ICT Specialists in Europe

Leogrande, Angelo and Magaletti, Nicola and Cosoli, Gabriele and Giardinelli, Vito and Massaro, Alessandro

LUM UNIVERSITY GIUSEPPE DEGENNARO-LUM ENTERPRISE S.R.L., LUM UNIVERSITY GIUSEPPE DEGENNARO-LUM ENTERPRISE S.R.L., LUM UNIVERSITY GIUSEPPE DEGENNARO-LUM ENTERPRISE S.R.L., LUM UNIVERSITY GIUSEPPE DEGENNARO-LUM ENTERPRISE S.R.L., LUM UNIVERSITY GIUSEPPE DEGENNARO-LUM ENTERPRISE S.R.L.

5 March 2022

Angelo Leogrande[1], Nicola Magaletti[2], Gabriele Cosoli[3], Vito O.M. Giardinelli[4], Alessandro Massaro[5]

ICT Specialists in Europe

Abstract

The following article estimates the value of ICT Specialists in Europe between 2016 and 2021 for 28 European countries. The data were analyzed using the following econometric techniques, namely: Panel Data with Fixed Effects, Panel Data with Random Effects, WLS and Pooled OLS. The results show that the value of ICT Specialists in Europe is positively associated with the following variables: "*Desi Index*", "*SMEs with at least a basic level of digital intensity*", "*At least 100 Mbps fixed BB take-up*" and negatively associated with the following variables: "*4G Coverage*","*5G Coverage*", "*5G Readiness*", "*Fixed broadband coverag*e", "*e-Government*", "*At least Basic Digital Skills*", "*Fixed broadband take-up*", "*Broadband price index*", "*Integration of Digital Technology*". Subsequently, two European clusters were found by value of "ICTG Specialists" using the k-Means clustering algorithm optimized by using the Silhouette coefficient. Finally, eight different machine learning algorithms were compared to predict the value of "*ICT Specialists*" in Europe. The results show that the best prediction algorithm is ANN-Artificial Neural Network with an estimated growth value of 12.53%. Finally, "*augmented data*" were obtained through the use of the ANN-Artificial Neural Network, through which a new prediction was made which estimated a growing value of the estimated variable equal to 3.18%.

Keywords: Innovation, and Invention: Processes and Incentives; Management of Technological Innovation and R&D; Diffusion Processes; Open Innovation.

JEL Classification: O30; O31, O32; O33; O36.

1. Introduction

In the following article an analysis of the determinants of the ICT sector in Europe has been carried out. In particular, the data of the DESI Index database were analyzed with reference to 28 countries in the period 2014-2021. The data were analyzed using a set of econometric techniques, i.e., panel data, and also machine learning techniques for prediction. The presence of ICT specialists is very important to ensure the development of an evolved digitization system at country level. An analysis of the reference literature is provided below.

[1] Assistant Professor at Lum University Giuseppe Degennaro and Researcher at Lum Enterprise s.r.l. Email: leogrande.cultore@lum.it. Strada Statale 100 km 18, 70010 Casamassima BA, Puglia, Italy, European Union.

[2] Chief Operation Officer-COO and Senior Researcher at Lum Enterprise s.r.l. Email: magaletti@lumenterprise.it. Strada Statale 100 km 18, 70010 Casamassima BA, Puglia, Italy, European Union.

[3] Senior IT Specialist and Solutions Architects and Researcher at Lum Enterprise s.r.l. Email: cosoli@lumenterprise.it. Strada Statale 100 km 18, 70010 Casamassima BA, Puglia, Italy, European Union.

[4] Business Developer and Researcher at Lum Enterprise s.r.l. Email: Giardinelli@lumenterprise.it Strada Statale 100 km 18, 70010 Casamassima BA, Puglia, Italy, European Union.

[5] Professor at Lum University Giuseppe Degennaro, and Chief Research Officer-CRO at Lum Enterprise s.r.l. Email: massaro@lum.it. Strada Statale 100 km 18, 70010 Casamassima BA, Puglia, Italy, European Union.

[1] address the issue of the relationship between digital skills, the labor market, and the structure of productivity in European countries in relation above all to Covid. The authors emphasize that the need for digitalization that has been produced by Covid 19 has met many limits due to the lack of an adequate workforce prepared to offer the services and products required by the health emergency. In particular, the use of a set of new technologies such as artificial intelligence, big data requires an increasingly experienced workforce in the digitalization sector. It therefore follows that to create the necessary workforce to support the need for digitization of the European Union, it is necessary to invest in training and in the requalification of human capital in Europe. [2] refer to the need to introduce transversal skills in the training of ICT specialists. The growth of transversal skills is necessary to increase the competitiveness of small and medium-sized enterprises operating in the ICT sector. ICT specialists need technical skills and soft skills. Soft skills include creativity, teamwork, and personal skills. It therefore follows that ICT experts must be equipped with a broad set of skills that are both technical and human. The authors therefore believe that for ICT specialists to be able to establish themselves in the labor market, they are also able to exercise soft skills. However, to effectively develop transversal skills, the company must have at its disposal a set of tools such as: the e-learning platform, workshops, tutoring, coaching. [3] analyze the case of the relationship between ICT technologies and the job market in Russia. The authors point out that there are delays in the development of ICT technologies in Russia. Therefore, the authors believe that a growth in skills in the IT sector can be a useful tool to allow companies to find human capital with hard skills capable of enabling digitization. [4] takes into consideration the relationship between the level of digitization of countries and the degree of productivity. In particular, the authors analyze the case of the Romanian economy with respect to the average of the European Union. The authors show that Romania has a certain gap vis-à-vis the EU and that this gap can effectively be bridged through a more significant investment in digital capabilities that can lead to productivity growth. [5] highlight the existence of the positive connection between the teaching of ICT disciplines in schools and the use of the same technologies in the workplace. The authors verified through the analysis of a sample of 122 teachers that the improvement of school education in relation to ICT disciplines also has a positive impact in terms of the adoption of the same technologies in the workplace. It follows therefore that the improvement of pedagogical-educational methodologies connected to ICT disciplines is a predictor of the future ability to use information technologies in the workplace. [6] underline the value of training and growth of professionals in the ICT sector as an essential element to solve the digital divide in Europe. [7] refer to the use of ICT technologies within the cultural production industry. In this regard, the authors underline the role of ICT specialists in increasing the possibility of enhancing cultural heritage also through digitization initiatives.

