Business Problem:

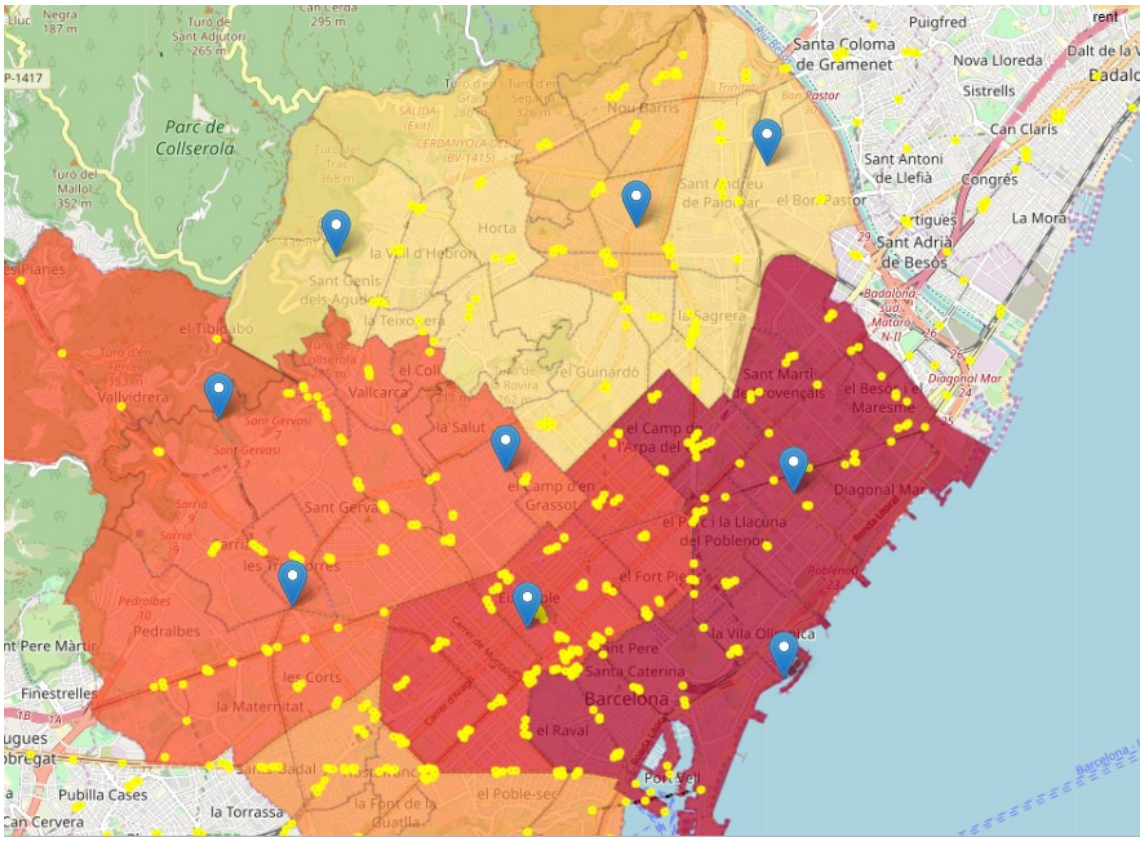
- How to locate an optimal place to rent a house when you move to Barcelona.

- We have taken in to consideration population as well the availability of public transportation infrastructure in each neighbourhood.

DATA

- Rent and buy average price on December 2018 from the biggest real state portal.www.idealista.com
- Public transportation data from OPEN DATA BARCELONA
- JSON data from OPEN DATA BARCELONA
- Academic level of the population of each neighbourhood. OPEN DATA BARCELONA. Note: The results have been the same with/without this data therefore in the last version, this data has been not used.

DAFRAME creation



- Here is the data that we have used. Neighbourhood using blue icons
- Choropleth has been used for the rental price
- and Yellow markers identifies the public transportations.

