Elderly People Treated in Integrated Home Care in Italian Regions: A Metric Approach

Resta, Onofrio and Resta, Emanuela and Costantiello, Alberto and Leogrande, Angelo

University of Bari Aldo Moro, University of Foggia, Lum University Giuseppe Degennaro, Lum Enteprise S.r.l.

30 December 2023

Onofrio Resta[1]°, Emanuela Resta[2]^, Alberto Costantiello[3], Angelo Leogrande[4]*,*

°University of Bari "*Aldo Moro*", Bari, Puglia, Italy, EU
^ University of Foggia, Foggia, Puglia, Italy, EU
*LUM University Giuseppe Degennaro, Casamassima, Bari, Puglia, Italy, EU

Elderly People Treated in Integrated Home Care in Italian Regions: A Metric Approach

Abstract

In this article, we analyse the ESG determinants of the "*Elderly People Treated in Integrated Home Care*"-EPIHC in the Italian regions between 2004 and 2022. We used data from the ISTAT-BES database. We used different econometric techniques i.e.: Panel Data with Random Effects, Panel Data with Fixed Effects, Pooled Ordinary Least Squares-OLS and Weighted Least Squares-WLS. The results show that the EPIHC is positively associated with "*Nurses, midwives, and Soil sealing by artificial cover*" and negatively associated with "*Museum heritage density and relevance*" and "*Trust in law enforcement agencies and firefighters fire*". Furthermore, we have applied a k-Means algorithm with the Silhouette Coefficient and we find the presence of two clusters. Finally, we propose a confrontation among eight different machine-learning algorithms and we find that Linear Regression is the best predictive algorithm.

Keywords: *Analysis of Health Care Markets, Health Behaviors, Health Insurance, Public and Private, Health and Inequality, Health and Economic Development, Government Policy • Regulation • Public Health.*

JEL CODE: *I11, I12, I13, I14, I15, I18*

1) Introduction- Research Question

In the following article, we analyse the value of "*Elderly People Treated in Integrated Home Care*"-EPIHC in the Italian regions between 2004 and 2022. The EPIHC variable is analysed in light of the ISTAT-BES dataset, which captures elements of the Environment, Social and Governance-ESG type in the Italian regions. The problems connected to the management of the health economy pose on the one hand the need to offer increasingly adequate care to the elderly population, and on the other hand, they underline the needs of reducing the costs of health services by increasing the benefits for patients also using new digital technologies. In particular, the aging of the Italian population creates risk conditions also from an economic-financial point of view for the public and private organizations that exercise healthcare management. It is therefore necessary to reorganize the healthcare system by considering the new technological opportunities offered by telemedicine for homecare and

[1] Professor at University of Bari Aldo Moro, email: onofrioresta2150@gmail.com
[2] Assistant Professor at University of Foggia, email: restaemanuela@gmail.com
[3]Professor of Economics at LUM University Giuseppe Degennaro, Strada Statale 100 km 18, Casamassima, Bari, Puglia, Italia. Email: costantiello@lum.it
[4]Researcher at LUM Enterprise s.r.l. Email: leogrande.cultore@lum.it, Strada Statale 100 km 18, Casamassima, Bari, Puglia, Italia.

highlighting the advantages in economic-financial, social and environmental terms connected to home care. The article continues as follows: the second section contains a brief analysis of the literature, the third section shows the results of the econometric model, the fourth section discusses the cluster analysis with k-Means algorithm optimized with the Silhouette Coefficient, the fifth section contains a comparison between machine algorithms learning for prediction of the future values of EPIHC, the sixth paragraph concludes. The appendix contains further materials, data, and graphs.

2) Literature Review

Below we consider a brief analysis of the literature analysing the costs and benefits of offering home care services for the elderly, under a financial and economic point of view.

Home care, robotics, and telehealth. Robotics can be used to improve the effectiveness of home care for the elderly [1]. The probability to increase efficacy in using telemedicine in the treatment of depression in elderly patients treated at home is positively associated to the knowledge of healthcare professionals [2]. It is relevant to analyse either direct either indirect costs in the economic evaluation of the application of eHealth of elderly people treated at home [3].

Patient care at the home is a necessity of almost all healthcare systems at country level. The home care of patients also requires investments in the training of the human capital employed who must acquire skills in telemedicine and in the applications of artificial intelligence to offer integrated domestic healthcare services especially to elderly patients [4].

Home care and treatments. There are economic advantages in home treatment of dementia even in non-elderly patients [5]. Home care is positively associated with increasing in cost-effectiveness in the case of the treatment of severe heart disease in a cohort of New York patients: specifically home tele monitoring generates an increase in cost-effectiveness ratio of 12,479 euros in respect to usual care while nurse telephone support improves the same measures of 8,270 euros [6]. The increase in the level of home care services in an elderly population delays the entry into a permanent residential care organization [7]. Elderly people who receive homecare show have lower levels of sensory and cognitive deficits than elderly people assisted in Long Term Care-LTC [8]. The cost-effectiveness of telehealth is greater than the cost-effectiveness of traditional care in an empirical study of North Denmark [9]. The cost of homecare can increase significantly if patients are subject to frailty. In a study carried out in Ontario, Canada, the cost of the onset of frailty in patients receiving home care was calculated. The onset of frailty increases the cost of home care by $10,845.00 in patients with dementia and by $12,360.00 in patients without dementia [10]. Providing home care to the elderly at home has various advantages, especially in the case of people with long-term illnesses. From the point of view of the cost-benefit analysis, there is a reduction in the use of healthcare, visits to the emergency room, hospital admissions, and use of a general practitioner. Furthermore, elderly people cared for at homecare level can also have various advantages in terms of reducing depression and loneliness if a multidisciplinary team cares for them [11]. The development of home care for patients has generated the emergence of private companies, which, for profit reasons, are specialized in offering homecare services to patients i.e.: Home Healthcare Agencies-HHAs. From an economic-financial point of view, the assignment of healthcare to HHAs presents numerous advantages with an increase in overall efficiency of 28%, technical efficiency by 23% and scale efficiency by 6% [12]. Physical exercise has positive impacts for elderly people treated at home with frailty. An empirical study conducted in Finland shows that physical exercise increases life expectancy by 0.04 quality-

adjusted years. From an economic-financial point of view, elderly people treated at home and carrying out sporting activities used less healthcare services, generating savings for the healthcare system during the observation period [13].

Alternatives to home care. Under a cost-benefit analysis point of view, there are medical interventions that can reduce the cost of home care as in the case of thrombectomy [14].

Negative home care outcomes. The cost effectiveness of collaborative dementia care management is greater for people that live alone, as showed in a cohort of 444 patients with a reduction of costs equal to 569,00 euros [15]. Caring for the elderly at home can generate nutritional problems. In fact, elderly people who are cared for at home may not have adequate nutrition. The lack of adequate nutrition could induce depression and increase suicidal tendencies in elderly people cared for at home [16]. Home care of older adults with dementia can significantly reduce the well-being of caregivers. Healthcare workers who assist the elderly at home show high level of anxiety, depression and sleep disorders [17].

3) The Econometric Model for the Estimation of the Value of Elderly Treated in Integrated Home Care-EPIHC

In the following section, we applied various econometric techniques to estimate the value of EPIHC using ISTAT-BES data at a regional level between 2015 and 2021 in Italy. The data were analyzed through the use of Panel Data with Random Effects, Panel Data with Fixed Effects, Pooled Ordinary Least Squares-OLS and Weighted Least Squares-WLS. Specifically, we estimated the following equation:

$$\begin{aligned} \boldsymbol{EPIHC_{it}} = \boldsymbol{\alpha_1} + \boldsymbol{\beta_1}(\boldsymbol{UL})_{it} + \boldsymbol{\beta_2}(\boldsymbol{ER})_{it} + \boldsymbol{\beta_3}(\boldsymbol{SWD})_{it} + \boldsymbol{\beta_4}(\boldsymbol{ESF})_{it} + \boldsymbol{\beta_5}(\boldsymbol{TPF})_{it} \\ + \boldsymbol{\beta_6}(\boldsymbol{WPRL})_{it} + \boldsymbol{\beta_7}(\boldsymbol{POF})_{it} + \boldsymbol{\beta_8}(\boldsymbol{DRMH})_{it} + \boldsymbol{\beta_9}(\boldsymbol{CLD})_{it} + \boldsymbol{\beta_{10}}(\boldsymbol{WT})_{it} \\ + \boldsymbol{\beta_{11}}(\boldsymbol{SWAC})_{it} + \boldsymbol{\beta_{12}}(\boldsymbol{CCC})_{it} + \boldsymbol{\beta_{13}}(\boldsymbol{IPS})_{it} + \boldsymbol{\beta_{14}}(\boldsymbol{NM})_{it} \end{aligned}$$

Where $\boldsymbol{i = 18}$ and $\boldsymbol{t = [2015; 2021]}$.

A synthesis of the statistical results is showed in Table 1.

Statistical Results For the Estimation of EPIHC																
Model		**const**	**UL**	**ER**	**SWD**	**ESF**	**TPF**	**WPRL**	**POF**	**DRMH**	**CLD**	**WT**	**SWAC**	**CCC**	**IPS**	**NM**
Pooled OLS	*Coefficient*	0,024	-0,032	0,040	-0,045	0,022	-0,726	0,022	-0,044	-0,120	-0,050	0,004	0,143	0,082	-0,005	0,372
	p-value		***	***	***	***	***	***	***	***	***	**	***	***	***	***
	Standard Deviation	0,043	0,007	0,007	0,010	0,003	0,099	0,004	0,011	0,029	0,009	0,001	0,011	0,011	0,001	0,018
Fixed-effects	*Coefficient*	0,009	-0,025	0,032	-0,036	0,022	-0,657	0,020	-0,028	-0,088	-0,039	0,004	0,151	0,067	-0,004	0,346
	p-value		***	***	***	***	***	***	*	**	***	***	***	***	***	***
	Standard Deviation	0,052	0,008	0,008	0,011	0,004	0,120	0,006	0,015	0,041	0,012	0,002	0,016	0,013	0,001	0,022
Random-effects	*Coefficient*	0,016	-0,027	0,033	-0,037	0,022	-0,669	0,020	-0,032	-0,095	-0,041	0,004	0,149	0,070	-0,004	0,352
	p-value		***	***	***	***	***	***	**	**	***	***	***	***	***	***
	Standard Deviation	0,081	0,008	0,008	0,011	0,004	0,118	0,006	0,015	0,040	0,012	0,001	0,016	0,013	0,001	0,021
WLS	*Coefficient*	0,002	-0,016	0,034	-0,039	0,022	-0,633	0,020	-0,026	-0,092	-0,028	0,004	0,153	0,061	-0,004	0,342
	p-value		**	***	***	***	***	***	**	***	***	***	***	***	***	***
	Standard Deviation	0,043	0,007	0,007	0,010	0,003	0,099	0,004	0,011	0,029	0,009	0,001	0,011	0,011	0,001	0,018

	Average of Coefficients	0,013	-0,025	0,035	-0,039	0,022	-0,671	0,021	-0,032	-0,099	-0,039	0,004	0,149	0,070	-0,004	0,353

Table 1. Statistical Results for the Estimation of EPIHC with Pooled OLS, Panel Data with Fixed Effects, Panel Data with Random Effects, WLS-Weighted Least Squares.

The level of EPIHC is positively associated to:

- *Nurses and Midwives-NM:* is a variable that considers the number of nurses and midwives per 1,000 inhabitants. There is therefore a positive relationship between the EPIHC value and the NM value. Regions that have a high EPIHC value also tend to have a high NM value. Specifically we can see that many regions that have an NM value above the average also have an EPIHC value equal to or above the average. For example, if we consider 2021, these regions are: Molise with a NM value of 8.6 and an EPIHC value of 2.9, followed by Liguria with a value of 7.7 and an amount of 2.9, Basilicata with a value of 7.4 and a value of 3.9, Friuli Venezia Giulia with a value of 7.2 and 3.1, Tuscany with a value of 7 and 3.7 and Abruzzo with a value of 7 and 4.7. It follows that in the regions where the EPIHC value is used more widely there is also a greater value in terms of NM. This relationship may be due precisely to the presence of medium-high levels of the EPIHC value. That is, since the demand for home care for the elderly tends to grow, there is also a growth in employment in professions specialized in care and assistance for the population. Furthermore, if we take into consideration the average of the EPIHC value and the average of the NM value between 2015 and 2021 for the regions considered we can notice the existence of a positive relationship also from a graphic as well as numerical point of view (Figure 1).

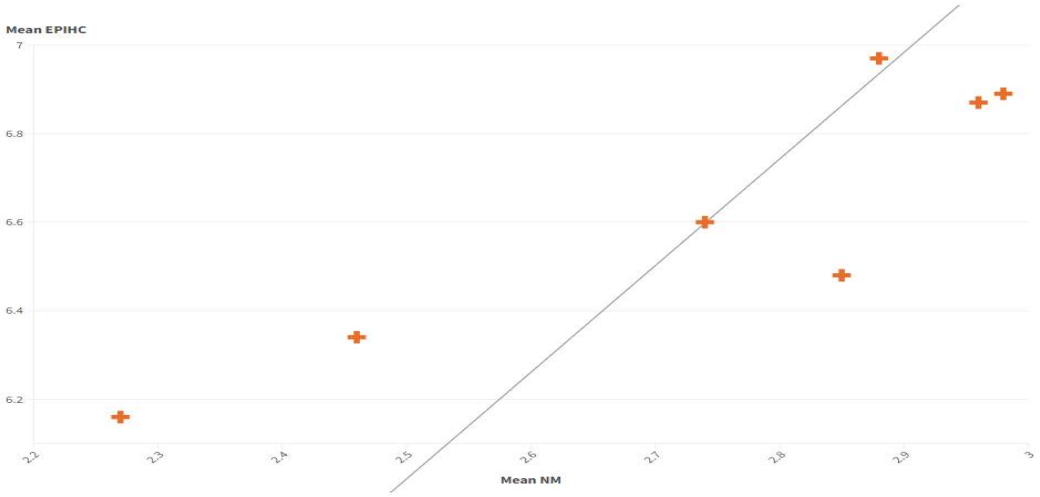


Figure 1. Relationship between the average EPIHC value and the average NM value for the regions considered between 2015 and 2021. Source: ISTAT-BES.

- *Soil Waterproofing from Artificial Cover-SWAC:* is a variable that considers the percentage of waterproofed soil on total land area. There is a relationship between the value of SWAC and the value of EPIHC in the Italian regions in the period considered between 2015 and 2021. In fact we can note that if we take 2021 into consideration it occurs that many of the regions that have a high value in terms of SWAC also have a high value in terms of EPIHC. Among these regions we can note Lombardy with a SWAC value equal to 12.37 units and EPIHC equal to 2.8, Veneto with a SWAC value equal to 11.93 and an EPIHC value equal to 4.3, followed from Emilia Romagna with a SWAC value of 8.95 and an EPIHC value of 3.6, Friuli Venezia Giulia with a SWAC value of 8.09 and an EPIHC value of 3.1, and finally Liguria with a SWAC value of 7.84 units and an EPIHC value of 2.9 units. It follows that the regions

in which there is a greater capacity for home care for the elderly are also the same regions in which a phenomenon of overbuilding and land consumption persists. However, it is very likely that this phenomenon will be reversed in the future by applying the rules relating to Environment Social and Governance-ESG systems (Figure 2).