[8] highlight the difficulty for companies, especially small and medium-sized ones, to hire skilled workers in the IT disciplines. In fact, the demand for work in the ICT sector is very large, even if due to this demand there is a lack of adequate professional profiles. In this regard, it would be necessary for companies to enter into agreements with universities to increase ICT training to have an impact on the skilled workforce. [9] analyze the relationship between digitalization and entrepreneurship in Europe. The authors underline the need for the European Union to invest significantly in the formation of human capital in the ICT field to increase the supply capacity of digital products and services. In particular, the authors underline the need for the European Union to invest significantly in start-ups to increase the IT content of employees' professional skills. [10] analyzes the case of Romania and its ability to attract international companies operating in the information communication technology sector. The international opening of Russia has led to the opening of about 15,000 companies operating in the ICT sector with a growing demand for operators in the IT sector. In this regard, the authors propose investments in the education sector to train a new workforce suitable to meet the demand of ICT

companies. The article therefore allows us to bring together in a unitary vision the question of industrial economic policies aimed at the creation of new information companies with economic policies aimed at education to create the foundations of a new workforce technically specialized in ICT disciplines.

[11] considers the role of the IT skills of the workforce as an essential component in the process of localization of industries within the local economy. In fact, one of the variables that companies must consider in determining those that support investment choices is the selection of areas with human capital endowed with the necessary professional skills. It follows therefore that the development of high IT skills in a certain population can allow to attract more investments in the ICT sector. [12] tackles the issue of the development of the digital economy emphasizing that the element of human capital formation trained in ICT disciplines is an essential component for the development of an economic system based on new information technologies. [13] present a comparison between the degree of digitization of Serbia compared to the level of the European Union. The authors verify that Serbia has high levels compared to the European average about the degree of use of new technologies within companies. Furthermore, the authors verify that in Serbia there is an excess of demand for jobs in the ICT sector which are not adequately answered by the local job offer. However, it should be noted that Serbia has a higher level of employees in the ICT sector in companies than the European average. This condition highlights the fact that the dynamism of the technological sector in Serbia could find greater capacity for development and production of added value if the labor market is more ready to provide skilled workers in the IT disciplines.

[14] addresses the issue of the development of the technological skills that are necessary to create the conditions for the growth of smart cities. The authors verify that some elements, such as the presence of experts in the ICT sector, are an essential component in the affirmation process of smart cities. The analysis was carried out on about 26 European countries and the results show that for the creation of smart cities it is necessary to invest in a series of factors, among which those relating to the presence and training of ICT personnel stand out. [15] analyze the case of the vocational training of ICT specialists in Europe. The data show that in most of the countries considered ICT specialists have a tertiary education qualification, even if there are some countries such as Italy, Portugal, and Germany where ICT experts do not have tertiary education qualifications. This lack of training in the field of ICT specialists in some countries can have an impact in terms of the ability to develop the entire IT sector from an industrial point of view, even if it may not impact on the overall degree of digitalization of the country from the point of view of view of networks and infrastructures. [16] addresses the issue of the development of gender equality in the STEM disciplines or Science, Technology, Engineering and Mathematics in Latvia. The analysis shows how it would be necessary to introduce changes in the system of training and the integration of women into the labor market to reduce the gender difference and create the conditions for a growth in general employment in the ICT sector.

[17] considers the limitations that come to be produced in terms of the gender gap between men and women in India in internet access. The authors verify that there are indeed distinctions of internet access that are based on gender differences and that these differences can also impact the training of specialist operators in ICT disciplines. Social rules therefore have a very negative impact in promoting access to the use of digital tools by women who in some cases even limit the use of the internet and mobile phones. It therefore follows that to strengthen the presence of workers operating in the ICT sector in India it is necessary to remove the digital divide and promote wider access to information technology by women. [18] addresses the issue of organizing ICT workers by emphasizing how these types of workers are poorly unionized. The authors highlight that due to the typical characteristics of workers in the ICT sector, it appears that the degree of autonomy in carrying out the work activity is very high and that for this reason the degree of unionization is low. [19] refer to the impact of digital skills for socio-economic

development in Russia. The authors highlight the need to connect the demand for work to the university and vocational training of Russian workers as a strategy to increase the digitization of the labor market. [20] consider the role of ICT specialists in promoting a digitization-oriented labor market in Russia. The possibility for ICT workers to participate in the wide phenomenon of digitization essentially depends on the complex of training tools made available to Russian students and workers through the university system and through training within companies.

Furthermore, it should also be considered that the presence of ICT specialists can be significantly associated with broadband penetration [21], fixed broadband [22], the price of broadband [23] ,e-government [24], and the technologies used in production lines [25].

The article continues as follows: the second paragraph presents an analysis of the correlation matrix, the third paragraph contains an analysis of econometric models, the fourth paragraph refers to predictive analysis, the fifth paragraph analyzes the prediction with use of augmented data, the sixth paragraph concludes. Finally, the appendix contains a summary of the results obtained using three different software, namely: Gretl, Orange and KNIME.

2. Correlation Matrix

An analysis was also carried out using the correlation matrix to analyze the relationships between the variables of the model. However, only the particularly significant values in terms of value were considered, i.e., those higher than a correlation value of 0.7 and the negative ones close to zero. It appears that the value of the *ICT Specialists* variable is positively correlated to the following variables, that is:

- *SMEs with at least a basic level of digital intensity* equal to an amount of 0.8228;
- *Desi Index Aggregate Score* with a correlation index value equal to a value of 0.798 units;
- *At Least Basic Digital Skills* with a correlation index value equal to a value of 0.745;
- *Integration of Digital Technology* with a correlation index value equal to a value of 0.737;
- *e-Government* with a correlation index value equal to an amount of 0.70;
- *Broadband Price Index* with a correlation index value equal to a value of -0.03.