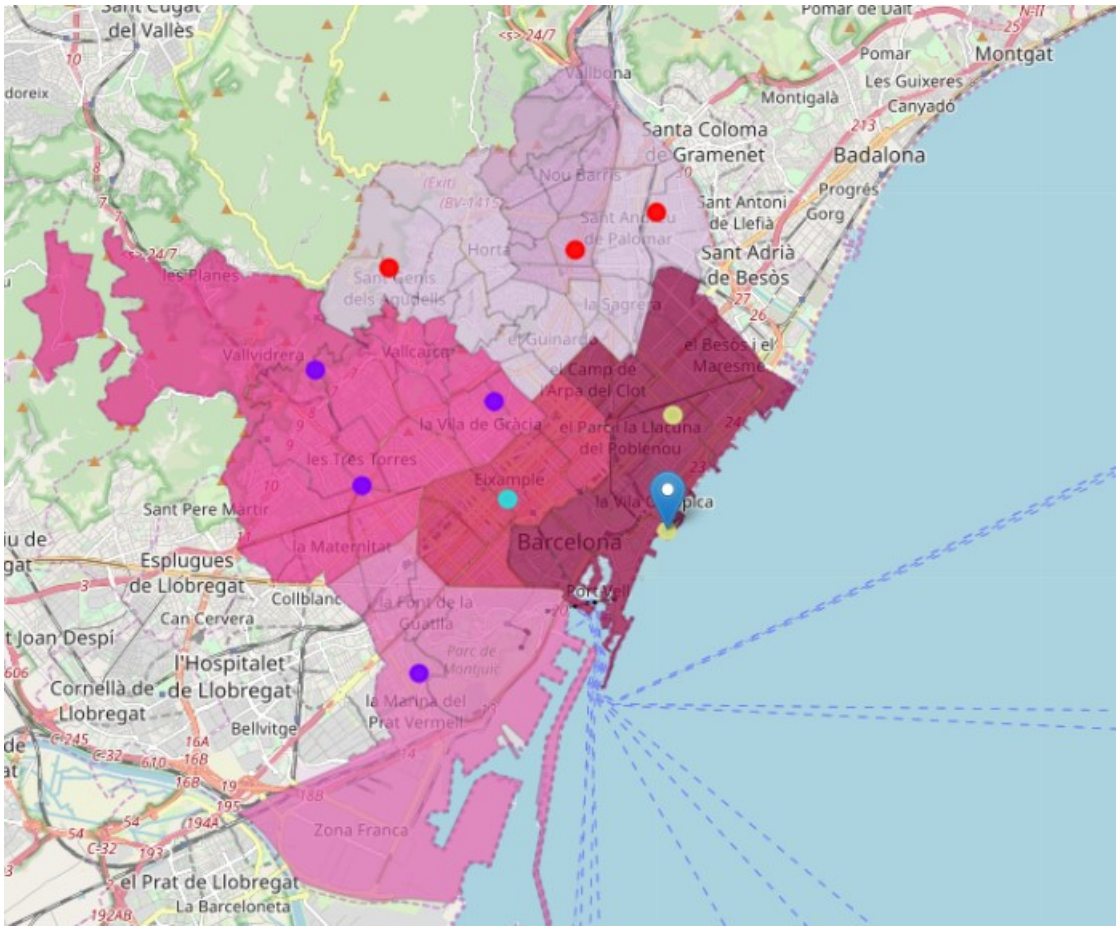
Methodology

The user has to introduce the desired neighbourhood and the program will automatically calculate the linear distance between that neighbourhood and the rest and update the information into the dataframe.

For the recommendation system we have used a 4-K clustering system.

RESULTS

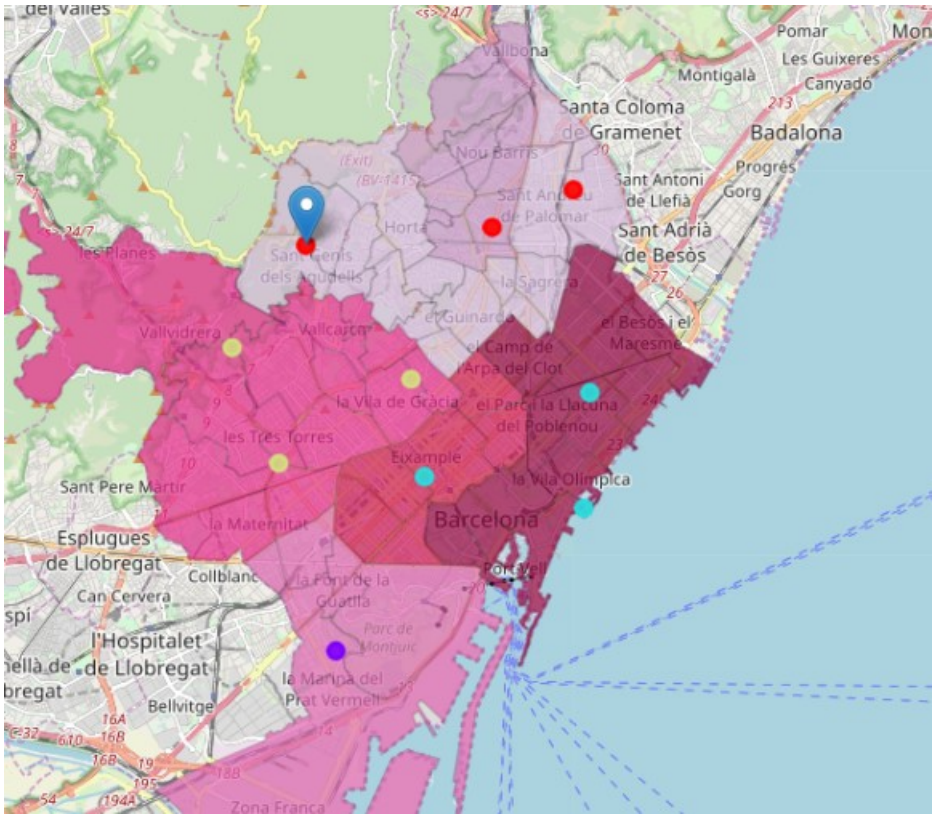
0 Ciudad Vella



Red:0, Purple:1 LightBlue:2 DarkYellow: 3

KMeans	METRO	price_rent	price_buy	CommuteKM
0	36.666667	13.466667	2836.333333	6.331599
1	40.000000	16.150000	4553.500000	5.705630
2	120.000000	17.900000	5005.000000	3.090517
3	48.500000	18.700000	4324.000000	1.098717

3 HORTA



Red:0, Purple:1 LightBlue:2 DarkYellow: 3

KMeans	METRO	price_rent	price_buy	CommuteKM
0	36.666667	13.466667	2836.333333	2.923415
1	53.000000	15.800000	3621.000000	7.765799
2	72.333333	18.433333	4551.000000	6.120077
3	35.666667	16.266667	4864.333333	3.276328

CONCLUSION

As we can see the results vary from each location and as well the mean price for each cluster.

Moreover, data visualization is a very powerful tool in order to show results.

AS a future project we can expand the scope of the project and perform the analysis by quarter. Data is available but not as recent as the one used in this study.