



THE FACTORS INFLUENCING THE LIVING STANDARDS IN BELGIUM

Data Management project by:
Massimo Buranel
Brian Dastroy
Patrick Zeni

PART ONE: DATA RETRIEVAL

We collected data from the statbel website, in particular we collected data by province from 2010 to 2022 on population, unemployment and average house price and national data on inflation, consumer price index and health index. We processed the data until we obtained this final dataframe

	Year	Province	Population	Unemployment	House_Price	\
0	2010	Antwerp	18775	0.061024	215000	
1	2011	Antwerp	20096	0.057756	225000	
2	2012	Antwerp	19097	0.052988	230000	
3	2013	Antwerp	19082	0.062531	237250	
4	2014	Antwerp	23058	0.061422	237500	
..
138	2018	Brussels	11174	0.133527	390000	
139	2019	Brussels	11706	0.127234	412000	
140	2020	Brussels	10501	0.124268	450000	
141	2021	Brussels	12485	0.125043	470500	
142	2022	Brussels	12559	0.115249	500000	
		Consumer price index	Inflation	Health index		
0		92.88	0.021782	93.37		
1		96.17	0.035422	96.22		
2		98.90	0.028387	98.77		
3		100.00	0.011122	100.00		
4		100.34	0.003400	100.40		
..	
138		107.24	0.020556	107.35		
139		108.78	0.014360	108.92		
140		109.59	0.007446	110.00		
141		112.26	0.024364	112.21		
142		123.03	0.095938	122.59		