An analysis of the correlation matrix was then carried out which considers the *Fixed Broadband Take Up* value which presents the following correlations:

- *At Least 100 Mbps fixed BB* take up with a correlation value equal to an amount of 0.7423;
- *Broadband price index* with a correlation index value equal to a value of -0.1567.

Fixed broadband coverage variable has the following correlation:

- *At Least 100 Mbps fixed BB Take Up* with a correlation index value equal to a value of 0.7197.

The *Broadband price index* is then considered, which has the following correlations, that is:

- *DESI Index Aggregate score* with a correlation index value equal to a value of -0.0369;
- *e-Government* with a correlation index value equal to a value of -0.577;
- *At Least Basic Digital Skills* with a correlation index value equal to a value of -0.1239;

- *SMEs with at least a basic level of digital intensity* with a correlation index value equal to a value of -0.1337;
- *Integration of Digital Technology* with a correlation index value equal to a value of -0.2039.

The value of the *e-Government* appears to have the following correlations, namely:

- *DESI Index Aggregate Score* with a correlation index value equal to a value of 0.9281;
- *Integration of Digital Technology* with a correlation index value equal to a value of 0.7467;
- *SMEs with at least basic level of digital intensity* with a correlation index value equal to an amount of 0.7458.

The value of the *Integration of Digital Technology* variable has the following correlations, that is:

- *SMEs with at least a basic level of digital intensity* with a correlation index value of 0.91 units;
- *DESI Index aggregate score* with a correlation index value equal to a value of 0.8914.

The value of the *DESI Index with Aggregate Score* has the following correlations, that is:

- *SMEs with at least a basic level of digital intensity* with a correlation index value equal to a value of 0.8626;
- *At Least Basic Digital Skills* with a Correlation Index Value of 0.733

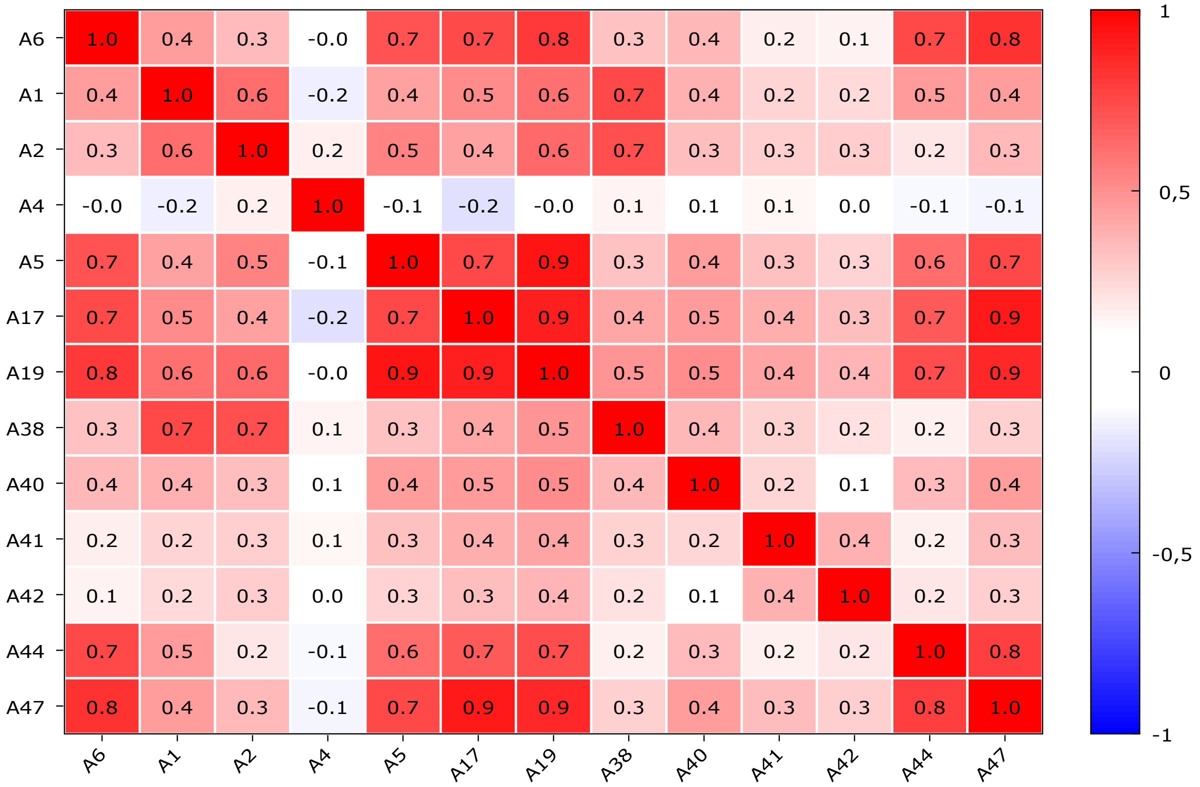


Figure 1. Correlation Matrix.