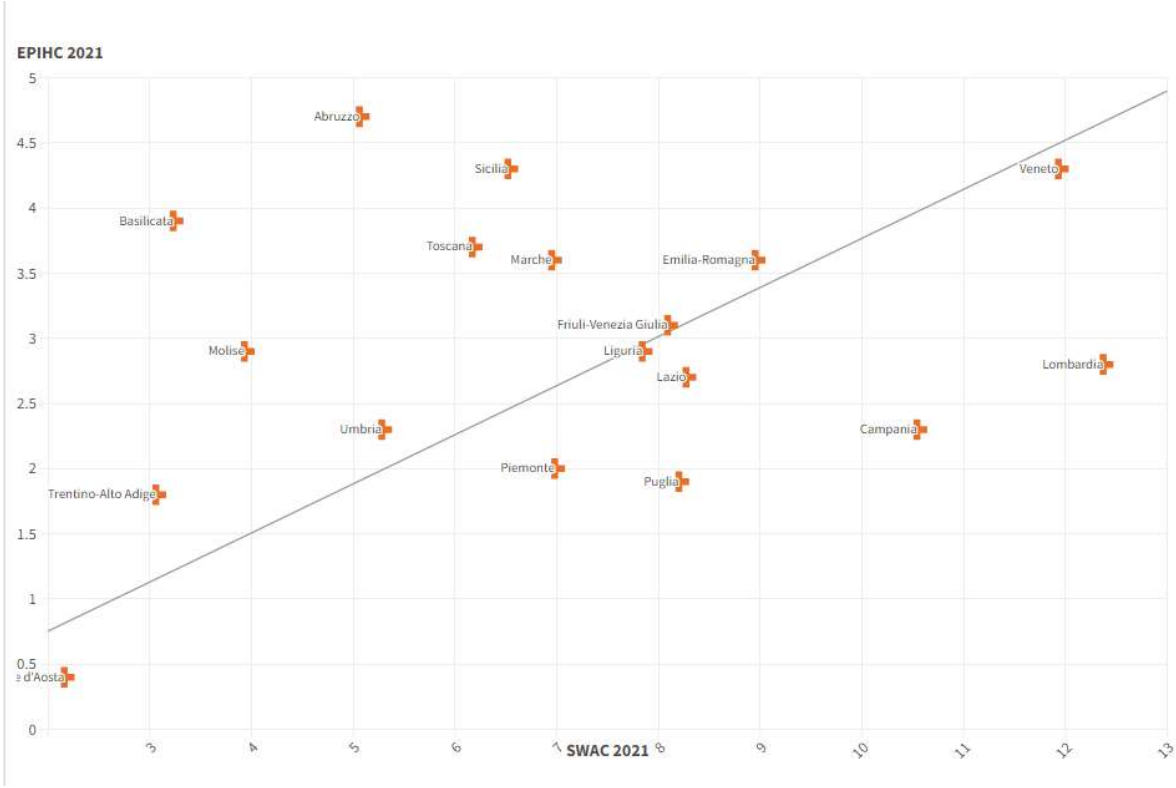


Figure 2. Relationship between the EPIHC value and the SWAC value in 2021. In regions where there is a high level of EPIHC there is also significant land consumption.

- *Concern about Climate Change-CCC:* is a variable that considers the value of the Percentage of people aged 14 and over indicate the ruin of the landscape caused by excessive building construction among the five problems most worrying environmental issues among the total number of people aged 14 and over. There is a positive relationship between the value of regions in which there is a high level of EPIHC and the regions in which there is a significant value of CCC. For example, if we take into consideration the data relating to 2021 we can notice that many regions that have a CCC value above the average also have an EPIHC value above the average. In particular, we note that Tuscany has a CCC value of 70.1 and an EPIHC value of 3.7, the Marche has a CCC value of 69.3 and an EPIHC value of 3.6 , Veneto has a CCC value equal to 68.6 and an EPIHC value equal to an amount of 4.3, Emilia Romagna has a CCC value equal to 68.2 units and an EPIHC value equal to an amount of 3.6, Abruzzo has a CCC value equal to 67.5 and an EPIHC value equal to an amount of 4.7 units, Friuli Venezia Giulia has a CCC value equal to 67.3 and a EPIHC value equal to an amount of 3.1. It therefore follows that many of the regions that have a high level of CCC also have a high level of EPIHC. This relationship may be due to the presence of medium-high level human and social capital. In fact, where human and social capital improves, there is an improvement in attention towards the environment and in solidarity as in the case of the provision of care systems for the elderly (Figure 3).

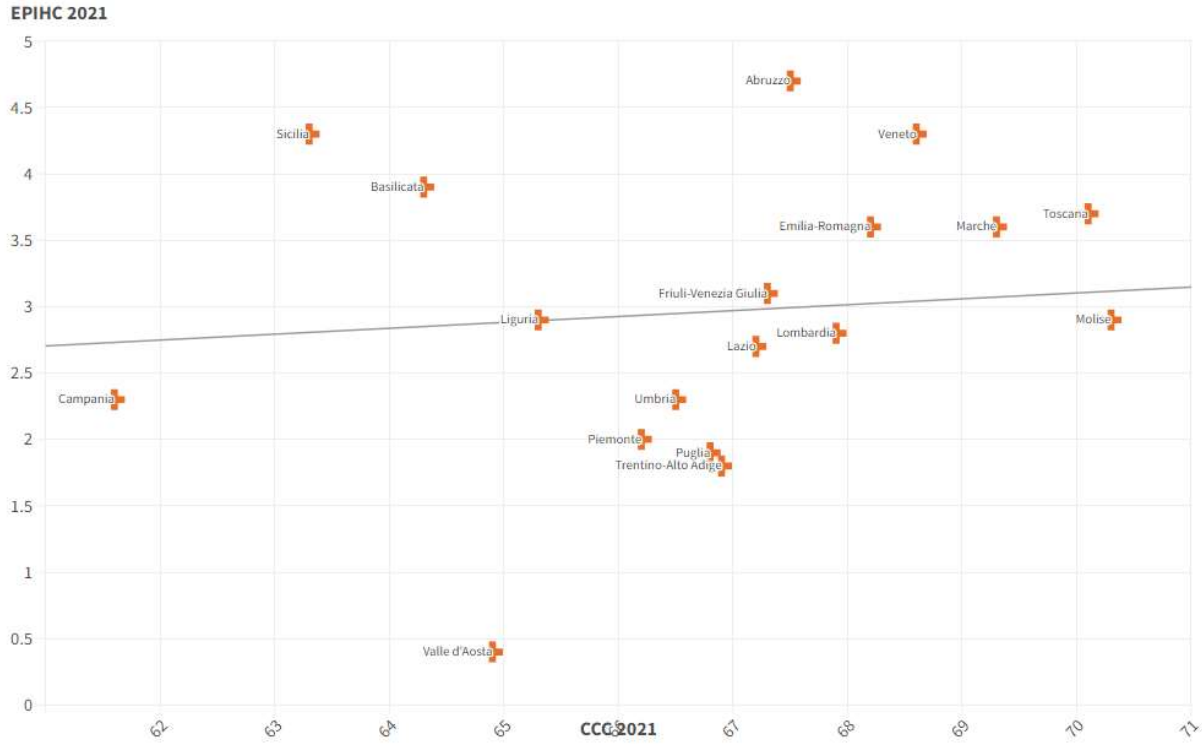


Figure 3. Positive relationship between the value of CCC and the value of EPIHC in 2021. Regions that have a high level of EPIHC also have a high level of CCC. Source: Istat-BES.

- *Employment Rate-ER:* is a variable that considers the percentage of employed people aged 20-64 in the population aged 20-64. There is a positive relationship between the value of ER and the value of EPIHC. The available data refers to the period 2018-2021. If we average each region in the period 2018-2021 for both EPIHC and ER and then compare the results obtained, we can see that many regions have a high value in both variables. That is: Trentino Alto Adige with an average ER value of 75.45 and an average EPIHC value of 20.45, Emilia Romagna with an average ER value of 74.13 and an EPIHC value of 21.53, Valle d'Aosta with an average ER value equal to 72.28 and an EPIHC value equal to an amount of 18.60, Lombardy with an average ER value equal to 72.18 and an EPIHC value equal to an amount of 20.38, Friuli Venezia Giulia with an ER value equal to 71.48 units and an EPIHC value equal to an amount of 20.28, Veneto with an ER value equal to an amount of 71, 30 and an EPIHC value equal to a value of 21.08, Tuscany with an ER value equal to 70.93 and an EPIHC value equal to an amount of 20.58, Piedmont with an ER value equal to an amount of 70.00 and an EPIHC value of 19.53. It follows that most of the regions that have a high value of the employment rate also have a high value of home care for the elderly. This relationship could also be because in regions where there is greater employment there are also greater economic-financial resources to support home care services (Figure 4).

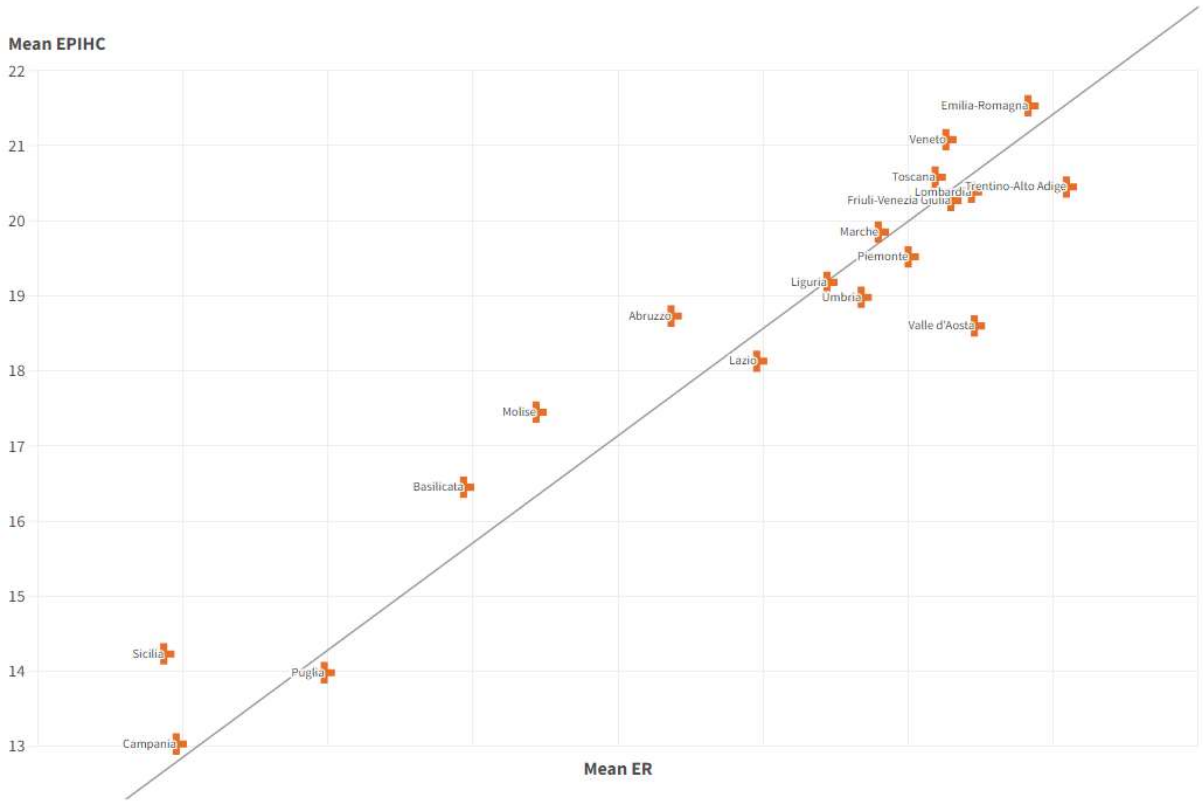


Figure 4. Positive relationship between the ER value and the EPIHC value. Regions where there is greater employment also have higher levels of home care for the elderly.

- *Economic Situation of the Family-ESF:* is a variable that considers the value of the Families who declare their economic situation has worsened or significantly worsened compared to the previous year. There is a positive relationship between the regions in which the condition of families has worsened and the regions in which home care for the elderly has grown. Considering the average for each region in the period 2016-2021 for both variables, we verified that many regions that have a high value in terms of ESF also have high values in terms of EPIHC. For example, Sicily has an ESF value equal to 35.93 and an EPIHC value equal to an amount of 3.95, Veneto with 31.33 ESF and 3.82 EPIHC, Abruzzo with 30.3 ESG and an EPIHC value equal to 3.83, Molise with an ESF value equal to 29.95 and an EPIHC value equal to an amount of 4.27, Friuli Venezia Giulia with an ESF value equal to an amount of 29.9 units and an EPIHC value equal to 3.05, Tuscany with an ESF equal to 29.83 and an EPIHC value equal to 3.28, Liguria with an ESF value equal to 29.45 and a of EPIHC equal to 3.1. This result may appear paradoxical. In fact, the possibility of increasing domestic care for the elderly should be directly proportional to the economic possibilities of families. That is, positive increases in family income should lead to increases in domestic home care. However, it is necessary to consider that the ESF value decreased in almost all regions between 2015 and 2021 with the exception of Tuscany, Trentino Alto Adige and Valle d'Aosta (Figure 5).

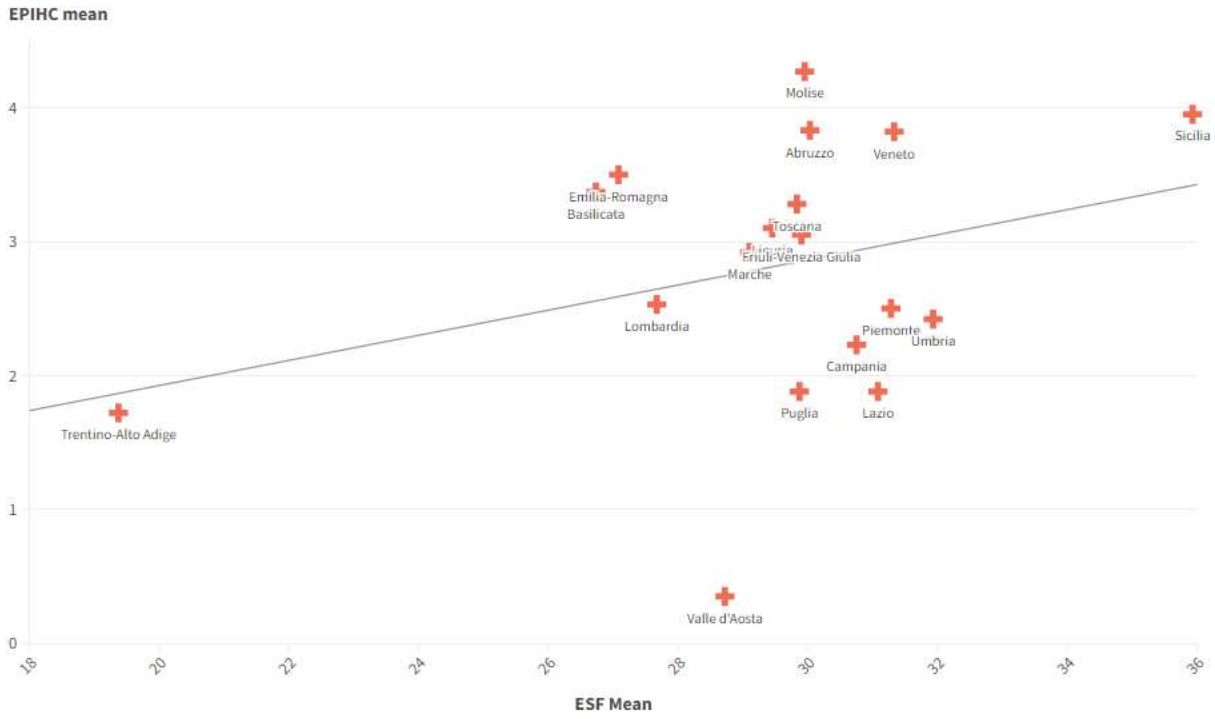


Figure 5. The positive relationship between the mean ESF value and the mean EPIHC value. In regions where the economic insecurity of families is growing, home care for the elderly increases.