LIVING STANDARD CALCULATION

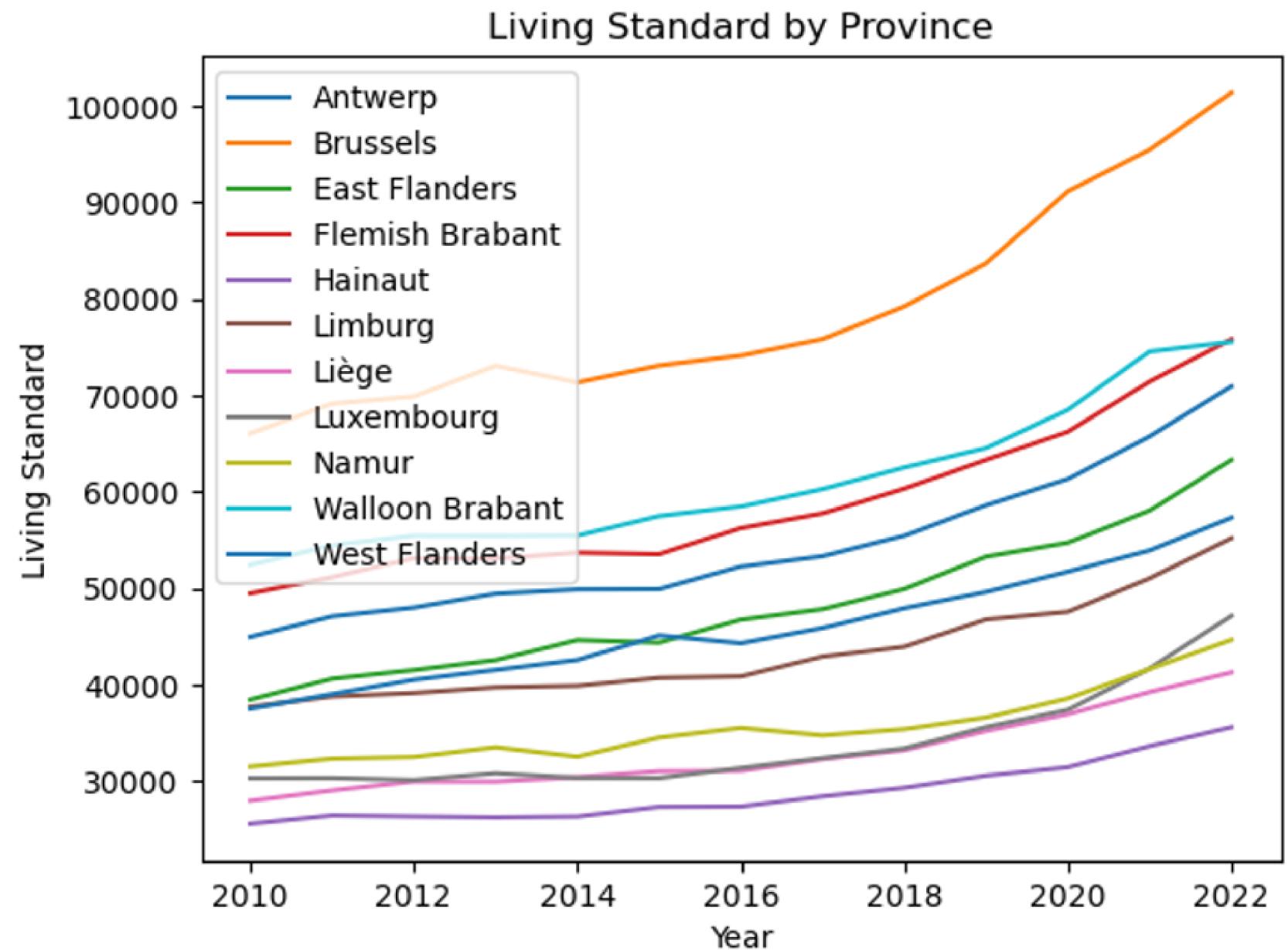
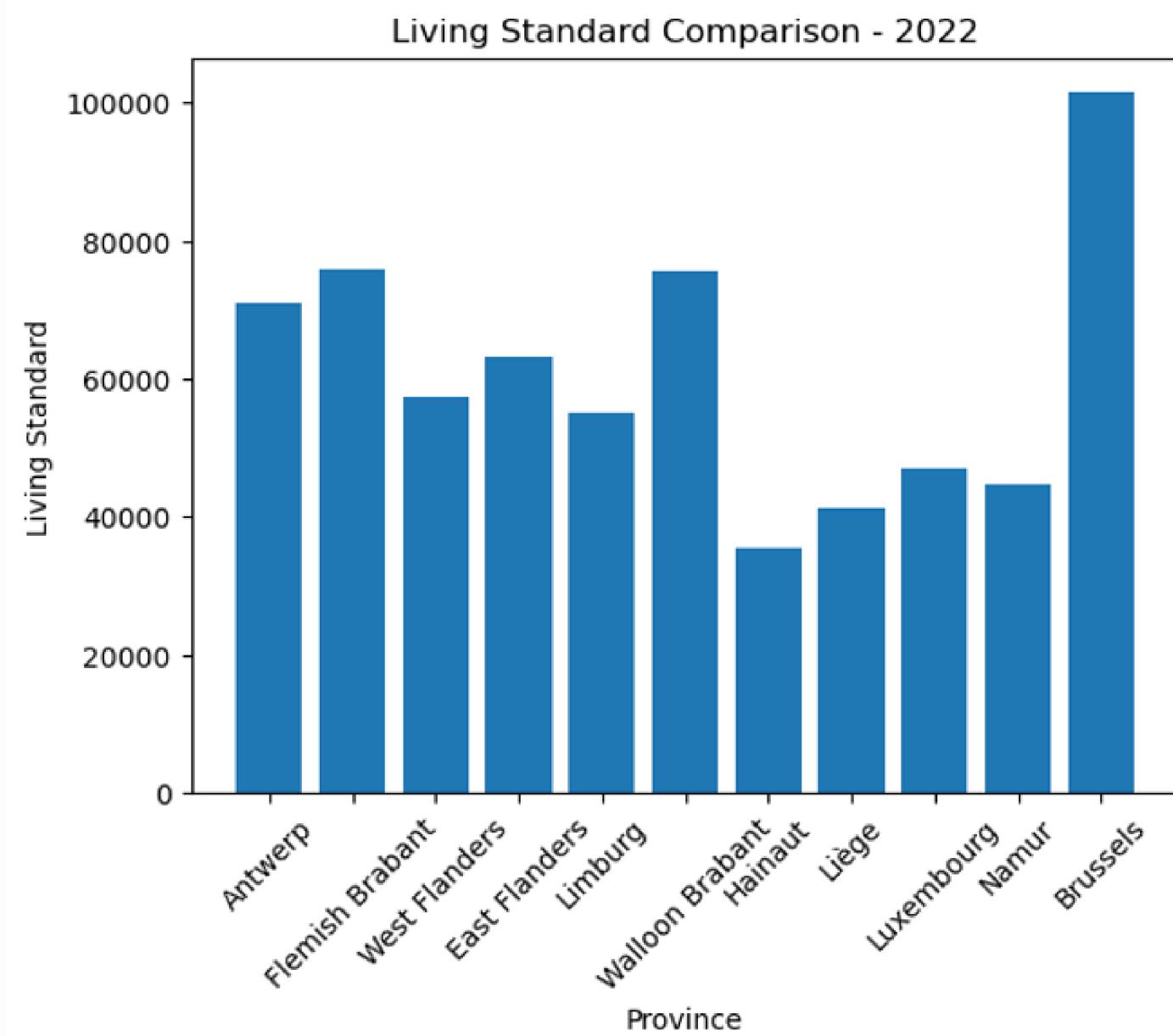
We then determined the weight of each of these variables in order to calculate the standard Living Index per province for all years from 2010 to 2022.