3. The Econometric Model

An econometric model is then created to estimate the value of ICT Specialists in the European Union. Data from the DESI-Digital Economy and Society Index for 28 countries[6] of the European Union between 2014 and 2021 were used. The data were analyzed using the following econometric techniques, namely: Panel Data with Fixed Effects, Panel Data with Random Effects, WLS and Pooled OLS. In particular, the following formula was estimated in an explicit form:

$$\begin{aligned}\boldsymbol{ICTSpecialists_{it}} &= \boldsymbol{a_1 + b_1(FixedBroadbandTakeUp)_{it} + b_2(FixedBroadbandCoverage)_{it}} \\ &+ \boldsymbol{b_3(BroadbandPriceIndex)_{it} + b_4(eGovernment)_{it}} \\ &+ \boldsymbol{b_5(IntegrationOfDigitalTechnology)_{it} + b_6(DESI)_{it}} \\ &+ \boldsymbol{b_7(AtLeast100MbpsFixedBBTakeUp)_{it} + b_8(4GCoverage)_{it}} \\ &+ \boldsymbol{b_9(5GReadiness)_{it} + b_{10}(5GCoverage)_{it}} \\ &+ \boldsymbol{b_{11}(AtLeastBasicDigitalSkills)_{it}} \\ &+ \boldsymbol{b_{12}(SMEsWithAtLeastABasicLevelOfDigitalIntensity)_{it}}\end{aligned}$$

$$\boldsymbol{Where\ i = 28\ and\ t = [2016; 2021]}$$

The analysis shows that the value of "*ICT Specialists*" is positively associated with the following variables

- *Desi Index Aggregate Score:* it is an indicator consisting of the sum of the following variables, namely "*Human Capital*", "*Connectivity*", "*Integration of Digital Technology*" and "*Digital Public Services*". In extended form it is possible to explain the following relationship as indicated below, that is:

$$\begin{aligned}\boldsymbol{DesiAggregateScore_{it}} &= \boldsymbol{a_1 + b_1(HumanCapital)_{it} + b_2(Connectivity)_{it}} \\ &+ \boldsymbol{b_3(IntegrationOfDigitalTechnology)_{it}} \\ &+ \boldsymbol{b_4(DigitalPublicServices)_{it}}\end{aligned}$$

$$\boldsymbol{Where\ i = 28\ and\ t = [2016; 2021]}$$

[6] Countries are: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, European Union, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden.

The data therefore show the existence of a positive relationship between the value of ICT Specialists in Europe and the value of the DESI Index, this relationship indicates that within the European countries considered and limited to the 2016-2021 period, the percentage of workers operating in the ICT sector tends to grow with the growth in the value of the overall degree of digitization. Obviously, it follows that digitization, i.e., as a structure consisting of human capital, connectivity tools, digital public services and integration of digital services tends to be associated with a growth in experts operating in the informatics sector. It therefore follows that to increase the overall value of ICT specialists, it is necessary to invest in an extended approach to digitization.

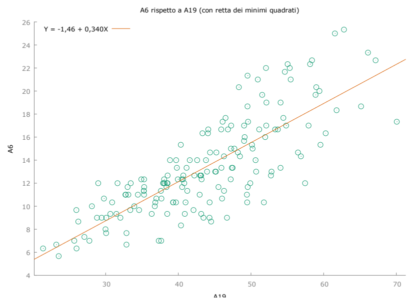


Figure 2. Relationship between ICT specialists and Desi Index.

- *At Least 100 Mbps fixed BB Take UP:* It is a variable that considers the "*percentage of households subscribing to fixed broadband of at least 100 Mbps, calculated as overall fixed broadband take up multiplied with the percentage of fixed broadband lines of at least 100 Mbps*". The presence of internet networks with at least 100 Mbps is positively associated with the presence of "ICT Specialists" in Europe. This relationship is since the network that has at least 100 Mbps is particularly widespread in European countries as it constitutes the basic value of connections to internet. It follows that to ensure that there is a growth of specialists in the ICT sector it is not necessary to invest in particularly fast networks, on the contrary, even a substantially basic network such as 100 Mbps could be sufficient to generate the conditions for training of the human capital and of a class of professionals operating in ICT. These conditions indicate that although infrastructural investments are important in determining the possibility of creating professionals operating in ICT, it is also true that the essential element for the training of ICT professionals is the dimension of human capital in its connections. with the education system, including university and tertiary education.

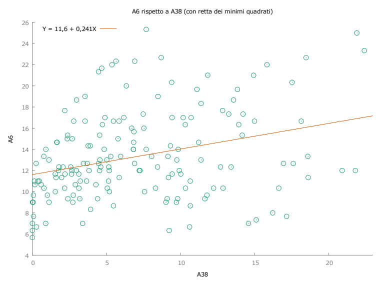


Figure 3. Relationship between the value of ICT Specialists and the value of At Least 100 Mbps fixed BB Take UP.

- *SMEs with at least a basic level of digital intensity:* it is a variable that considers the digital intensity based on counting how many out of 12 selected technologies are used by enterprises. A basic level requires usage of at least 4 technologies. There is a relationship between the value of ICT Specialists and the value of companies that have a minimum level of digitization. This positive relationship can be understood considering that very often it is the companies that train employees in the use of information technologies and that where there are companies that have high levels of digitization it is also possible that there are training programs for employees in the ICT sectors. For example, companies make training programs for employees such as academies or they can also organize hackathons for the selection of personnel, or they can also

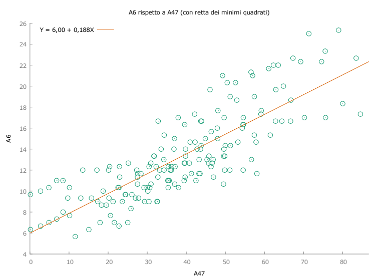


Figure 4. Relationship between the value of ICT Specialists and the value of SMEs with at least a basic level of digital intensity

accompany the insertion and training of employees with programs for the growth of computer skills.