- *Women and Political Representation at Local Level-WPRL:* it is a variable that considers the Percentage of women elected to Boards regional out of the total elected. There is a positive relationship between the WPRL value and the EPIHC value. Taking the average for each region in the period 2015-2021 for each variable, i.e. both WPRL and EPIHC, it turns out that many regions have high values both in terms of EPIHC and in terms of WPRL. Specifically we can note for example that Emilia-Romagna has a WPRL value equal to 34.86 and an EPIHC value equal to 3.49, followed by Tuscany with a WPRL value equal to 29.14 and EPIHC equal to 3.19, Veneto with WPRL equal to 25.51 and EPIHC equal to 3.57, Molise with WPRL equal to 22.47 and EPIHC equal to 4.20, Marche with a WPRL value equal to 22.14 and an amount of EPIHC equal to 2.74. Regions that have a high level of women who are present in local institutions also have medium-high levels of value of home care for the elderly. This positive relationship indicates the presence of medium-high level human and

social capital, which allows for the reduction of social differences both in the sense of gender inequalities and at the level of intergenerational inequalities (Figure 6).

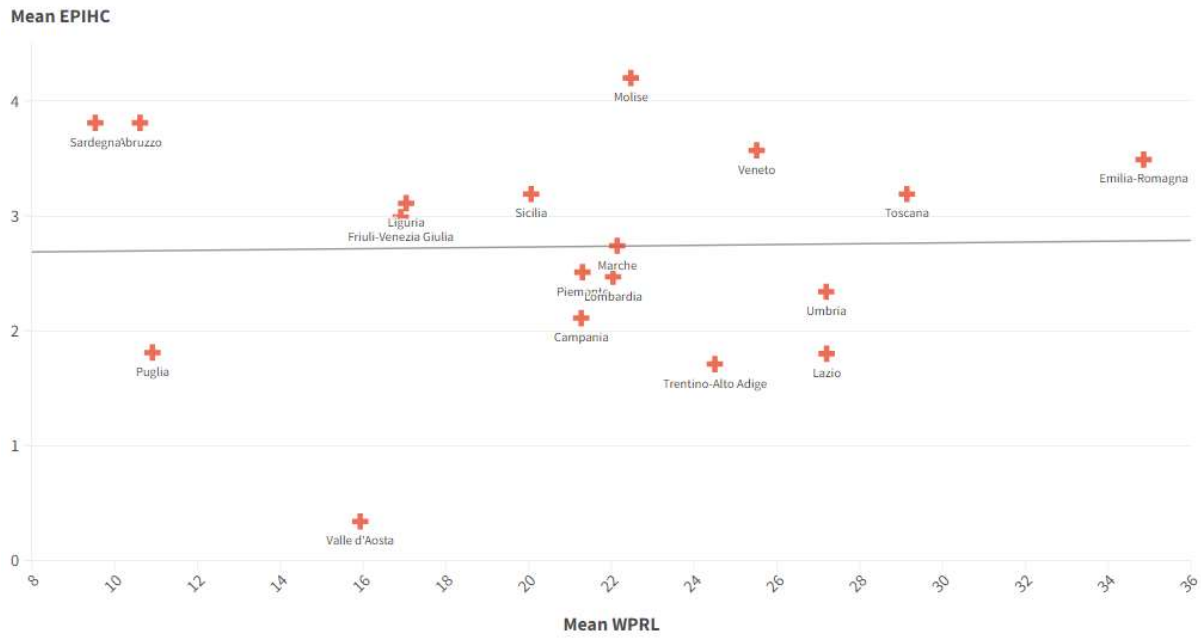


Figure 6. Positive relationship between the average EPIHC value and the average WPRL value. Regions that have a high number of women in local institutions also have high levels of home care for the elderly.

- *Wastewater Treatment-WT:* is a variable that considers the percentage share of polluting loads flowing into secondary plants or advanced, in equivalent inhabitants, compared to the loads urban totals (Aetu) generated. There is a positive relationship between the WT value and the EPIHC value. Specifically we can note that, considering the year 2015 it is possible to note that there are regions that have a WT value above the average in connection with high levels of EPIHC. For example, Piedmont with a WT value equal to 69.70 and EPIHC equal to 2.6, Emilia Romagna with a WT value equal to 67.70 units and EPIHC equal to 3.4, and Abruzzo with a of WT equal to 63.90 and an EPIHC value equal to an amount of 3.7. We can see that the value of WT has grown significantly in the time series considered. This orientation of the time series could be further increasing in the future, determining a positive trend between the WT value and the EPIHC value. It is therefore necessary to intervene through economic policies to reduce the value of WT.

The level of EPIHC is negatively associated to:

- *Innovation of the Production System-IPS:* indicates the percentage of companies that have introduced product innovations and processes in the three-year reference period out of the total of companies with at least 10 employees. There is a negative relationship between the IPS value and the EPIHC value at a regional level between 2015 and 2021. If for example we consider the 2020 value we can notice that there are many regions that have an IPS value above the average and a EPIHC value lower than the average i.e.: Piedmont with an IPS value equal to 54.57 and an EPIHC value equal to 2.63, Lombardy with an IPS value equal to 53.97

and an EPIHC value equal to 2, 57, Marche with an IPS value of 50.13 and an EPIHC value of 2.67, Umbria with a PIS value of 49.03 and an EPIHC value of 2.17, Trentino Alto Adige with an IPS value of 48.47 and an EPIHC value of 1.70. However, we must consider that even if the value of the relationship between IPS and EPIHC is negative, it is a value very close to zero. In fact, in the case of Pooled OLS this value is equal to -0.005, Fixed Effects equal to -0.004 units, Random Effects equal to -0.004, WLS equal to -0.004 with an average final value equal to -0.004 units. It follows that marginal variations in both variables could transform the negative relationship between IPS and EPIHC into a positive one.

- *Use of Libraries-UL:* represents the Percentage of people aged 3 years and older who went to the library at least once in the 12 months preceding the interview out of the total number of people aged 3 years and older. There is a negative relationship between the value of UL and the value of EPIHC. The regions in which home care for the elderly is very high are also regions in which the population uses library services insufficiently. In fact, if we consider 2021 we can note that there are various regions that have a UL value higher than average despite having an EPIHC value lower than average. These regions are indicated below: Trentino Alto Adige with a UL value equal to an amount of 29.40 units and an EPIHC value equal to 1.77, Valle d'Aosta with a UL value equal to 24.47 and a UL value equal to 0.50 units, Lombardy with a UL value equal to 17.33 and an EPIHC value equal to 2.73, Piedmont with a UL value equal to 13.40 units and a value of EPIHC equal to 2.40 units (Figure 7).

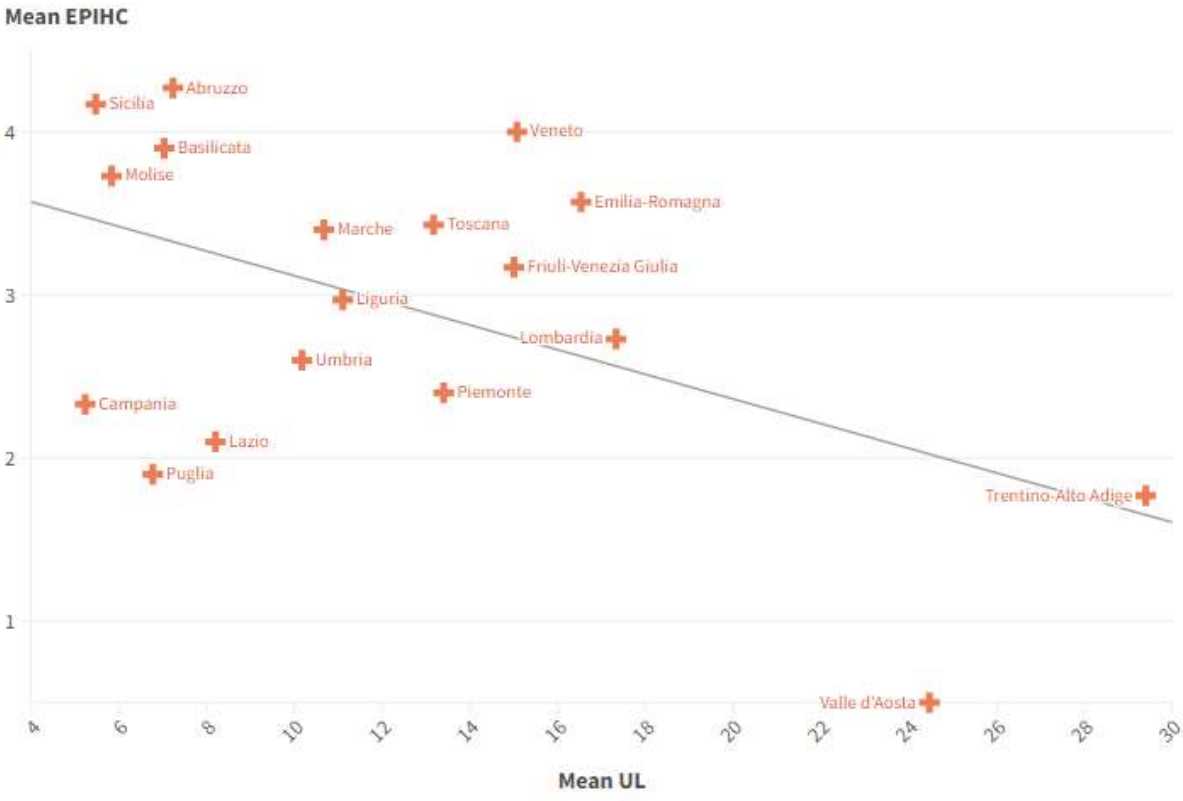


Figure 7. Relationship between the UL value and the EPIHC value in the Italian regions. The UL value tends to decrease with the growth of EPIHC.

- *Positive Opinion on Future Prospects-POF:* represents the percentage of people aged 14 and over who believe that their personal situation will improve in the next 5 years out of the total number of people aged 14 and over. There is a negative relationship between the POF value and the EPIHC value. The growth of integrated home care for the elderly is associated with a worsening of the value of positive expectations relating to the future. By averaging both the EPIHC value and the POF value for each Italian region between 2015 and 2021 we can see that there are many regions that have a high level of EPIHC and a reduced value in terms of POF, that is: Lombardy has a POF value of 32.09 and an EPIHC value of 2.47, Valle d'Aosta has a POF value of 30.86 and an EPIHC value of 0.34, Lazio has a POF value equal to 30.56 and an EPIHC value equal to 1.8, Trentino Alto Adige has a POF value equal to 28.64 and an EPIHC value equal to 1.71 units (Figure 8).

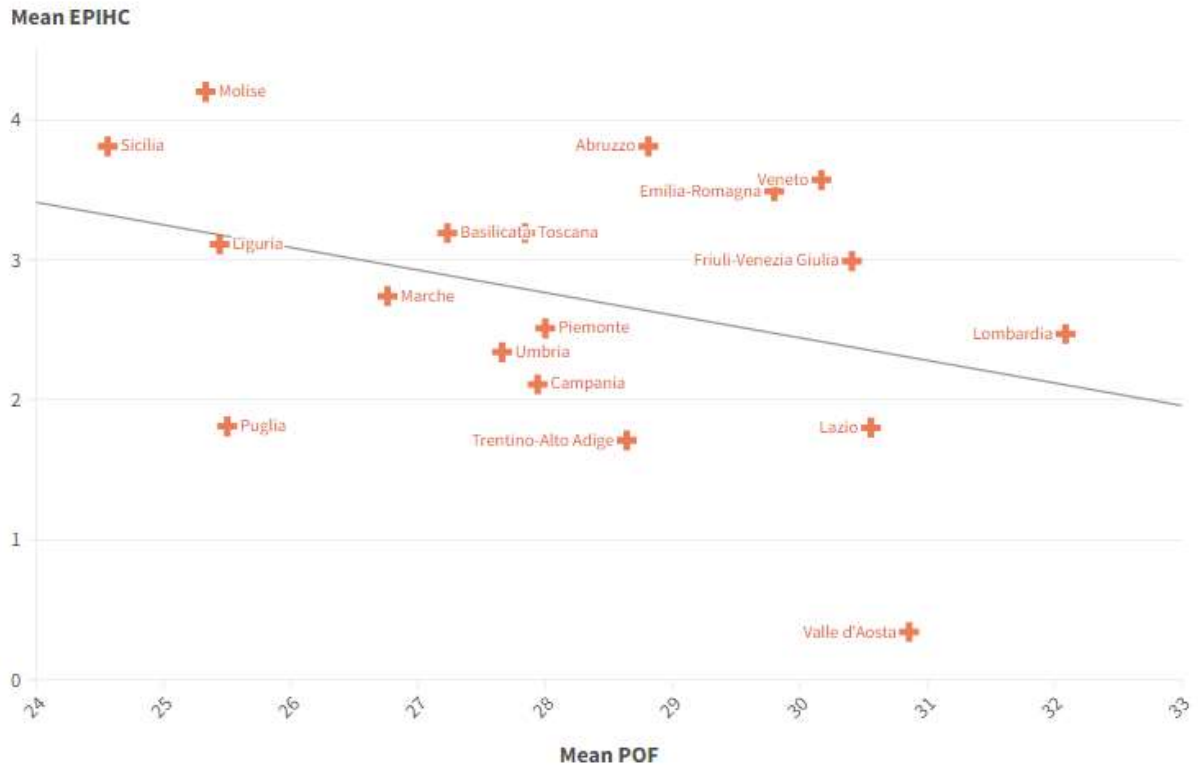


Figure 8. Relationship between the average EPIHC value and the average POF value in the Italian regions between 2015 and 2021.

- *Satisfaction with the Work Done-SWD:* represents the Percentage of employed people who expressed an average score of satisfaction between 8 and 10 for the following aspects of the work performed: earnings, career opportunities, number of hours worked, job stability, distance from home to work, interest in the job. There is a negative relationship between the SWD value and the EPIHC value. Specifically, regions that have high levels of EPIHC have reduced levels of SWD. In fact, if we take the average of the values for each region of both SWD and EPIHC we can see that many countries that have high levels of EPIHC also have reduced levels in terms of SWD. For example, Trentino Alto Adige has an average SWD value between 2018 and 2021 equal to 61.58 and an EPIHC value equal to an amount of 1.7 units, followed by Valle d'Aosta with a value in terms of SWD equal to 57.33 and an EPIHC

value equal to 0.43 units, Piedmont with a SWD value equal to 52.5 and an EPIHC value equal to 2.5, Umbria with SWD equal to 52.5 units and an EPIHC value of 2.45, Lombardy with an SWD value of 48.6 and an EPIHC value of 2.68 (Figure 9).