The weight was determined subjectively, in particular, the unemployment index was given a negative weight so as to lower the living standard as unemployment increased.

	Year	Province	Living Standard
0	2010	Antwerp	44942.753152
1	2011	Antwerp	47076.939533
2	2012	Antwerp	47978.860080
3	2013	Antwerp	49428.189718
4	2014	Antwerp	49876.056396
..
138	2018	Brussels	79192.500406
139	2019	Brussels	83646.793425
140	2020	Brussels	91126.994636
141	2021	Brussels	95427.036864
142	2022	Brussels	101341.797138

PART TWO: DATA VISUALIZATION

We used different techniques to visualise the data and to make it easier to understand

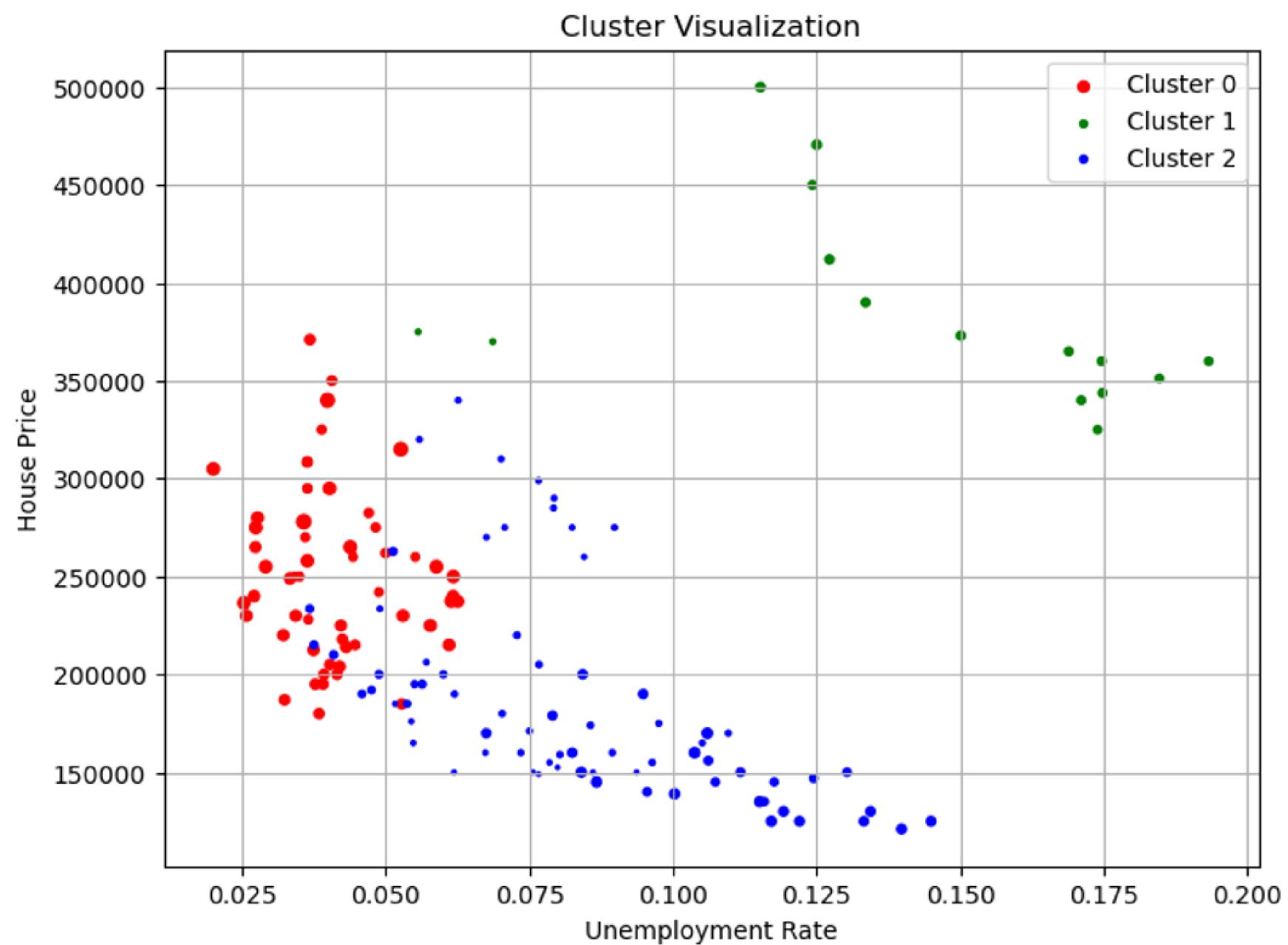


CLUSTER ANALYSIS

We have performed a cluster analysis by grouping in 3 clusters.