The analysis shows that the value of "*ICT Specialists*" is negatively associated with the following variables:

- *Fixed Broadband Take Up:* is a variable made up of three different sub-variables, namely "*Overall Fixed Broadband Take Up", "At Least 100 Fixed Broadband Take Up*", "*At Least 1 Gbps Take Up*". There is a negative relationship between the value of "ICT Specialists" and the value of "*Fixed Broadband Take Up*". This negative relationship indicates the fact that the possibility of European countries to grow in the training of ICT Specialists does not depend so much on the internet structures that are present within the countries. In fact, to create experienced professionals in ICT it is necessary to invest more in human capital than in networks, although at least a certain minimum level of networks is necessary to develop the sector. The investment in networks is important for developing the overall digitization of a certain country, however if policy makers want to ensure that the percentage of workers active in the ICT sector increases, more investments need to be made in the human capital such as through education and training institutions and through IT training programs implemented in companies. it is possible to represent the constitutive expression of fixed broadband take up in explicit form as indicated below:

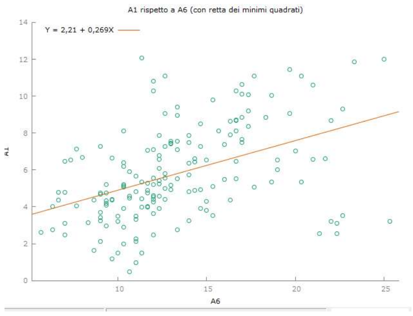


Figure 5. Relationship between the value of ICT Specialists and the value of Fixed Broadband Take Up in Europe.

$$\begin{aligned} \boldsymbol{FixedBroadTakeUp_{it}} &= \boldsymbol{a_1 + b_1(OverallFixedBroadbandTakeUp)_{it}} \\ &\quad \boldsymbol{+ b_2(AtLeast100FixedBroadbandTakeUp)_{it}} \\ &\quad \boldsymbol{+ b_3(AtLeast1GbpsTakeUp)_{it}} \end{aligned}$$

$$\boldsymbol{Where\ i = 28\ and\ t = [2016; 2021]}$$

- *Fixed Broadband Coverage:* it is a variable consisting of two sub-variables namely "*Fast Broadband NGA Coverage*" and "*Fixed Very High-Capacity Network VHCN Coverage*". There is therefore a negative relationship between the value of ICT Specialists in Europe and the value of "*Fixed Broadband Coverage*". This negative relationship indicates that the fact that the coverage of the fixed internet network is not a determining factor for the growth in the number of operators specializing in IT disciplines. This relationship may appear counterfactual. However, it must be considered that although the coverage of networks is a relevant element in determining the overall structure of digitization in a certain European country, it is also true that the variables that have the greatest impact on the formation of human capital in the ICT field are those related to the formation of capital. human. In fact, to

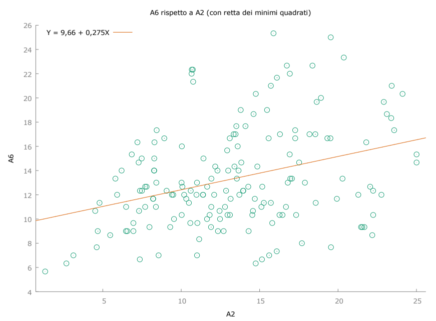


Figure 6. Relationship between ICT Specialists and Fixed Broadband Coverage value.

increase the presence of specialists in the ICT sector, it is necessary to create training courses both within the university system and in companies. It is also possible to create an explicit representation in extended form of the "*Fixed Broadband Coverage*" variable as indicated below, that is:

$$\begin{aligned} \boldsymbol{FixedBroadbandCoverage_{it}} &= \boldsymbol{a_1 + b_1(FastBroadbandNGACoverage)_{it}} \\ &+ \boldsymbol{b_2(FixedVeryHighCapacityNetworkVHCNCoverage)_{it}} \end{aligned}$$

$$\boldsymbol{Where\ i = 28\ and\ t = [2016; 2021]}$$

- *Broadband Price Index:* it is a variable that measures the prices of representative baskets of fixed, mobile, and converged broadband offers in a score between 0 and 100. There is a negative correlation between the value of the fixed network price and the value of the specialists operating in the ICT sector. This relationship indicates that where the price of the fixed network tends to increase, the distribution of ICT Specialists operators tends to decrease. This condition indicates that for the number of ICT operators to grow, the price of the fixed network must be accessible. It follows that for businesses, families and even institutions operating in the training of human capital, the price of the fixed network must be reduced to create the conditions for greater diffusion and use. Low fixed network prices can stimulate the creation of new companies operating in the ICT sector that offer training programs to their employees, as well as could push people to invest in online training courses to acquire computer-type hard skills. the increase in the price of the network can therefore constitute an "*entry barrier*" for some countries that are interested in developing a workforce with a marked professionalism in the ICT sector.

Figure 7. Relationship between ICT Specialists and Broadband Price Index.

- *E-Government:* is a variable made up of a set of sub variables that is "*e Government Users*", "*Pre-Filled Forms*", "*Digital Public Services for Citizens*", "*Digital Public Services for Businesses*", "*Open Data*". There is a negative relationship between the value of e-government and the value of ICT Specialists. This negative relationship indicates that the fact that a country increases its investments in e-government does not necessarily imply that there is a growth in ICT specialists. This negative relationship can be better understood considering that the growth of specialists in the ICT sector tends to grow with the growth of investment in university and professional educational facilities and education which are oriented towards the training of operators in the sector. Therefore, although investment in the e-government sector is important for the overall degree of digitization of a country, it fails to have an impact in terms of growth of ICT operators. If policy makers intend to increase the presence of ICT operators, it is necessary for

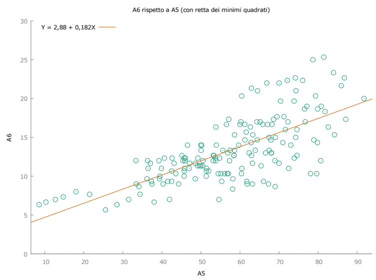


Figure 8. Relationship between the value of ICT Specialists and the value of e-Government.