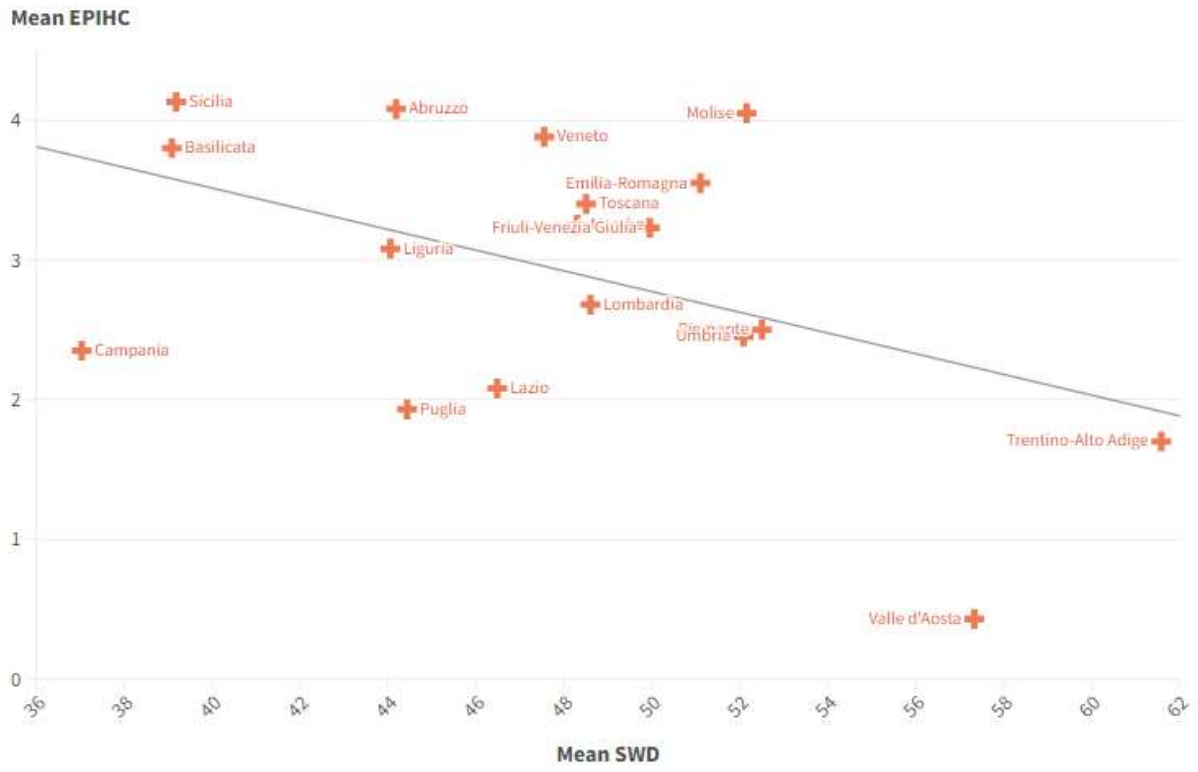


Figure 9. Relationship between the EPIHC value and the SWD value in the Italian regions.

- *Concern about Landscape Deterioration-CLD:* represents the percentage of people aged 14 years and over and indicates the ruin of the landscape caused by excessive construction among the five most worrying environmental problems out of the total number of people aged 14 years and over. There is a negative relationship between the CLD value and the EPIHC value. In fact, if we consider the average of the values of the variables in the Italian regions between 2015 and 2021 we can notice that many regions that have a high value in terms of CLD also have a reduced value in terms of EPIHC. For example, Trentino Alto Adige has a CLD value of 18.47 and an EPIHC value of 1.71, Lombardy has a CLD value of 17.64 and an EPIHC value of 2.47, Valle d'Aosta has a CLD value of 16.41 and an EPIHC value of 0.34, Piedmont has a CLD value of 15.13 and an EPIHC value of 2.51, the Lazio has a CLD value equal to an amount of 13.63 and an EPIHC value equal to 1.8 units (Figure 10).

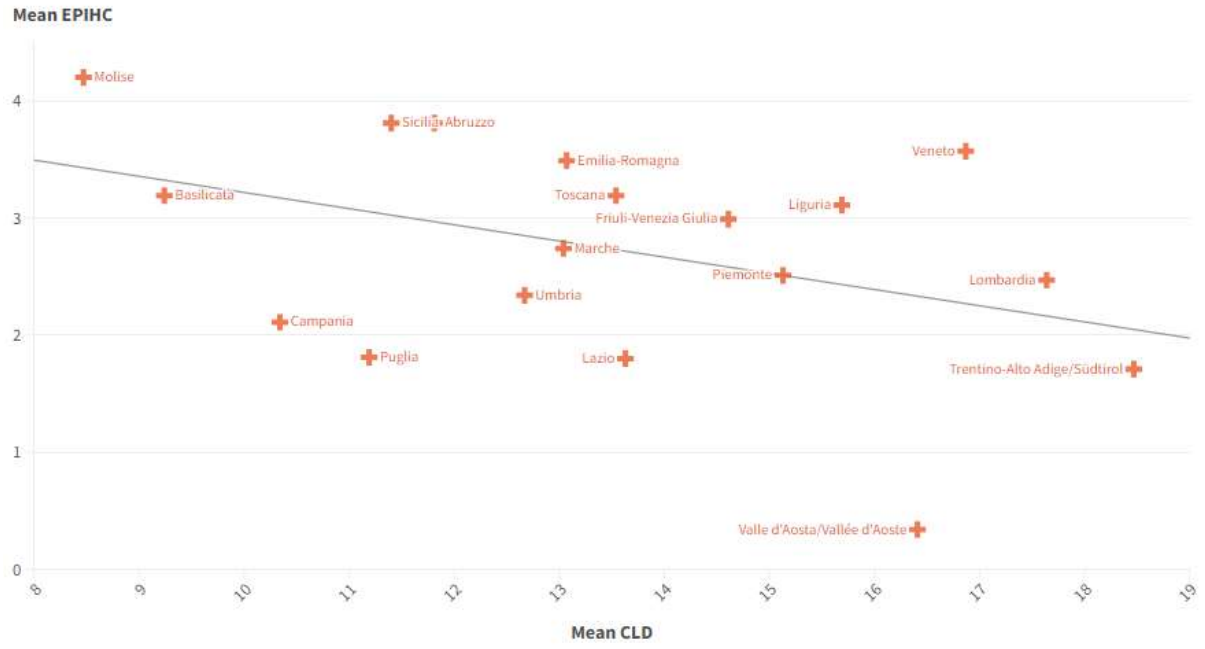


Figure 10. Relationship between mean EPIHC and mean CLD in Italian regions between 2015 and 2021.

- *Density and Relevance of Museum Heritage-DRMH:* is a variable that considers the Number of permanent exhibition structures per 100 km2 (museums, archaeological areas and monuments open to the public), weighted by the number of visitors. The weight of each structure is assumed to be equal to (Vi/VM), where Vi is the number of visitors to the structure, M the total structures and V the total visitors. There is a negative relationship between the value of DRMH and EPIHC. Considering the average value of the two variables i.e. DRMH and EPIHC between 2015 and 2021, it appears that many of the variables that have a DRMH value higher than the average have an EPIHC value lower than the average. For example, Lazio with an average DRMH value of 5.91 and an EPIHC value of 1.88, Campania with a value of 3.59 units and an EPIHC value of 2.13, and Lombardy with a DRMH value of 1.53 and an EPIHC value of 2.48. The relationship between the EPIHC value and the DRMH value affects many regions which, despite having a DRMH value lower than the average, also have a high EPIHC value such as Friuli Venezia Giulia with a DRMH value equal to 1.43 and an EPIHC value of 3.05, Emilia Romagna with a DRMH value of 1.10 and an EPIHC value of 3.52, Sicily with an EPIHC value of 0.98 and a DRMH value of at 3.87, and Molise with a DRMH value of 0.17 and an EPIHC value of 4.23 (Figure 11).

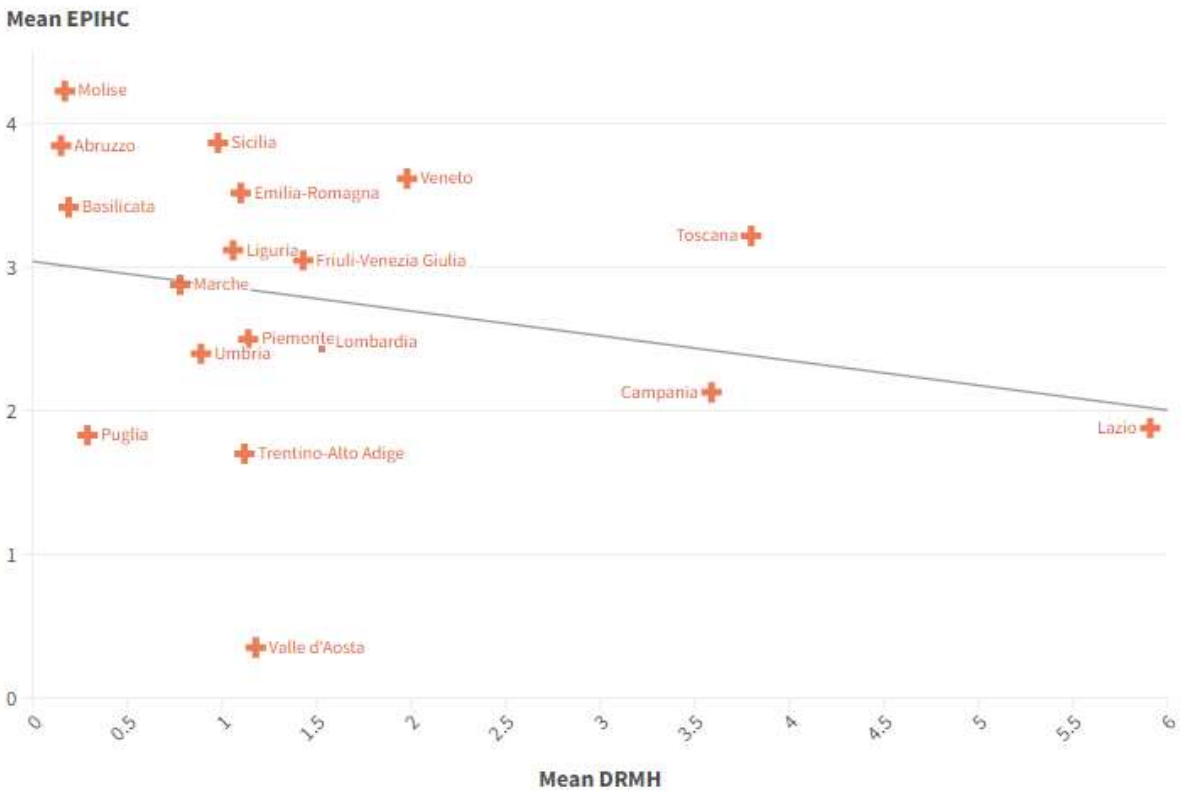


Figure 11. Relationship between the value of EPIHC and the value of DRMH in the Italian regions between 2015 and 2021.

- *Trust in the Police and Firefighters-TPF:* is a variable that calculates the average score of trust in the police and firefighters (on a scale from 0 to 10) expressed by people aged 14 and over. There is a negative relationship between the TPF value and the EPIHC value. In fact, by calculating the average value of TPF and EPIHC in the Italian regions between 2015 and 2021 we can see that there are many regions that have a TPF value higher than the average and an EPIHC value lower than the average. For example, Trentino Alto Adige has a TPF value of 7.57 and an EPIHC value of 1.71, Friuli Venezia Giulia with a TPF value of 7.56 and a value of 2.99 units, Piedmont with a TPF value of 7.43 and an EPIHC value of 2.51, Umbria with a TPF value of 7.41 and an EPIHC value of 2.34, Lombardy with a TPF value of 7.36 and an EPIHC value of 2.47, Valle d'Aosta with a TPF value of 7.36 and an EPIHC value of 0.34 (Figure 12).

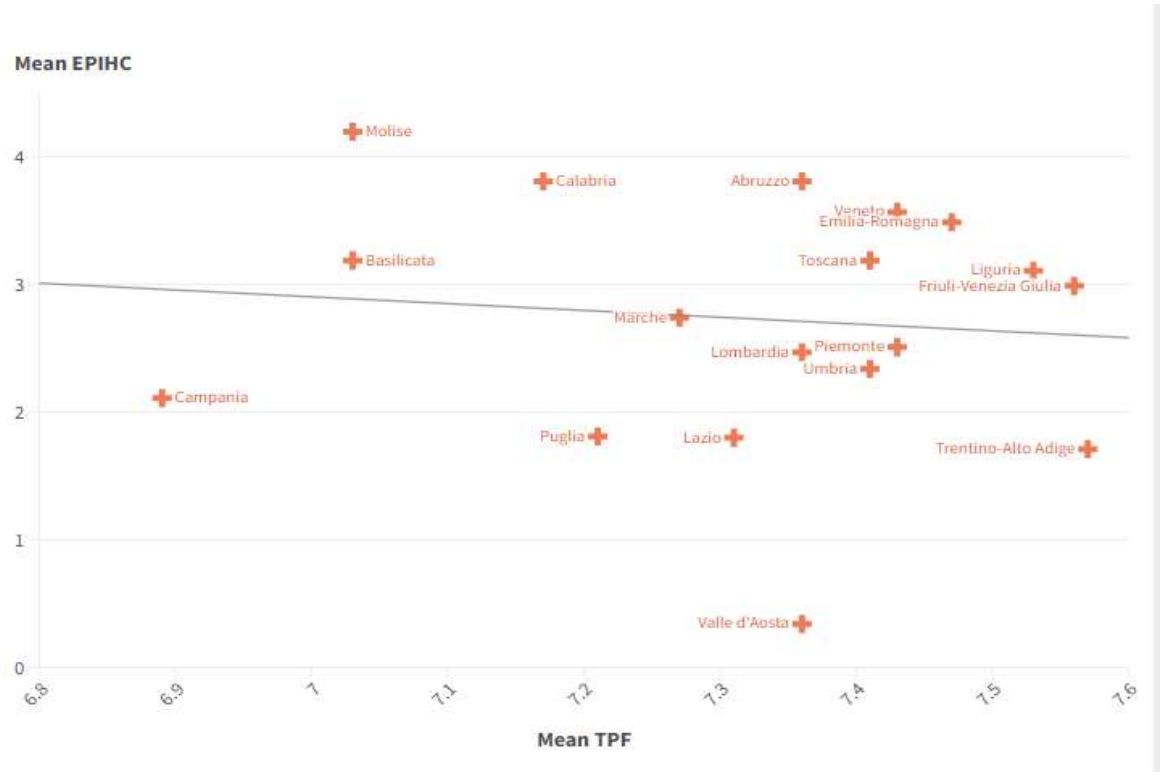


Figure 12. There is a negative relationship between the EPIHC value and the TPF value in the Italian regions between 2015 and 2021.

4) Rankings and Clusterization with k-Means algorithm optimized with Silhouette Coefficient

Below we take into consideration the trend of the EPIHC variable in the context of the Italian regions as indicated in the ISTAT-BES dataset between 2015 and 2021. Before analysing the data from the point of view of clustering with the k-Means algorithm, we verify the performance of the Italian regions by EPIHC value. The data relating to Sardinia and Calabria were not considered due to the lack of completeness in the historical series.

Ranking of the Italian regions by EPIHC value in 2021. Abruzzo is in first place by EPIHC value with an amount of 4.7, followed by Veneto and Sicily both with a value of 4.3. In the middle of the table are Liguria and Molise with a value of 2.9 units followed by Lombardy with 2.8 units. Puglia closes the ranking with a value of 1.9 units, followed by Trentino Alto Adige with a value of 1.8 units and Valle d'Aosta with an amount of 0.4 units.

Ranking of the Italian regions by value of the percentage change in the EPIHC value between 2015 and 2021. The Marche is in first place for the value of the percentage change in the EPIHC value with an amount equal to 111.76% equivalent to a change from 1 .7 units up to a value of 3.6 units. Lazio follows with a value of 107.69% corresponding to a variation from an amount of 1.3 units up to a value of 2.7 units or equal to a variation of 1.4 units. Veneto follows, where between 2015 and 2021 the value of the percentage change in EPIHC grew by 104.76% or from an amount of 2.1 units up to a value of 4.3 units. In the middle of the table are Valle d'Aosta and Piedmont with a variation of 33.33% followed by Abruzzo with a variation of 27.03% corresponding to a modification from an amount of 3.7 units up to a value of 4.7 units. Liguria closes the ranking with a value equal to -9.38%

equivalent to a variation from an amount of 3.2 units up to 2.9 units, followed by Piedmont with a variation equal to -23.08% corresponding to an amount from 2.6 units up to 2 units, followed by Molise with a variation of -23.68% (Figure 13).