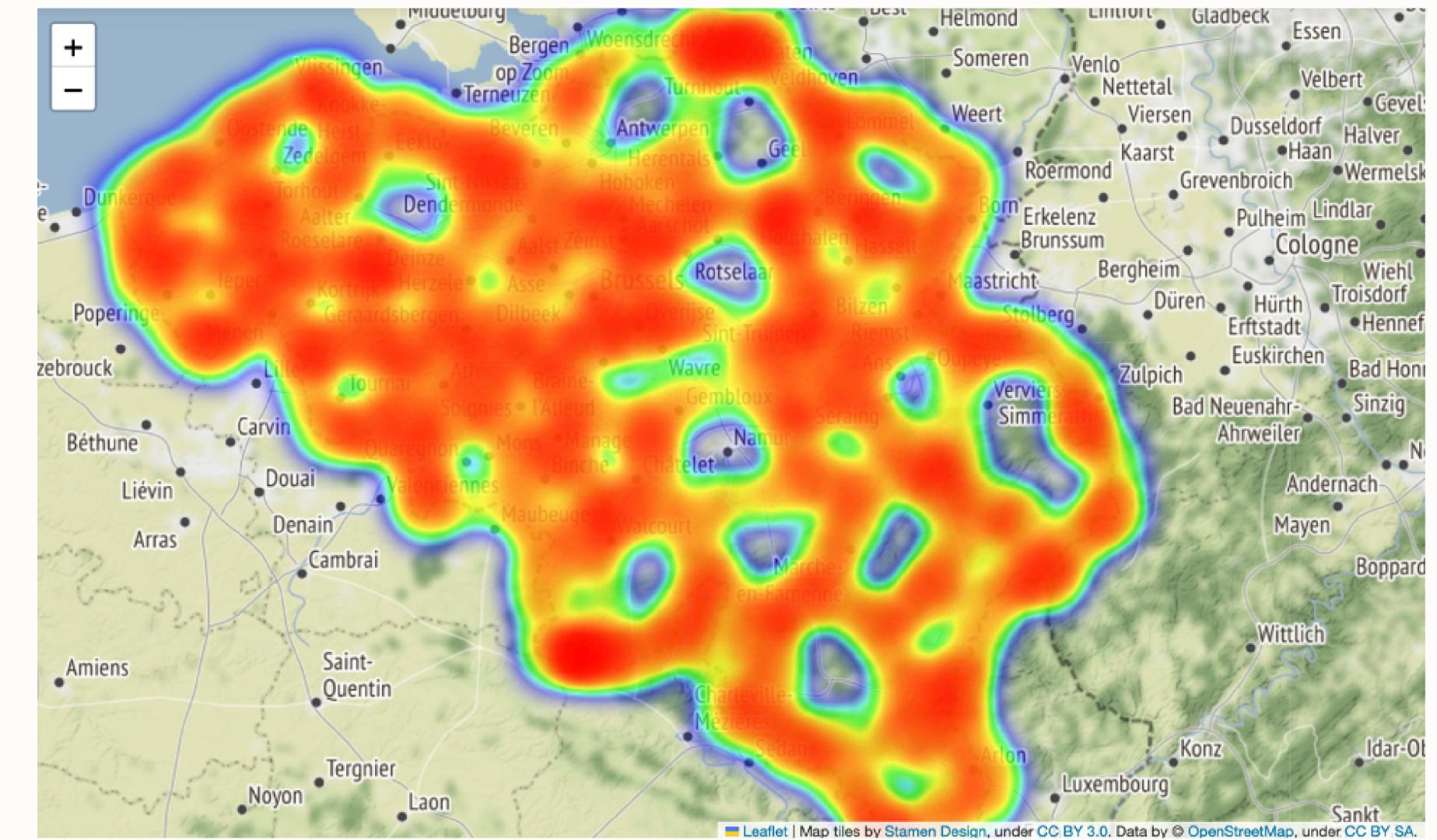
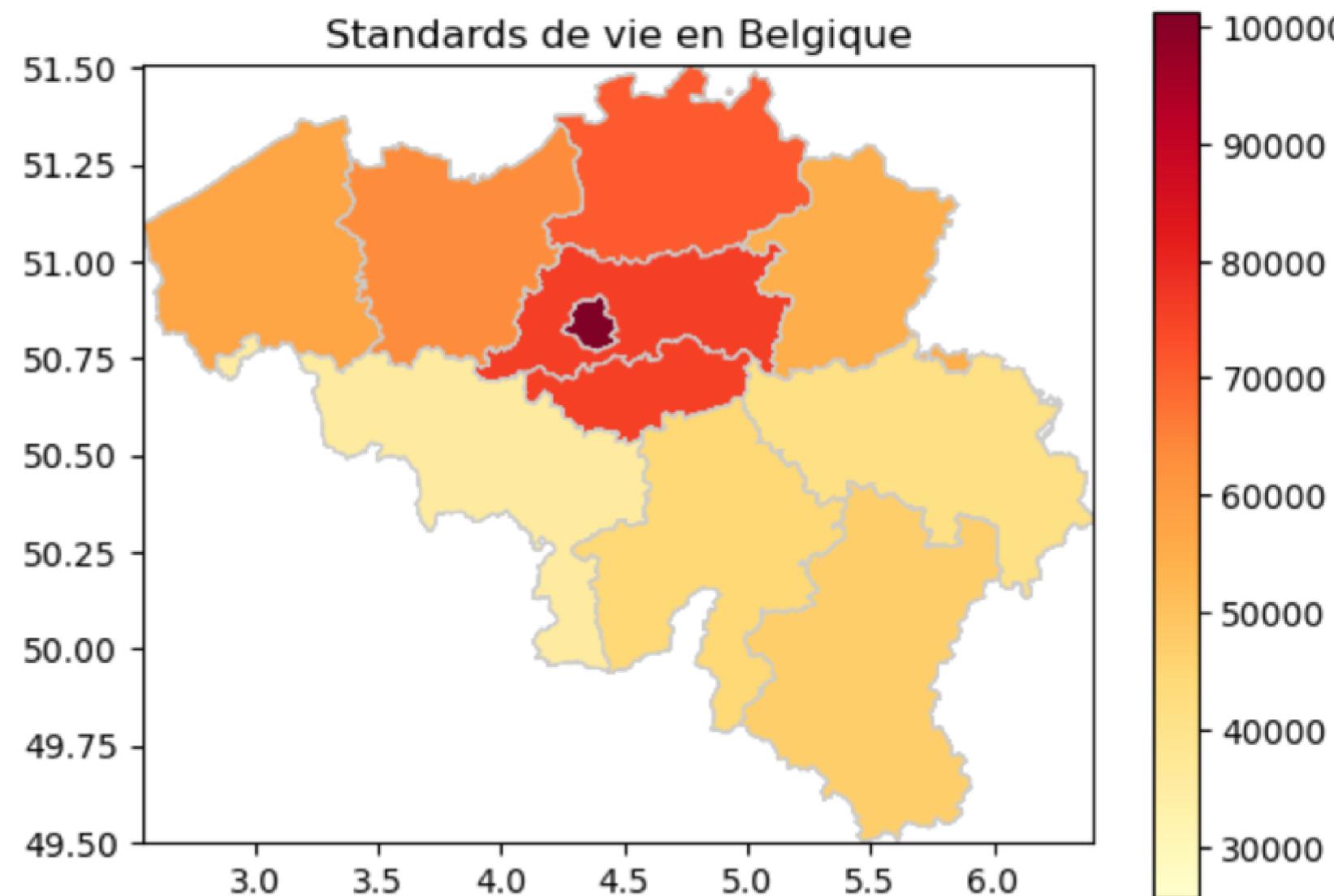
We standardised the data using the StandardScaler class to ensure that all variables had the same scale. Next, we applied the K-Means clustering algorithm with 3 clusters using the KMeans class.