training to be carried out both at university and company level. The equation can be explicitly expressed in the following form:

$$\begin{aligned} &\boldsymbol{eGovernment_{it}} \\ &= \boldsymbol{a_1 + b_1(eGovernmentUsers)_{it} + b_2(PreFilledForms)_{it}} \\ &+ \boldsymbol{b_3(DigitalPublicServicesForCitizens)_{it}} \\ &+ \boldsymbol{b_4(DigitalPublicServicesForBusinesses)_{it} + b_5(OpenData)_{it}} \end{aligned}$$

$$\boldsymbol{Where\ i = 28\ and\ t = [2016; 2021]}$$

- *Integration of Digital Technology:* is made up of 3 sub-dimensions: digital intensity, take-up of selected technologies by enterprises and e-commerce. SMEs with at least a basic level of digital intensity, take-up of Big Data, Cloud and AI are targets of the Digital Decade Compass. There is a negative relationship between the value of Integration of Digital Technology and the value of ICT Specialists in Europe. This relationship indicates that the fact that companies invest and integrate digital technologies does not increase the development of professionalism in the ICT sector. Such a relationship might appear illogical. However, it should be borne in mind that to create professionals in the ICT sector, it is necessary to invest in organizations capable of training human capital. Therefore, the growth of experienced professionals in the ICT sector can only take place by investing more in university and post-graduate training in IT disciplines. The fact that companies adopt digital technologies could be seen as a fact of context, however, to ensure that there is real growth among IT professionals, it is necessary to invest in training. In this regard, policy makers can provide incentives to ensure that the acquisition of digital technologies is accompanied by the presence of adequate training courses in companies.

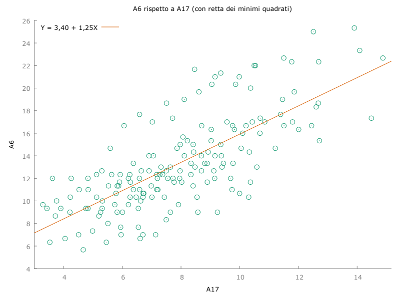


Figure 9. Relationship between the value of ICT Specialists and the value of Integration of Digital Technologies.

- *4G Coverage:* is a variable that considers the percentage of populated areas with coverage by 4G. There is a negative relationship between the presence of professionals operating in the ICT sector and the presence of populated areas covered by 4G. This negative relationship may seem unreasonable as one might imagine that the growth of network coverage through 4G could in some way promote a greater information technology culture and therefore the growth of professionals in the information technology sector. However, to increase the percentage of workers who are operating in the ICT sector, it is necessary to invest in training rather than in communication networks. In fact, communication networks can constitute the premise for a digitized system, however they do not say anything about the possibility of training expert workers in the ICT sector. To increase the percentage of ICT operators, it is necessary to invest in training at both university and

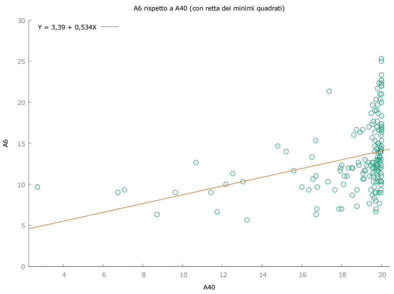


Figure 10. Relationship between the value of ICT Specialists and coverage with 4G network.

company level. Certainly, the presence of 4G networks can be a useful tool for online training, but only if there are institutions that are specialized in providing courses remotely.

- *5G Readiness:* is a variable that considers the amount of spectrum assigned and ready for 5G use withing the so-called 5G pioneer bands. These bands are 700 MHz (703-733 MHz and 758-788 MHz) and 26 GHz (1000 MHz within 24520-27500 MHz). All three spectrum bands have an equal weight. There is therefore a negative relationship between the value of 5G Readiness and the value of ICT Specialists. This negative relationship could appear counterfactual as it could seem that the countries that have greater 5G readiness are also those that are more evolved from the point of view of digital culture with the highest values in terms of professional endowment. However, as is evident, it appears that the possibility of a country to increase the supply of professional specialists in the ICT sector essentially depends on the structure of training both at the medical and corporate level. 5G can increase training and the offer of training products also online, however, only if there are organizations that are engaged in providing distance training.

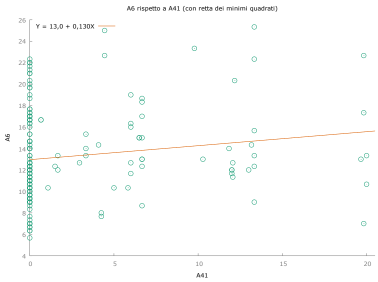


Figure 11. Relationship between the value of ICT Specialists and the value of 5G Readiness.

- *5G Coverage:* is a variable that considers the percentage of populated area with coverage by 5G. There is a negative relationship between the value of ICT Specialists and the value of the percentage of populated area covered by 5G. This relationship could appear counterfactual since it could be thought that the growth of 5G coverage could in some way stimulate the growth of workers who are operating in the ICT sector. However, this relationship is negative since the growth of operators who have training in the ICT sector does not depend on the presence of 5G but rather on the presence of university or corporate training courses that are specifically aimed at growth and dissemination of professional skills in terms of ICT. It follows therefore that although 5G is certainly necessary for the growth of digitalization, it does not have an immediate impact on the growth of ICT specialists, on the contrary it has a negative impact. If policy makers want to increase the presence of specialized ICT operators, they must invest in training structures, be they public or private, university or corporate, to increase the training of ICT personnel.

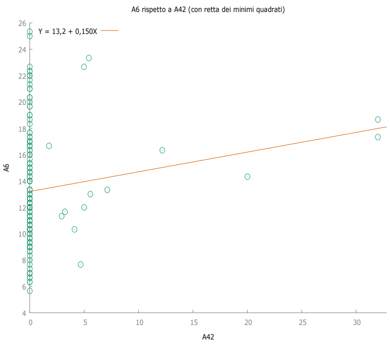


Figure 12. Relationship between the value of ICT Specialists and the value of 5G Coverage.