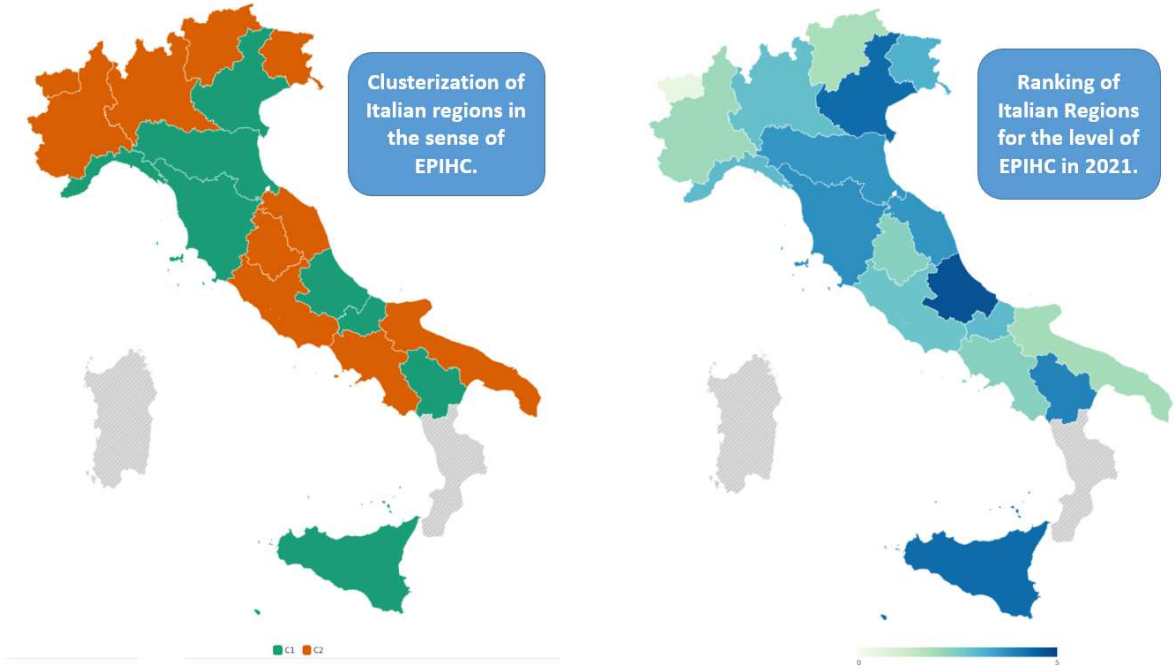


Figure 13. Clusterization with k-Means algorithm for Italian regions in the sense of EPIHC and ranking of Italian regions for the level of EPIHC in 2021. The data relating to Sardinia and Calabria were not considered due to the lack of completeness in the historical series.

Clusterization with k-Means algorithm optimized with Silhouette coefficient. Below we present a clustering with k-Means algorithm optimized with the Silhouette coefficient. The data shows the presence of two clusters. The two clusters are indicated below:

- *Cluster 1*: Emilia-Romagna, Sicily, Abruzzo, Veneto, Tuscany, Liguria, Molise, Basilicata, Friuli Venezia Giulia;
- *Cluster 2*: Puglia, Lazio, Trentino-Alto Adige, Campania, Valle d'Aosta, Umbria, Lombardy, Piedmont, Marche.

From the analysis we can see that Cluster 1 is dominant compared to Cluster 2. Specifically, the following condition occurs: *C2>C1*. We can note that the regions participating in both Cluster 1 and Cluster 2 are very heterogeneous from each other both from an economic point of view, i.e. in terms of GDP per capita, and from a geographical point of view, as it is not possible to find a clear contrast between the regions of the South and the regions of the Centre-North. However, it is very likely that the difference between the regions of Cluster 1 and Cluster 2 is instead to be found in the medical culture and in the health policy choices made at the regional level.

5) Machine Learning and Prediction for the Estimation of the Future Value of EPIHC

Below we use eight different machine-learning algorithms for predicting the future value of EPIHC in Italian regions. The algorithms are analysed because of their ability to maximize the R-squared and minimize the following statistical errors: Mean Absolute Error-MAE, Mean Squared Error-MSE, Root Mean Squared Error-RMSE, and Mean Signed Difference-MSD. The algorithms were trained with 70% of the data while the remaining 30% was used for the actual prediction. To choose the best

predictor algorithm we created rankings of the algorithms for each individual statistical indicator. We then added the positioning of each algorithm within the rankings. It follows that the algorithm that has a lower payoff value is considered the most efficient algorithm in predictive terms (Figure 14).

Figure 14. Statistical Errors of Algorithms.

We therefore obtained the following ordering of the algorithms:

- Linear Regression with a payoff value of 7;
- Gradient Boosted Tree with a payoff value of 12;
- Simple Regression Tree and Random Forest Regression with a payoff value of 19;
- PNN-Probabilistic Neural Network with a payoff value of 25;
- ANN-Artificial Neural Network with a payoff value of 27;
- Tree Ensemble Regression with a payoff value of 32;
- Polynomial Regression with a payoff value of 39 (Figure 15).

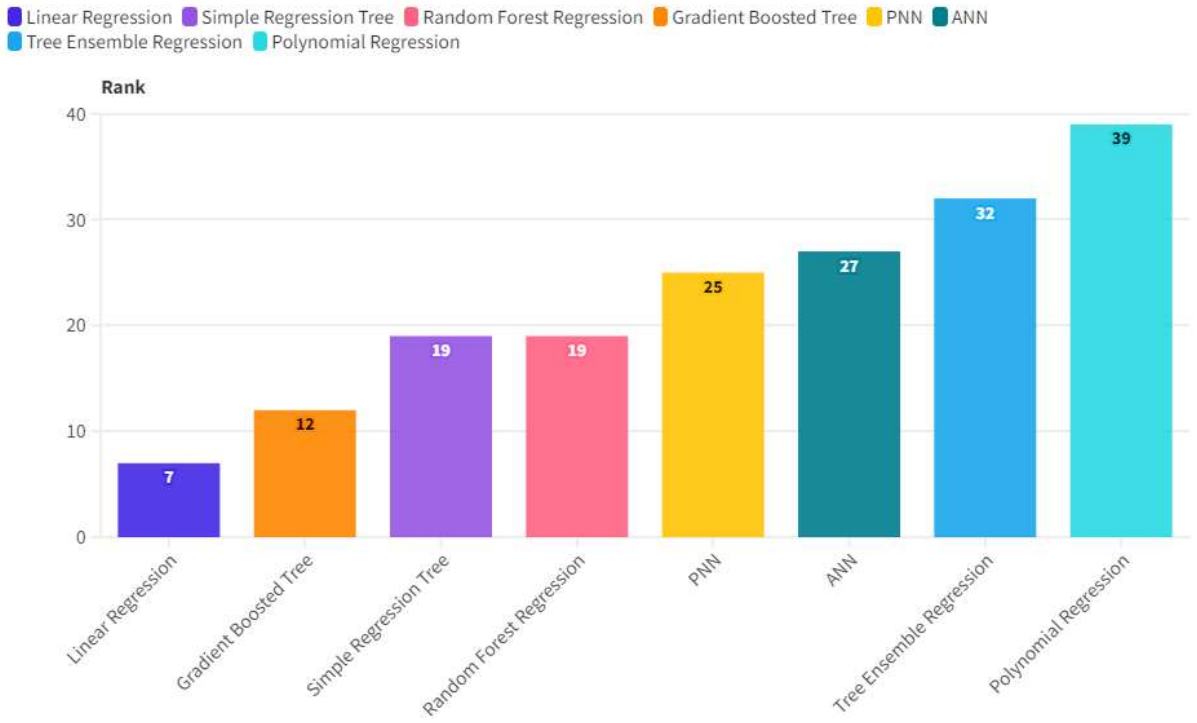


Figure 15. Ranking of Algorithms based on maximization of R-Squared and minimization of Statistical Errors .

It therefore follows that the best performing algorithm in predictive terms is Linear Regression. By applying the Linear Regression algorithm it is possible to predict the following trends in the EPIHC value in the Italian regions, namely:

- Valle d'Aosta with a predicted EPIHC value growing from 0.40 to 0.50 or equal to a value of 0.10 units equal to +25.50%;
- Trentino Alto Adige with a predicted EPIHC value growing from 1.80 to 1.95 or equal to a variation of 0.15 units equivalent to +8.56%;
- Emilia Romagna with EPIHC value predicted to increase from 3.60 to 3.75 or a change equal to +0.15 units equivalent to +4.22%;
- Umbria with a predicted EPIHC value growing from an amount of 2.30 units up to 3.09 units equivalent to a value of 0.79 units equal to +34.43%;
- Lazio with a predicted EPIHC value decreasing from an amount of 2.70 to a value of 1.97 units or equal to an amount of -0.73 units equal to -26.93%;
- Campania with a predicted EPIHC value decreasing from an amount of 2.30 units to 2.24 units or equal to -0.06 units equal to -2.43%.

On average, the EPIHC value is predicted to grow for the regions analyzed by an amount equal to 3.18% (Figure 16).

Figure 16. Prediction with the application of Linear Regression.

6) Discussion and Policy Implications

The value of EPIHC in the Italian regions between 2015 and 2021 grew on average by 30%. There are regions in which the EPIHC value has grown more than the average and these regions are: Marche, Lazio, Veneto, Basilicata, Campania, Sicily, Tuscany, Puglia, Valle d'Aosta, Lombardy. The regions in which the EPIHC value grew less than the average in the period between 2015 and 2021 are Abruzzo, Umbria, Friuli Venezia Giulia, Trentino Alto Adige, Emilia Romagna, Liguria, Piedmont and Molise. We therefore note that the EPIHC value has grown in almost all southern regions with the exception of Molise. On the contrary, many regions of Northern Italy with medium-high per capita incomes have experienced a reduction in the value of EPIHC between 2015 and 2021 as for example in the case of Friuli Venezia Giulia, Trentino Alto Adige, Emilia Romagna, Liguria, Piedmont. It follows that the regions of Southern Italy, characterized by medium-low per capita incomes, have achieved a growth in the EPIHC value, while many Northern regions with medium-high per capita incomes have experienced a reduction in the EPIHC value. The growth in the value of EPIHC therefore seems to be in some way independent of the value of the trend in per capita income. Therefore, it is possible to identify the following propositions:

- In 2021, the average EPIHC value was 3.33 in Southern Italy, 3.08 in Central Italy, 2.61 in Northern Italy.
- Between 2015 and 2021 the average EPIHC value grew by 64.00% in Central Italy, by 29.87% in Southern Italy, and by 16.11% in Northern Italy;

- The value of EPIHC tends to be inversely proportional to the trend in per capita income.

The fact that the EPIHC value tends to be higher in the southern regions raises a series of questions relating to the ability to offer adequate health services to the population. In fact, Southern Italy is the part of Italy with the smallest number of doctors and nurses in respect to the other Italian macro-regions, and therefore it is very unlikely that the growing elderly population can be adequately served, especially in the presence of low incomes. It is therefore probable that the increasing value of EPIHC in the regions of Southern Italy is not the consequence of a health policy at the service of the elderly but rather the consequence of the insufficient offer of hospitals and health facilities, which forces the elderly to be treated at home rather than in specialized health structures.

Policy implications. It should be considered that the fact that the elderly are treated at home is certainly a good thing from the point of view of reducing healthcare costs, especially in public hospitals. Currently in Italy the EPIHC program is supported by the State. The elderly are therefore assisted in their own home through the administration of various health services with a pool of professionals.

However, if we look at the issue of aging prospectively, we can note that Italians tend to have a low level of savings, low incomes, and have a lower propensity both to have children and to get married. It follows that in the future it is likely that the financial fragility of the elderly may increase, and therefore the elderly will probably be characterized by increasing conditions of loneliness and may not even own homes.

Health policies aimed at the elderly must therefore also take into consideration their overall socio-economic condition. It is in fact very paradoxical that in the rich part of Italy the EPIHC value is lower than the corresponding value in the part of Italy with medium-low incomes. It is therefore very likely that in the Italian context, home care is simply the palliative that highlights the lack of adequate hospital and healthcare facilities intended for the care of the elderly.

The policy maker must therefore intervene to ensure that the economic-social inequalities that are also manifested in healthcare are eliminated. In fact, it is very likely that people with lower per capita incomes are less assisted by the healthcare system and therefore do not have access to the necessary care even though they are entitled to it.

7) Conclusions

In this article, we have analysed the ESG determinants of EPIHC in the Italian regions between 2004 and 2022 using the ISTAT-BES database. The econometric results show that the value of EPIHC is positively associated with NM and negatively associated with DRMH and TPF. The cluster analysis with the k-Means algorithm has revealed the presence of two clusters that are heterogeneous either geographically either under a socio-economic point of view. Finally, we perform a confrontation among eight different machine-learning algorithms to predict the future level of EPIHC, and we found that it is expected to growth of 3.18% on average for the analysed regions.

8) References

[1] P. Rantanen, T. Parkkari, S. Leikola, M. Airaksinen e A. Lyles, «An in-home advanced robotic system to manage elderly home-care patients' medications: A pilot safety and usability study,» *Clinical therapeutics,* vol. 5, n. 1054-1061, p. 39, 2017.

[2] E. H. Kim, Z. D. Gellis, C. K. Bradway e B. Kenaley, «Depression care services and telehealth technology use for homebound elderly in the United States,» *Aging & Mental Health,* vol. 23, n. 9, pp. 1164-1173, 2019.

[3] S. Sülz, H. J. van Elten, M. Askari, A. M. Weggelaar-Jansen e R. Huijsman, «eHealth applications to support independent living of older persons: scoping review of costs and benefits identified in economic evaluations,» *Journal of medical Internet research,* vol. 23(3), n. e24363, 2021.

[4] J. A. Nasir, S. Hussain e C. Dang, «An integrated planning approach towards home health care, telehealth and patients group based care,» *Journal of network and computer applications,* vol. 117, pp. 30-41, 2018.

[5] P. Clarkson, L. Davies, R. Jasper, N. Loynes e D. Challis, «A systematic review of the economic evidence for home support interventions in dementia,» *Value in Health,* vol. 20, n. 8, pp. 1198-1209, 2017.

[6] A. S. Grustam, J. L. Severens, D. De Massari, N. Buyukkaramikli, R. Koymans e H. J. Vrijhoef, «Cost-effectiveness analysis in telehealth: a comparison between home telemonitoring, nurse telephone support, and usual care in chronic heart failure management,» *Value in Health,* vol. 21, n. 7, pp. 772-782, 2018.

[7] M. Jorgensen, J. Siette, A. Georgiou, A. Warland e J. Westbrook, «Modeling the association between home care service use and entry into residential aged care: a cohort study using routinely collected data,» *Journal of the American medical directors association,* vol. 19, n. 2, pp. 117-121, 2018.

[8] D. M. Guthrie, J. G. Davidson, N. Williams, J. Campos, K. Hunter, P. Mick e W. Wittich, «Combined impairments in vision, hearing and cognition are associated with greater levels of functional and communication difficulties than cognitive impairment alone: Analysis of interRAI data for home care and long-term care recipients in Ontario.,» *PloS one,* vol. 2, n. e0192971, p. 13, 2018.