We then represented the clusters graphically, in particular the population variable is represented by the size of the dot



MAP VISUALISATION

We also used map views to show the level of standard living by province.



PART THREE: DATA MODELLING

Finally, we built a model to predict the value of the future Living Standard.

We calculate the trend for the Living Standard so that we can predict the value of the Living Standard for each province for the year 2023.

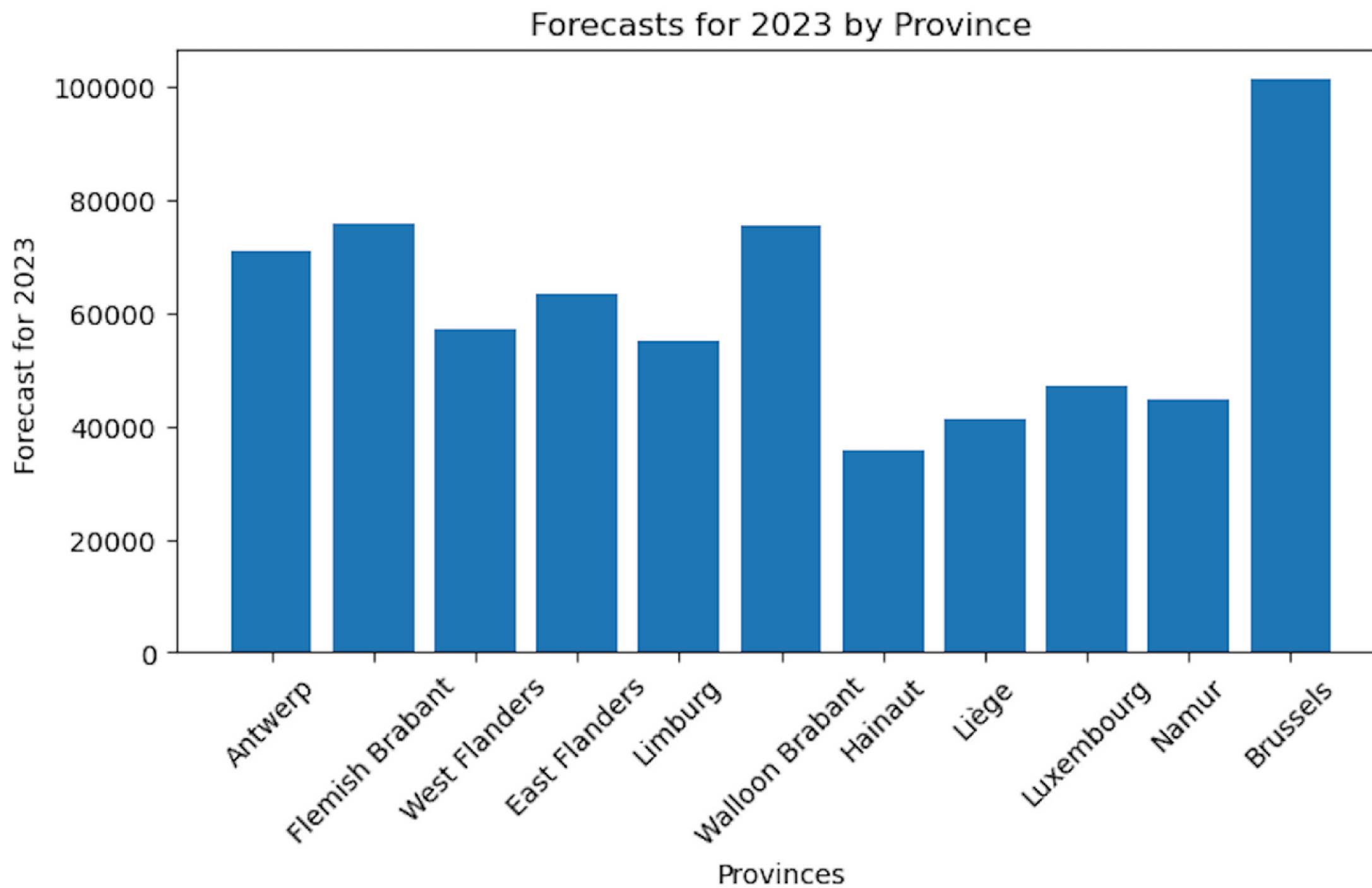
We are averaging the historical values of the target variable 'Living Standard' for each province and determining a trend based on the average difference of the historical values from the average. We are therefore generating future values of the target variable using this trend.

	Province	Population	Unemployment	House_Price \
0	Antwerp	28280	0.039851	340000
1	Flemish Brabant	14999	0.036815	371000
2	West Flanders	22080	0.027391	275000
3	East Flanders	21987	0.019973	305000
4	Limburg	10787	0.035970	270000
5	Walloon Brabant	4317	0.055652	375000
6	Hainaut	14949	0.105991	170000
7	Liège	11772	0.084294	200000
8	Luxembourg	3349	0.048982	233500
9	Namur	5546	0.072859	220000
10	Brussels	12559	0.115249	500000

	Consumer price index	Inflation	Health index	Living Standard
0	123.03	0.095938	122.59	54361.523909
1	123.03	0.095938	122.59	58839.248856
2	123.03	0.095938	122.59	45905.050664
3	123.03	0.095938	122.59	48140.488870
4	123.03	0.095938	122.59	43389.502388
5	123.03	0.095938	122.59	61143.720021
6	123.03	0.095938	122.59	28810.988273
7	123.03	0.095938	122.59	32886.314263
8	123.03	0.095938	122.59	33928.467044
9	123.03	0.095938	122.59	35686.302308
10	123.03	0.095938	122.59	78698.988117

LIVING STANDARD PREDICTION

Then using a linear regression model, we calculated the expected value of the standard of living for each province for 2023





THANK YOU