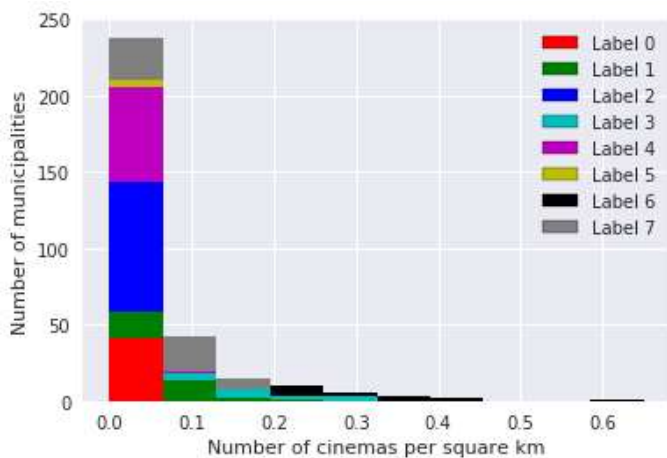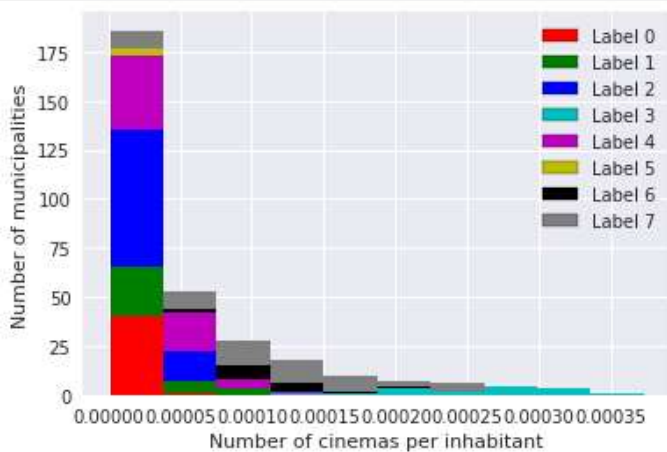
Four parameters were considered as metric:

1. Number of cinemas per inhabitant

2. Number of cinemas per square km

3. Number of cinemas per inhabitant times square km

4. Number of potential customers for a new cinema

The number of potential customers for a new cinema was chosen.
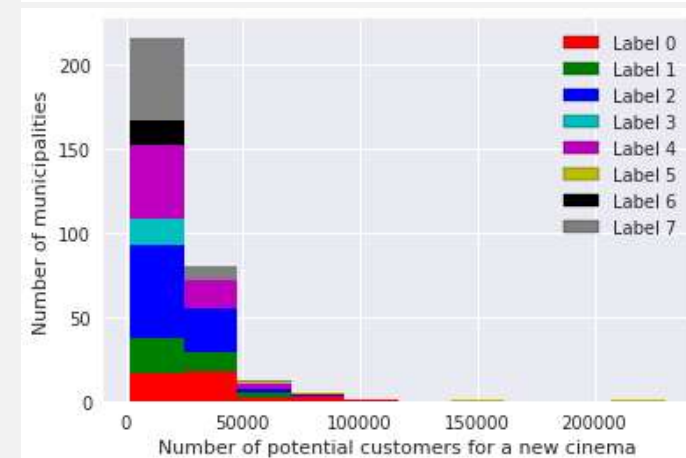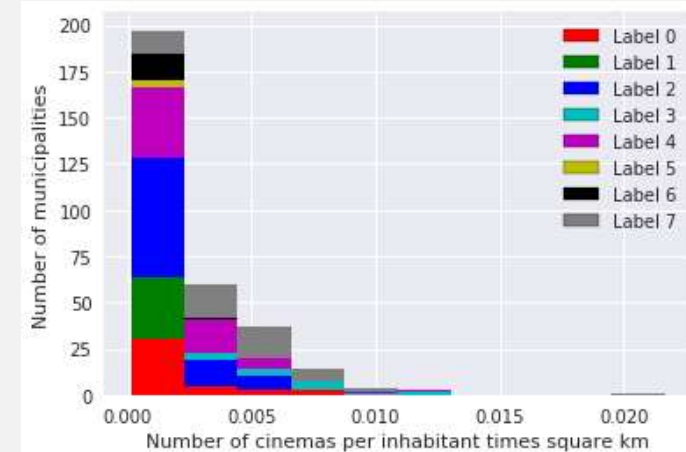
RESULTS

According to this analysis, the top 10 of best municipalities to build a new cinema are:

1. Rotterdam
2. Amsterdam
3. Almere
4. Den Haag
5. Súdwest-Fryslân
6. Apeldoorn
7. Oss
8. Emmen
9. Groningen
10. Venlo

DISCUSSION

- The top 10 consists of large municipalities or medium-sized municipalities with relatively few cinemas

- Only three labels feature in the top 10

- Nine different provinces feature in the top 10

- Average household income is lower in the top 10 than the Dutch average

- For a next study, I recommend:

  - taking consumer behavior into account

  - improve the Foursquare data

  - take data on a finer level than on municipal level

CONCLUSION

- The top 10 of best municipalities to build a new cinema was found

- The number of potential customers for a new cinema was used

- K-means clustering was performed with eight clusters, of which only three clusters are featured in the top 10.

EXTRA

Municipality	Number of potential customers for a new cinema	Inhabitants	Land area	Average household income	Number of cinemas	Label	Province
Rotterdam	230,188	644,527	218	26,200	1.8	5	Zuid-Holland
Amsterdam	159,852	863,202	166	29,700	4.4	5	Noord-Holland
Almere	109,378	207,819	129	28,800	0.9	0	Flevoland
Den Haag	89,664	537,988	82	28,300	5	5	Zuid-Holland
Súdwest-Fryslân	81,550	89,705	522	27,500	0.1	0	Friesland
Apeldoorn	81,228	162,456	340	29,200	1	0	Gelderland
Oss	76,197	91,437	163	29,100	0.2	2	Noord-Brabant
Emmen	71,407	107,111	335	26,200	0.5	0	Drenthe
Groningen (gemeente)	60,882	231,354	95	25,700	2.8	0	Groningen (provincie)
Venlo	59,751	101,578	124	26,800	0.7	2	Limburg