[9] F. W. Udsen, P. H. Lilholt, O. Hejlesen e L. Ehlers, «Cost-effectiveness of telehealthcare to patients with chronic obstructive pulmonary disease: results from the Danish 'TeleCare North'cluster-randomised trial,» *BMJ open,* vol. 5, n. e014616, p. 7, 2017.

[10] L. Mondor, C. J. Maxwell, D. B. Hogan, S. E. Bronskill, M. A. Campitelli, D. P. Seitz e W. P. Wodchis, «The incremental health care costs of frailty among home care recipients with and without dementia in Ontario, Canada,» *Medical care,* vol. 57, n. 7, 2019.

[11] T. Vandiver, T. Anderson, B. Boston, C. Bowers e N. Hall, «Community-based home health programs and chronic disease.,» *Professional Case Management,* vol. 23, n. 1, pp. 25-31, 2018.

[12] V. G. Valdmanis, M. D. Rosko, H. Leleu e D. B. Mukamel, «Assessing overall, technical, and scale efficiency among home health care agencies,» *Health care management science,* vol. 20, pp. 265-275, 2017.

[13] S. A. Suikkanen, P. K. Soukkio, E. M. Aartolahti, H. Kautiainen, S. M. Kääriä, M. T. Hupli e K. T. Kukkonen-Harjula, «Effects of home-based physical exercise on days at home and cost-effectiveness in pre-frail and frail persons: randomized controlled trial.,» *Journal of the American medical directors association, ,* vol. 22, n. 4, pp. 773-779, 2021.

[14] K. Steen Carlsson, G. Andsberg, J. Petersson e B. Norrving, «Long-term cost-effectiveness of thrombectomy for acute ischaemic stroke in real life: An analysis based on data from the Swedish Stroke Register (Riksstroke),» *International Journal of Stroke,* vol. 12, n. 8, pp. 802-814, 2017.

[15] B. Michalowsky, F. Xie, T. Eichler, J. Hertel, A. Kaczynski, I. Kilimann e W. Hoffmann, «Cost-effectiveness of a collaborative dementia care management—results of a cluster-randomized controlled trial,» *Alzheimer's & Dementia,* vol. 15(10), n. 129, 2019.

[16] M. R. Alam, S. Karmokar, S. Reza, M. R. Kabir, S. Ghosh e M. A. Al Mamun, «Geriatric malnutrition and depression: Evidence from elderly home care population in Bangladesh,» *Preventive medicine reports,* vol. 101478, n. 23, 2021.

[17] M. Lee, J. H. Ryoo, C. Campbell, P. J. Hollen e I. C. Williams, «Exploring the challenges of medical/nursing tasks in home care experienced by caregivers of older adults with dementia: An integrative review,» *Journal of clinical nursing,* vol. 28, n. 23-24, pp. 4177-4189, 2019.

9) Appendix

Variables of the Econometric Model			
Variables	**Acronym**	**Label**	**Average**
Elderly people treated in integrated home care	EPIHC	A108	
Use of libraries	UL	A10	-0,025
Employment rate (20-64 years)	ER	A11	0,035
Satisfaction with the work done	SWD	A20	-0,039
Economic situation of the family	ESF	A32	0,022
Trust in the police and firefighters	TPF	A46	-0,671
Women and political representation at local level	WPRL	A48	0,021
Positive opinion on future prospects	POF	A66	-0,032
Density and relevance of museum heritage	DRMH	A69	-0,099
Concern about landscape deterioration	CLD	A78	-0,039
Wastewater Treatment	WT	A85	0,004
Soil waterproofing from artificial cover	SWAC	A89	0,149
Concern about climate change	CCC	A94	0,070

Innovation of the production system	IPS	A99	-0,004
Nurses and midwives	NM	A122	0,353

List of Variables of the Econometric Model			
Variable	**Acronym**	**Definition**	**Source**
Elderly people treated in integrated home care	EPIHC	Elderly people treated in integrated home care: Percentage of elderly people treated in home care integrated into the total elderly population (65 years and over) resident.	Istat, Elaborazione su dati Ministero della Salute,
Use of libraries	UL	Percentage of people aged 3 years and older who went to the library at least once in the 12 months preceding the interview out of the total number of people aged 3 years and older.	Istat, Indagine Aspetti della vita quotidiana.
Employment rate (20-64 years)	ER	Percentage of employed people aged 20-64 in the population aged 20-64.	Istat, Indagine Aspetti della vita quotidiana.
Satisfaction with the work done	SWD	Percentage of employed people who expressed an average score of satisfaction between 8 and 10 for the following aspects of the work performed: earnings, career opportunities, number of hours worked, job stability, distance from home to work, interest in the job.	Istat, Rilevazione sulle Forze di lavoro.
Economic situation of the family	ESF	Families who declare their economic situation has worsened or significantly worsened compared to the previous year.	Istat, Aspetti della vita quotidiana
Trust in the police and firefighters	TPF	Average score of trust in the police and firefighters (on a scale from 0 to 10) expressed by people aged 14 and over.	Istat, Indagine Aspetti della vita quotidiana.
Women and political representation at local level	WPRL	Percentage of women elected to Boards regional out of the total elected.	Singoli Consigli regionali.
Positive opinion on future prospects	POF	Percentage of people aged 14 and over who believe that their personal situation will improve in the next few years 5 years on the total of people aged 14 and over.	Istat, Indagine Aspetti della vita quotidiana
Density and relevance of museum heritage	DRMH	Number of permanent exhibition structures per 100 km2 (museums, archaeological areas and monuments open to the public), weighted by the number of visitors. The weight of each structure is assumed to be equal to (Vi / VM), where Vi is the number of visitors to the structure, M the total structures and V the total visitors.	Istat, Indagine sui musei e le istituzioni similari.
Concern about landscape deterioration	CLD	Percentage of people aged 14 and over indicate the ruin of the landscape caused by excessive building construction among the five problems most worrying environmental issues among the total number of people aged 14 and over.	Istat, Indagine sui musei e le istituzioni similari.
Wastewater Treatment	WT	Percentage share of polluting loads flowing into secondary plants or advanced, in equivalent inhabitants, compared to the loads urban totals (Aetu) generated.	Istat, Censimento delle acque per uso civile.
Soil waterproofing from artificial cover	SWAC	Percentage of waterproofed soil on total land area.	ISPRA, Consumo di suolo, dinamiche territoriali e servizi
Concern about climate change	CCC	Percentage of people aged 14 and over indicate the ruin of the landscape caused by excessive building construction among the five problems most worrying environmental issues among the total number of people aged 14 and over.	Istat, Indagine Aspetti della vita quotidiana.
Innovation of the production system	IPS	Percentage of companies that have introduced product innovations and process in the three-year reference period out of the total of companies with at least 10 employees.	Istat, Cis (Community Innovation Survey)
Nurses and midwives	NM	Number of nurses and midwives per 1,000 inhabitants.	Co.Ge.A.P.S. - Banca dati Nazionale dei crediti ECM

Pooled OLS, using 372 observations

Included 20 cross-sectional units

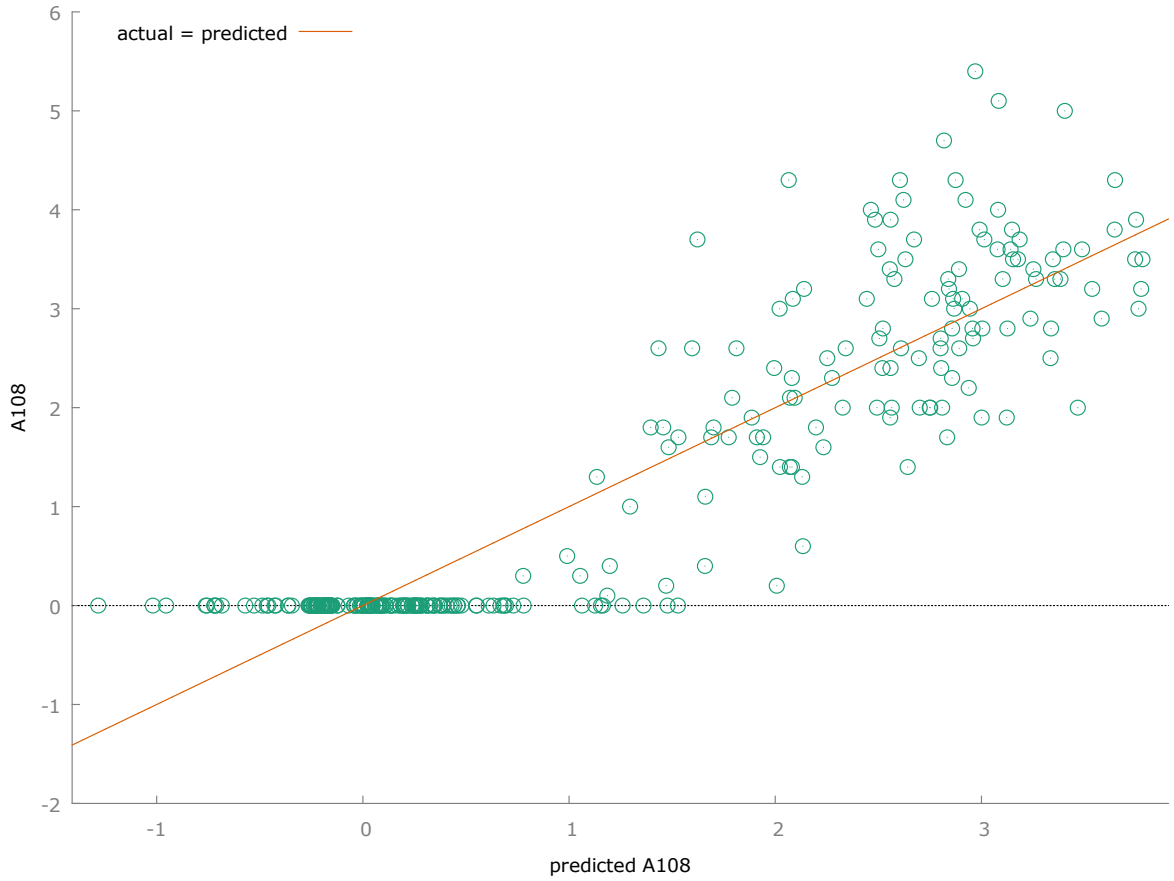
Time-series length: minimum 13, maximum 19

Dependent variable: A108

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	0.0244965	0.0557866	0.4391	0.6608	
A10	−0.0315480	0.00859316	−3.671	0.0003	***

A11	0.0396129	0.00838976	4.722	<0.0001	***
A20	−0.0451382	0.0113794	−3.967	<0.0001	***
A32	0.0215936	0.00399532	5.405	<0.0001	***
A46	−0.725750	0.122226	−5.938	<0.0001	***
A48	0.0221439	0.00544034	4.070	<0.0001	***
A66	−0.0436165	0.0141767	−3.077	0.0023	***
A69	−0.120287	0.0390302	−3.082	0.0022	***
A78	−0.0497352	0.0112356	−4.427	<0.0001	***
A85	0.00406651	0.00158730	2.562	0.0108	**
A89	0.143212	0.0148050	9.673	<0.0001	***
A94	0.0820359	0.0132757	6.179	<0.0001	***
A99	−0.00476756	0.00137400	−3.470	0.0006	***
A122	0.371931	0.0205353	18.11	<0.0001	***

Mean dependent var	0.917204	S.D. dependent var	1.415595
Sum squared resid	119.3045	S.E. of regression	0.578089
R-squared	0.839526	Adjusted R-squared	0.833233
F(14, 357)	133.4041	P-value(F)	2.6e-132
Log-likelihood	−316.3232	Akaike criterion	662.6464
Schwarz criterion	721.4298	Hannan-Quinn	685.9909
rho	0.500194	Durbin-Watson	0.985681



Fixed-effects, using 372 observations

Included 20 cross-sectional units

Time-series length: minimum 13, maximum 19

Dependent variable: A108

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	0.00860771	0.0523767	0.1643	0.8696	
A10	−0.0249888	0.00847119	−2.950	0.0034	***
A11	0.0317663	0.00849299	3.740	0.0002	***
A20	−0.0355933	0.0114837	−3.099	0.0021	***
A32	0.0224608	0.00380752	5.899	<0.0001	***
A46	−0.656708	0.119754	−5.484	<0.0001	***
A48	0.0200580	0.00615583	3.258	0.0012	***
A66	−0.0279420	0.0154784	−1.805	0.0719	*
A69	−0.0883874	0.0412554	−2.142	0.0329	**
A78	−0.0385103	0.0119545	−3.221	0.0014	***
A85	0.00418076	0.00151435	2.761	0.0061	***

A89	0.150844	0.0163590	9.221	<0.0001	***
A94	0.0669082	0.0129769	5.156	<0.0001	***
A99	−0.00414093	0.00129964	−3.186	0.0016	***
A122	0.345718	0.0215871	16.02	<0.0001	***

Mean dependent var	0.917204	S.D. dependent var	1.415595
Sum squared resid	97.28106	S.E. of regression	0.536483
LSDV R-squared	0.869149	Within R-squared	0.857379
LSDV F(33, 338)	68.03313	P-value(F)	8.6e-129
Log-likelihood	−278.3653	Akaike criterion	624.7306
Schwarz criterion	757.9730	Hannan-Quinn	677.6447
rho	0.432361	Durbin-Watson	1.091992

Joint test on named regressors -

Test statistic: F(14, 338) = 145.137

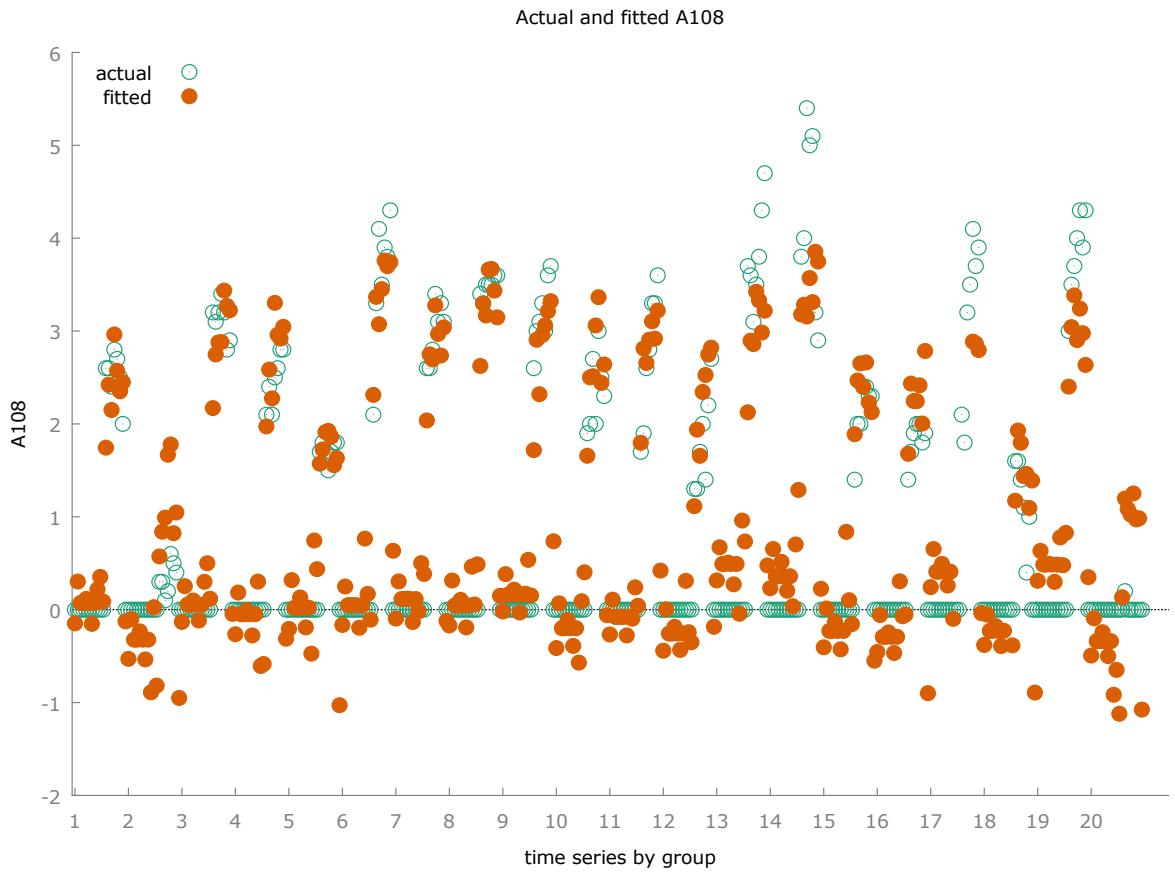
with p-value = P(F(14, 338) > 145.137) = 1.66665e-133