- *At Least Basic Digital Skills:* is a variable that considers individuals with basic or above basic digital skills in each of the following four dimensions: information, communication, problem solving, and software for content creation. There is a negative relationship between the value of ICT Specialists and the value of "*At Least Basic Digital Skills*". This relationship may appear counterfactual however it is evident that having basic skills does not make individuals experts in the ICT sector. Certainly, the fact that the population has relevant IT skills is a positive element that can certainly support the overall digitization of a country. However, for the

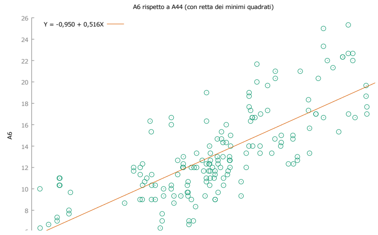


Figure 13. Relationship between the value of ICT Specialists and the value of "At Least Basic Digital Skills".

percentage of the population operating as ICT Specialists to increase, it is necessary to have skills that are not only basic and that are more advanced. A type of training that can be effectively produced only through investment in the university system or in company-type courses.

Variables		Fixed Effects		Random Effects		Pooled OLS		WLS		Coefficient Mean
		Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	
Variables	*const*	6,1349	***	3,1385	**	-1,9499		-1,9220	**	1,3504
Fixed broadband take-up	A1	-0,5945	***	-0,6126	***	-0,5733	***	-0,6168	***	-0,5993
Fixed broadband coverage	A2	-0,3497	***	-0,3847	***	-0,6141	***	-0,6012	***	-0,4874
Broadband price index	A4	-0,7049	***	-0,6225	***	-0,4633	***	-0,4557	***	-0,5616
e-Government	A5	-0,3665	***	-0,3721	***	-0,4643	***	-0,4627	***	-0,4164
Integration of Digital Technology	A17	-1,7612	***	-1,9254	***	-2,6570	***	-2,6756	***	-2,2548
Aggregate score	A19	1,3938	***	1,4756	***	1,9304	***	1,9365	***	1,6841
At least 100 Mbps fixed BB take-up	A38	0,1086	**	0,1118	***	0,2097	***	0,1908	***	0,1552
4G coverage	A40	-0,1130	***	-0,1302	***	-0,1709	***	-0,2009	***	-0,1538
5G readiness	A41	-0,1854	***	-0,1987	***	-0,2755	***	-0,2795	***	-0,2348
5G coverage	A42	-0,1741	***	-0,1814	***	-0,2244	***	-0,2357	***	-0,2039
At least Basic Digital Skills	A44	-0,3949	***	-0,3360	***	-0,3596	***	-0,3598	***	-0,3626
SMEs with at least a basic level of digital intensity	A47	0,1628	***	0,1602	***	0,1433	***	0,1539	***	0,1550

Figure 14. Summary of the econometric results obtained through the Panel, WLS and OLS estimates.

From the analysis there are two different types of elements that have the greatest impact both in a negative and a positive sense on the degree of "ICT Specialists" present at the national level. The Desi Index is the variable that has the most positive impact on the value of ICT Specialists. It therefore follows that the overall degree of digitization of a country is generically associated with the presence of a growth in human capital trained in ICT disciplines. Instead, the "Integration of Digital Technology" indicator is the negative element that has the greatest impact on the determination of ICT Specialists. This negative relationship indicates the fact that the simple adoption of information technologies by companies is not able to guarantee a growth of ICT specialists. In fact, it is not certain that the technologies that companies implement require the presence of qualified personnel, and sometimes in the context of process innovations, they can even generate a reduction in employment. It therefore follows that the digitization of companies, the fact that companies use e-commerce, to create a positive impact in terms of growth of ICT Specialists must also be associated with adequate training of employees also in connection with university institutions.

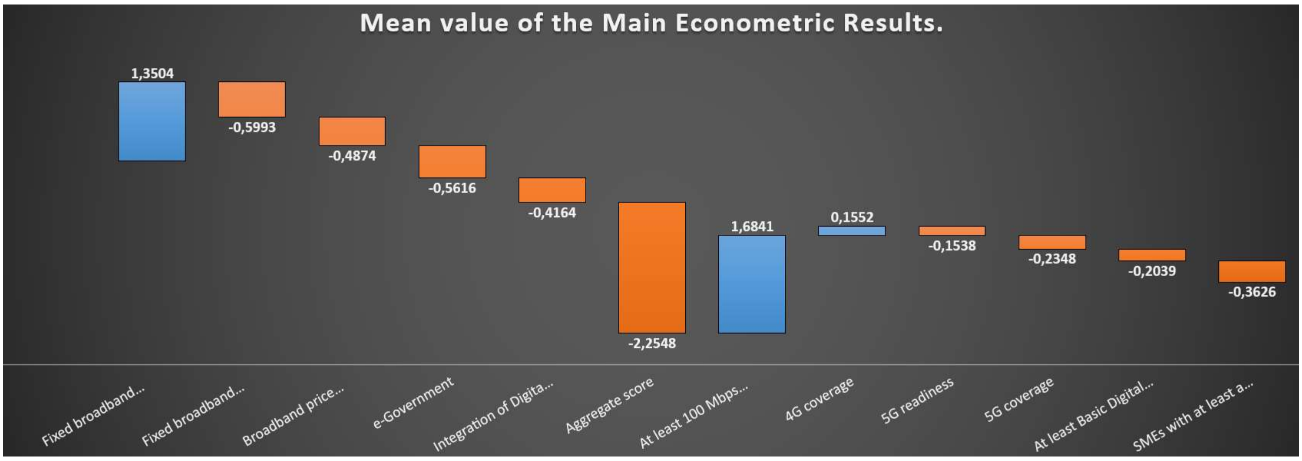


Figure 15. Mean Value of the Main Econometric Results.

4. Clusterization

Subsequently, a clusterization was carried out using the unsupervised k-Means algorithm optimized through the Silhouette coefficient. The Silhouette coefficient varies between -1 and 1. The positivity of the Silhouette coefficient suggests the optimal number of clusters for a variable. In the analyzed case, the optimal value of the clusters is equal to 2. The composition of the clusters is indicated below:

- *Cluster 1*: Belgium, Denmark, Estonia, Finland, Ireland, Luxembourg, Netherlands, Sweden;
- *Cluster 2*: Austria, Bulgaria, Croatia, Cyprus, Czechia, European Union, France, Germany, Greece, Hungary, Italy, Latvia, Lithuania, Malta, Poland, Portugal, Romania, Slovakia, Slovenia, Spain.