Test for differing group intercepts -

Null hypothesis: The groups have a common intercept

Test statistic: F(19, 338) = 4.02736

with p-value = P(F(19, 338) > 4.02736) = 7.10491e-08



Random-effects (GLS), using 372 observations
Using Nerlove's transformation
Included 20 cross-sectional units
Time-series length: minimum 13, maximum 19
Dependent variable: A108

	Coefficient	*Std. Error*	*z*	*p-value*	
const	0.0160898	0.0808156	0.1991	0.8422	
A10	−0.0265731	0.00830555	−3.199	0.0014	***
A11	0.0332317	0.00829591	4.006	<0.0001	***
A20	−0.0373123	0.0112211	−3.325	0.0009	***
A32	0.0223572	0.00376007	5.946	<0.0001	***
A46	−0.669215	0.117706	−5.685	<0.0001	***
A48	0.0203903	0.00586786	3.475	0.0005	***
A66	−0.0315292	0.0148971	−2.116	0.0343	**
A69	−0.0945432	0.0399832	−2.365	0.0181	**
A78	−0.0414042	0.0115258	−3.592	0.0003	***
A85	0.00418397	0.00149524	2.798	0.0051	***

A89	0.148601	0.0156737	9.481	<0.0001	***
A94	0.0700581	0.0127540	5.493	<0.0001	***
A99	−0.00429741	0.00128490	−3.345	0.0008	***
A122	0.352147	0.0208599	16.88	<0.0001	***

Mean dependent var	0.917204	S.D. dependent var	1.415595
Sum squared resid	120.8250	S.E. of regression	0.580948
Log-likelihood	−318.6787	Akaike criterion	667.3575
Schwarz criterion	726.1409	Hannan-Quinn	690.7020
rho	0.432361	Durbin-Watson	1.091992

'Between' variance = 0.070923

'Within' variance = 0.261508

mean theta = 0.59256

Joint test on named regressors -

Asymptotic test statistic: Chi-square(14) = 2081.96

with p-value = 0

Breusch-Pagan test -

Null hypothesis: Variance of the unit-specific error = 0

Asymptotic test statistic: Chi-square(1) = 47.3276

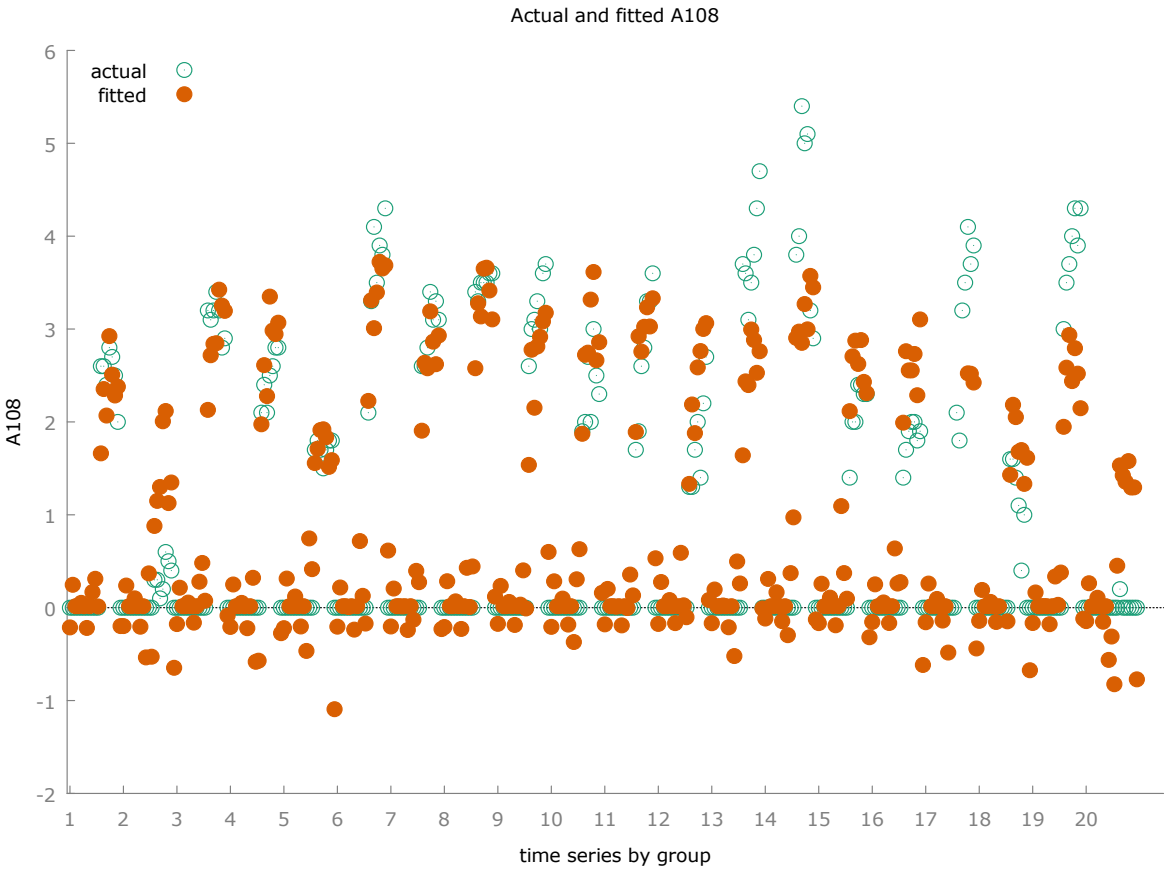
with p-value = 6.00607e-12

Hausman test -

Null hypothesis: GLS estimates are consistent

Asymptotic test statistic: Chi-square(14) = 8.82869

with p-value = 0.841865



WLS, using 372 observations					
Included 20 cross-sectional units					
Dependent variable: A108					
Weights based on per-unit error variances					
	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	0.00224777	0.0425337	0.05285	0.9579	
A10	−0.0155677	0.00662838	−2.349	0.0194	**
A11	0.0341924	0.00705658	4.845	<0.0001	***
A20	−0.0394755	0.0100755	−3.918	0.0001	***
A32	0.0220288	0.00324028	6.798	<0.0001	***
A46	−0.632843	0.0988149	−6.404	<0.0001	***
A48	0.0199823	0.00405249	4.931	<0.0001	***
A66	−0.0261940	0.0114618	−2.285	0.0229	**
A69	−0.0924719	0.0292089	−3.166	0.0017	***
A78	−0.0277423	0.00905178	−3.065	0.0023	***
A85	0.00447133	0.00121453	3.682	0.0003	***
A89	0.153353	0.0113473	13.51	<0.0001	***
A94	0.0605325	0.0112459	5.383	<0.0001	***
A99	−0.00367996	0.000998693	−3.685	0.0003	***

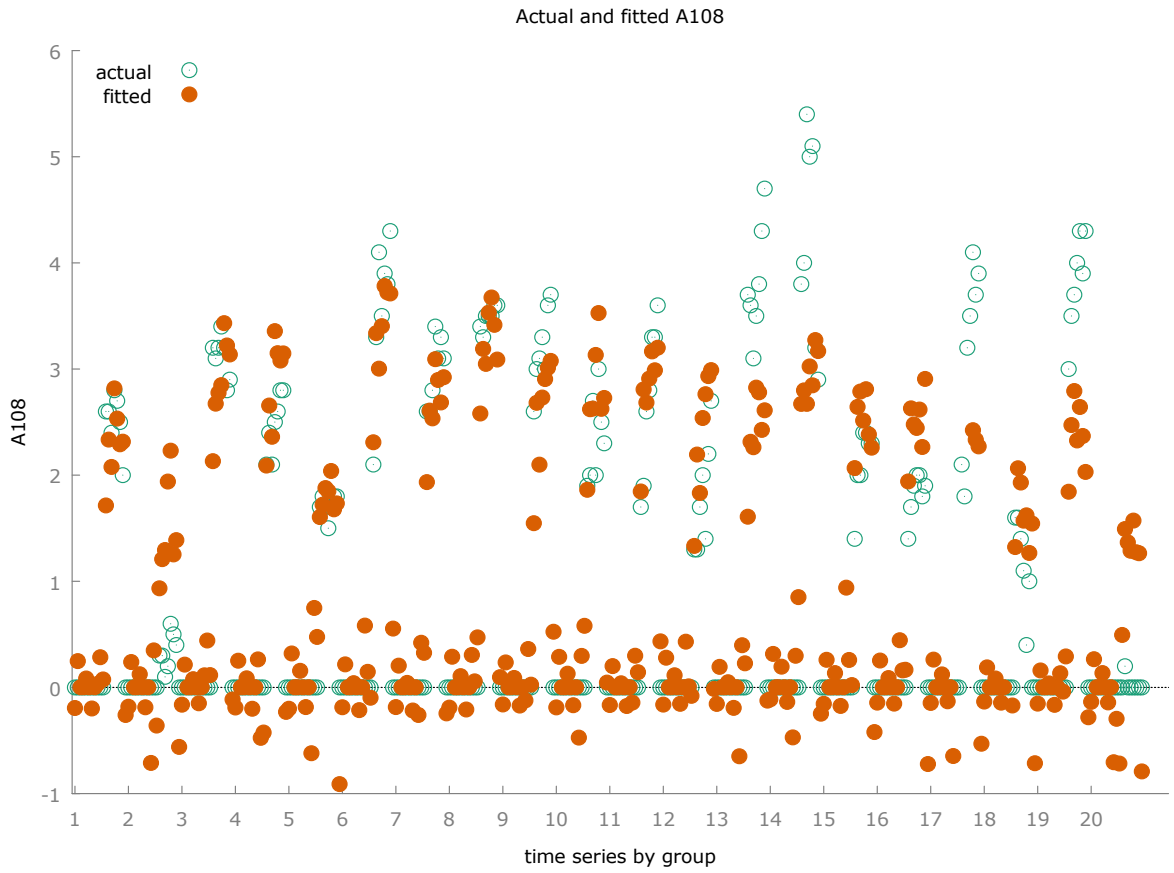
A122	0.341707	0.0175095	19.52	<0.0001	***

Statistics based on the weighted data:

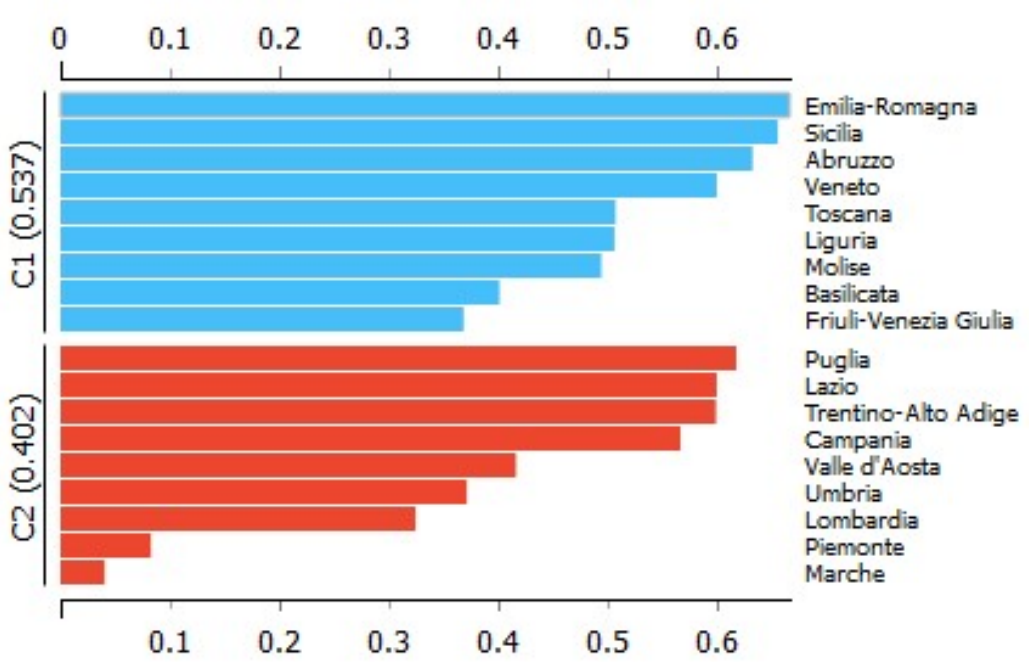
Sum squared resid	349.0734	S.E. of regression	0.988836
R-squared	0.904421	Adjusted R-squared	0.900673
F(14, 357)	241.2954	P-value(F)	2.7e-172
Log-likelihood	−516.0134	Akaike criterion	1062.027
Schwarz criterion	1120.810	Hannan-Quinn	1085.371

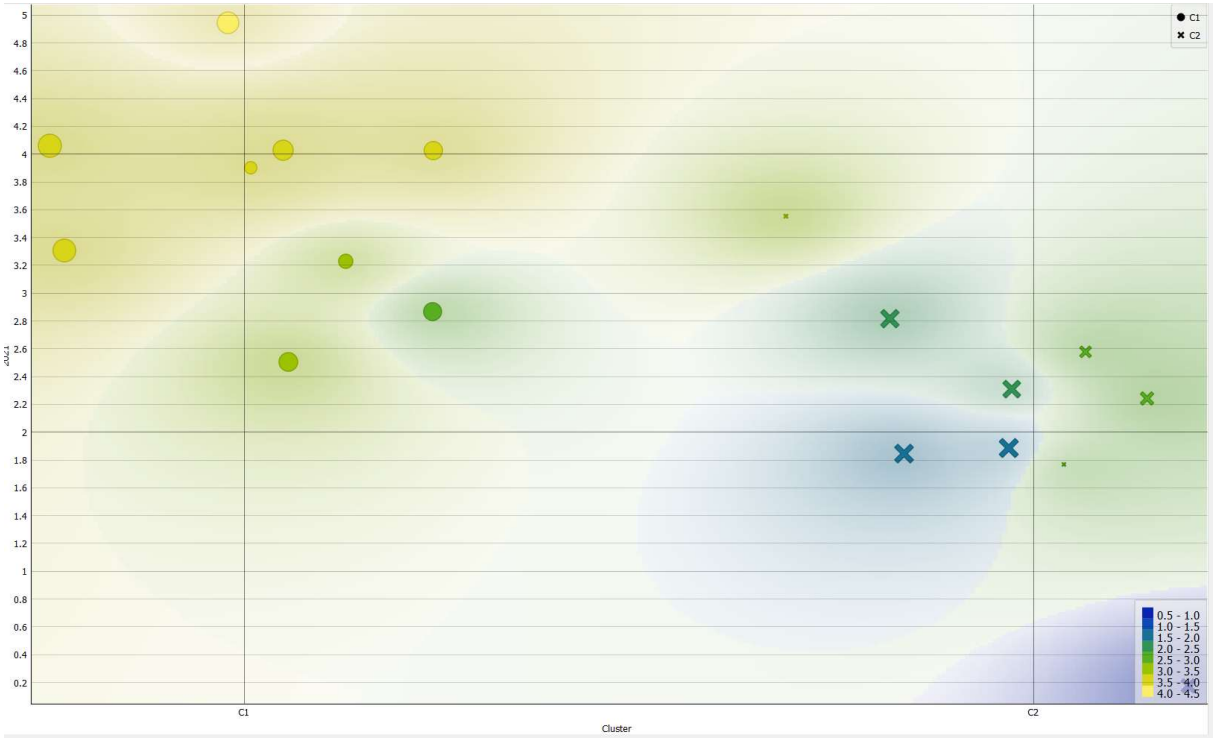
Statistics based on the original data:

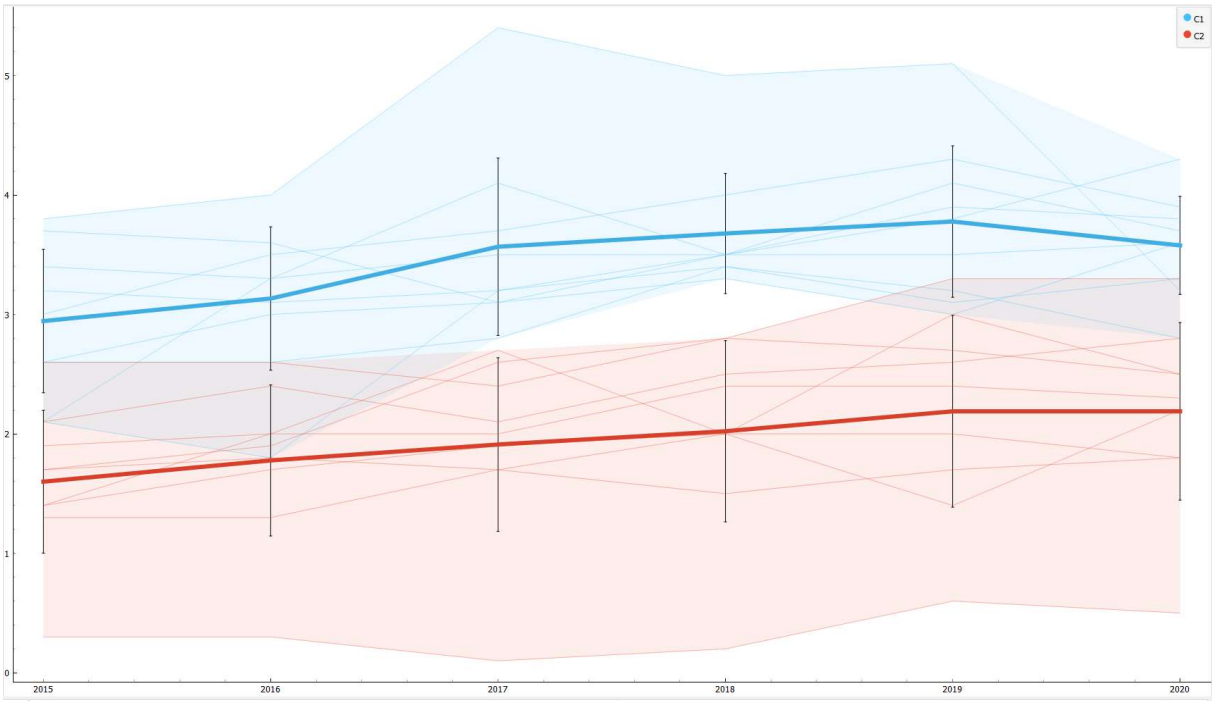
Mean dependent var	0.917204	S.D. dependent var	1.415595
Sum squared resid	124.9842	S.E. of regression	0.591689

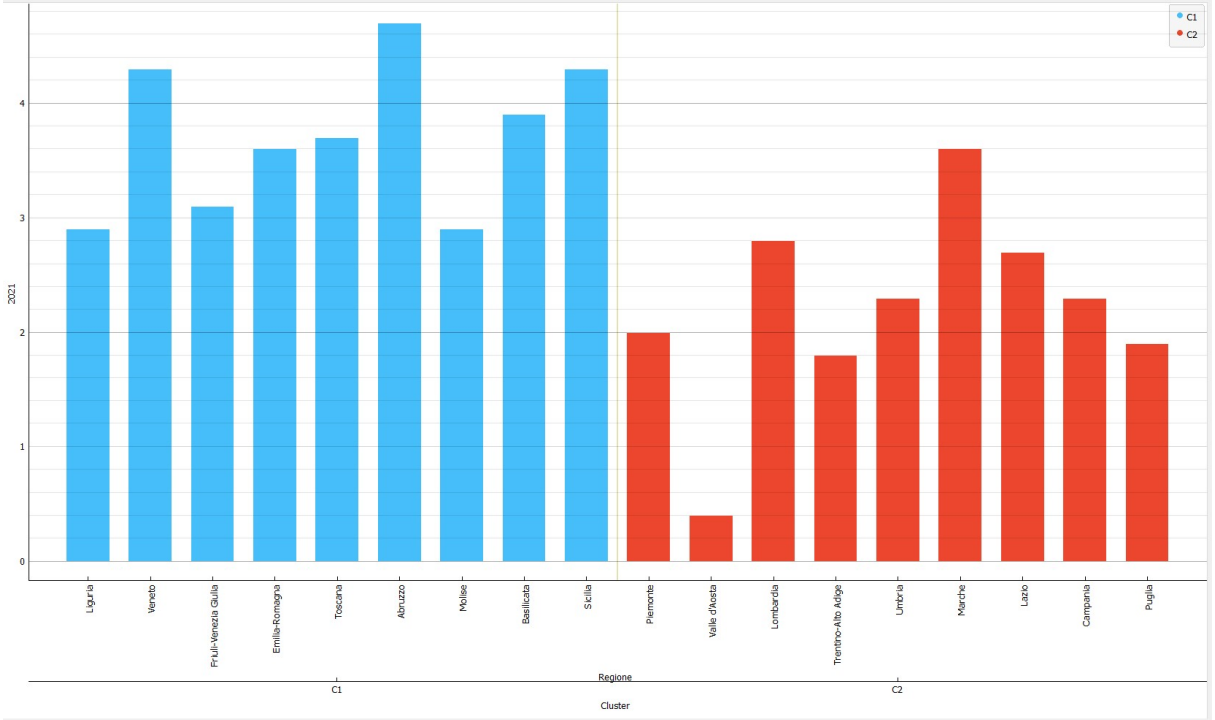


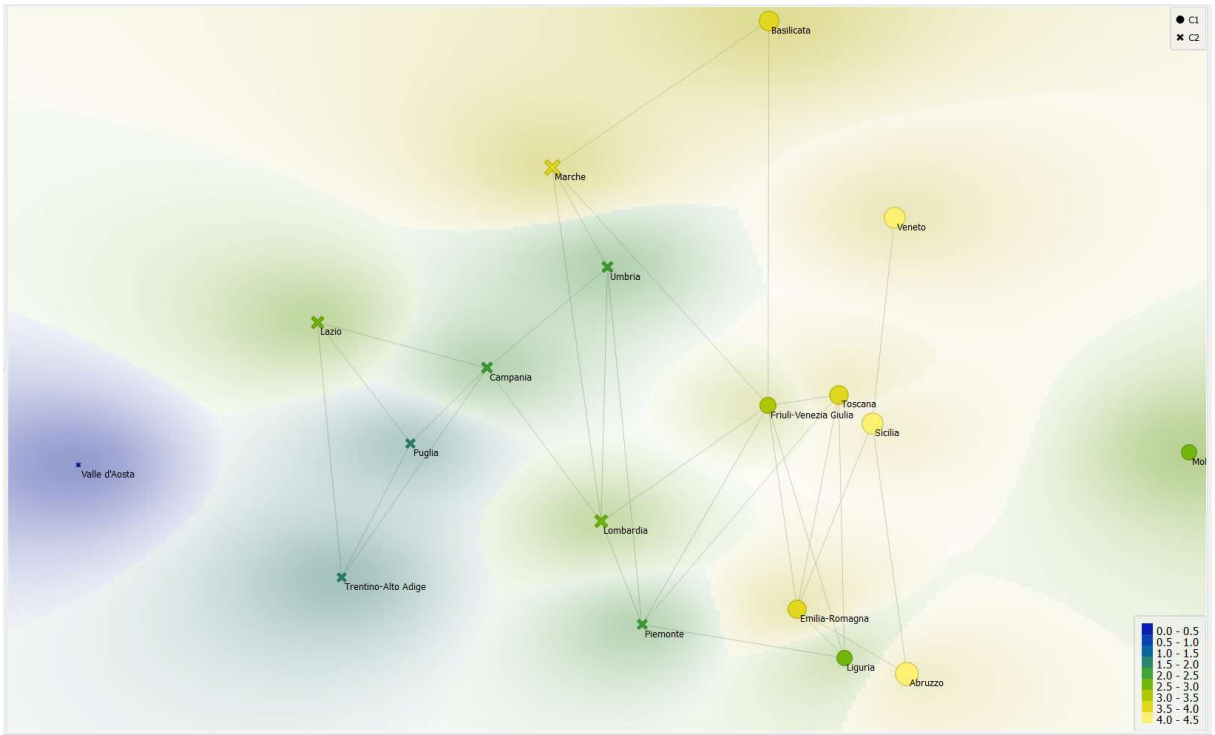
Clusterization with k-Means algorithm


0
0.1
0.2
0.3
0.4
0.5
0.6
C1 (0.537)
Emilia-Romagna
Sicilia
Abruzzo
Veneto
Toscana
Liguria
Molise
Basilicata
Friuli-Venezia Giulia
C2 (0.402)
Puglia
Lazio
Trentino-Alto Adige
Campania
Valle d'Aosta
Umbria
Lombardia
Piemonte
Marche

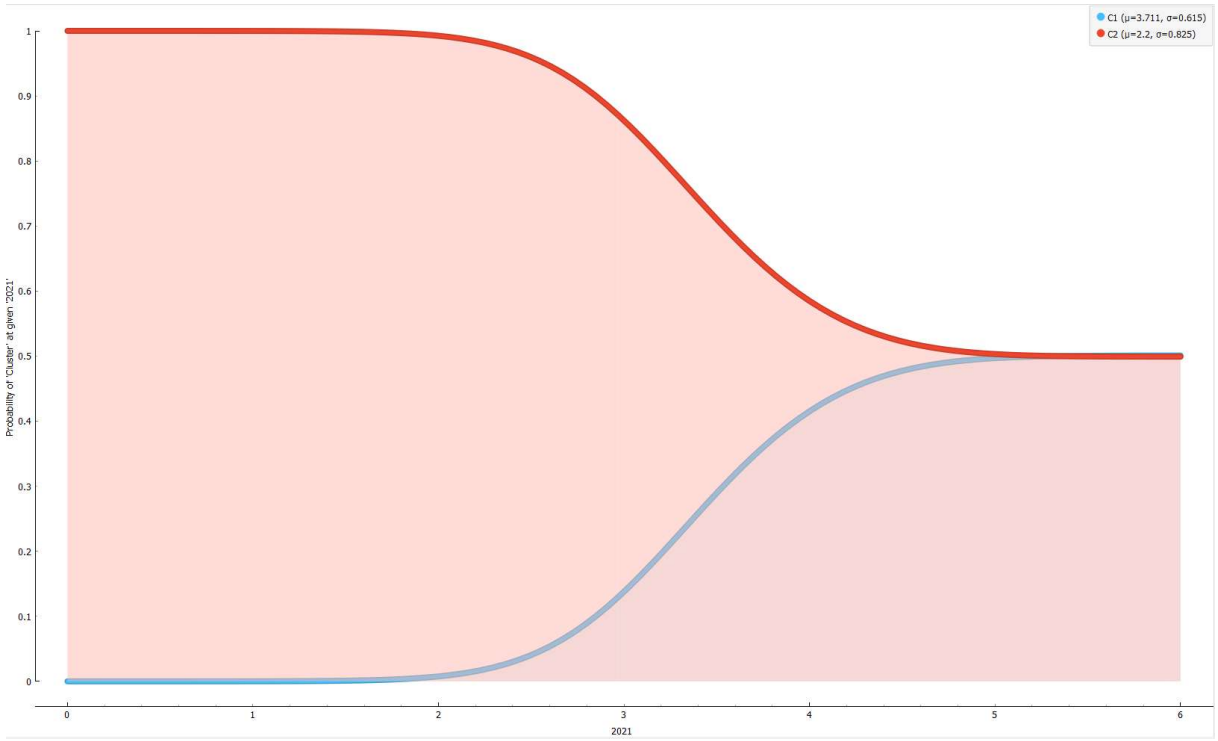

C1
C2
Cluster
0.5 - 1.0
1.0 - 1.5
1.5 - 2.0
2.0 - 2.5
2.5 - 3.0
3.0 - 3.5
3.5 - 4.0
4.0 - 4.5


C1
C2
5
4
3
2
1
0
2015
2016
2017
2018
2019
2020


C1
C2
2021
4
3
2
1
0
Liguria
Veneto
Friuli-Venezia Giulia
Emilia-Romagna
Toscana
Abruzzo
Molise
Basilicata
Sicilia
Piemonte
Valle d'Aosta
Lombardia
Trentino-Alto Adige
Umbria
Marche
Lazio
Campania
Puglia
Regione
C1
C2
Cluster



	2021	Regione	Cluster	Silhouette	2015	2016	2017	2018	2019	2020
3	2.9	Liguria	C1	0.631426	3.2	3.1	3.2	3.4	3.2	2.8
6	4.3	Veneto	C1	0.652672	2.1	3.3	4.1	3.5	3.9	3.8
7	3.1	Friuli-Venezia G...	C1	0.601924	2.6	2.6	2.8	3.4	3.1	3.3
8	3.6	Emilia-Romagna	C1	0.672746	3.4	3.3	3.5	3.5	3.5	3.6
9	3.7	Toscana	C1	0.635875	2.6	3.0	3.1	3.3	3.0	3.6
13	4.7	Abruzzo	C1	0.663037	3.7	3.6	3.1	3.5	3.8	4.3
14	2.9	Molise	C1	0.64053	3.8	4.0	5.4	5.0	5.1	3.2
17	3.9	Basilicata	C1	0.584818	2.1	1.8	3.2	3.5	4.1	3.7
18	4.3	Sicilia	C1	0.676589	3.0	3.5	3.7	4.0	4.3	3.9
1	2.0	Piemonte	C2	0.520843	2.6	2.6	2.4	2.8	2.7	2.5
2	0.4	Valle d'Aosta	C2	0.624995	0.3	0.3	0.1	0.2	0.6	0.5
4	2.8	Lombardia	C2	0.597204	2.1	2.4	2.1	2.5	2.6	2.8
5	1.8	Trentino-Alto A...	C2	0.666721	1.7	1.8	1.7	1.5	1.7	1.8
10	2.3	Umbria	C2	0.614973	1.9	2.0	2.7	2.0	3.0	2.5
11	3.6	Marche	C2	0.523723	1.7	1.9	2.6	2.8	3.3	3.3
12	2.7	Lazio	C2	0.664558	1.3	1.3	1.7	2.0	1.4	2.2
15	2.3	Campania	C2	0.656794	1.4	2.0	2.0	2.4	2.4	2.3
16	1.9	Puglia	C2	0.670646	1.4	1.7	1.9	2.0	2.0	1.8


C1 (μ=3.711, σ=0.615)
C2 (μ=2.2, σ=0.825)
Probability of 'Cluster' at given '2021'
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
2021