Considering the median of "*ICT Specialists*" it appears that the value of the nations in cluster 1 is equal to 19.33 while the median of the countries in cluster 2 is equal to 11.83. Therefore, the following ranking in terms of the median value of the ICT Specialists among the various clusters derives, namely: C1> C2. The cluster analysis highlights that there is a clear contrast between the Scandinavian countries, with the addition of Belgium and Luxembourg which have high values of ICT Specialists, and the rest of Europe with lower values. It should be considered that in this case the "*ICT Specialists*" variable takes into consideration the percentage value of ICT workers on the total number of workers considered. Therefore, even if it is possible that there are more ICT workers in absolute value in Germany than in Estonia, it follows that in percentage terms Estonia has more ICT workers than Germany. Furthermore, it is also necessary to make some strategic considerations regarding the fact that countries choose to invest significantly in the ICT sector by creating companies operating in the supply of IT services and products. In fact, since there is a significant demand for work in the ICT sector at the country level, it is likely that investments in the training of human capital in the ICT sector will also occur. This condition appears to be quite low in Europe as there are no European companies that can be competitive with US Big Tech giants such as Google, Facebook, and Amazon. This condition obviously also impacts on the training choices of workers who may not have the adequate incentives to acquire hard skills in the IT sector. In any case, the result is a clear contrast between an evolved Northern Europe from the point of view of the presence of skilled ICT workforce and the rest of Europe with lower values in terms of the presence of skilled workers in ICT disciplines. A contrast that the policy maker should try to resolve by investing on the one hand in the formation of human capital and on the other hand by stimulating a demand for work that is increasingly oriented towards operators specialized in ICT. In this regard, to increase the percentage of operators in the ICT sector, it is necessary to invest both through the economic policies of education, especially university and post-university education, and by investing through the implementation of specific industrial policies to increase the demand for specialized jobs. The mix of educational economic policies and industrial economic policies, i.e., the combination of professional training including university, and mobilization of the industrial system, should allow the economic system and the population to have the right stimuli to be able to increase the percentage of workers specialized in the ICT sector. In any case, the cluster analysis shows that there is at least a digital divide between Northern and Central Southern Europe in terms of professional training in the ICT sector. Even if by deepening the analysis it is possible to verify that there are two digital divides: one between Northern and Southern Europe and one between Western Europe and Eastern Europe. These contrasts evidently reflect the greater overall dynamism of the countries of Northern Europe. However, it is necessary that at least in part this gap be reduced through the preparation of adequate European economic policies.

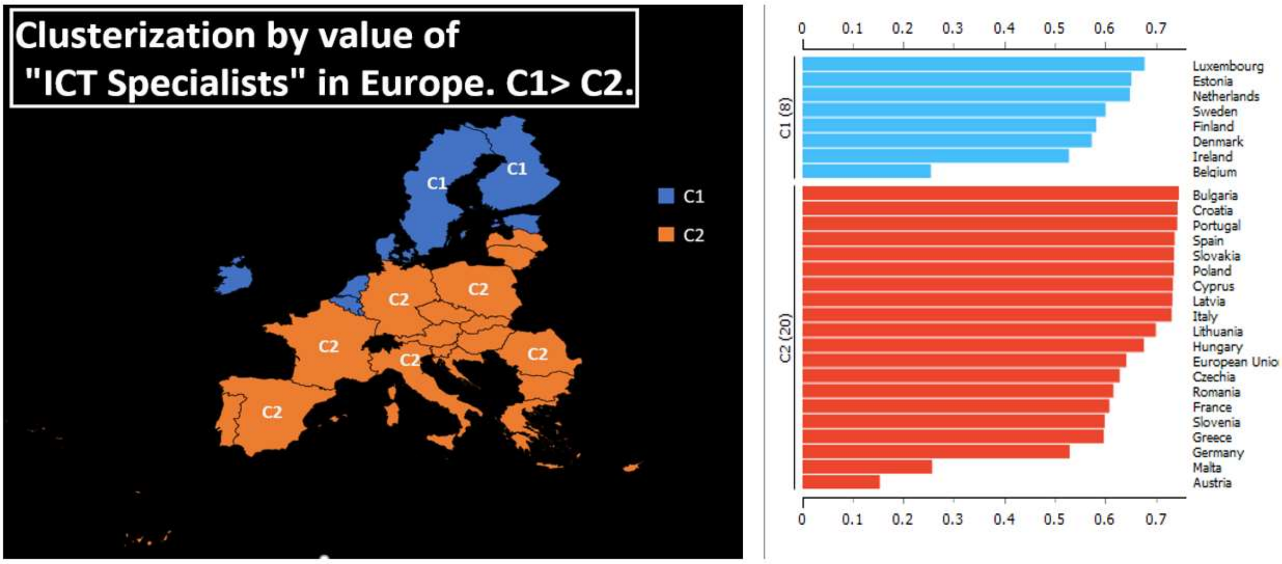


Figure 16. Clustering of countries by value of ICT Specialists with k-Means algorithm optimized by the Silhouette coefficient.

5. Machine Learning and Prediction

An analysis was then carried out using eight machine learning algorithms to predict the value of "*ICT Specialists*". Specifically, the algorithms were compared in terms of performance based on the ability to maximize the R-square and a set of statistical errors or "*Mean absolute error*", "*Mean squared error*", "*Root mean squared error*", "*Mean signed difference*", "*Mean absolute percentage error*". The data was trained with 70% of the available data while the remaining 30% was used for prediction. Following the performance analysis, the following ranking of the algorithms results, namely:

- *ANN-Artificial Neural Network* with a payoff value of 11;
- *PNN-Probabilistic Neural Network* with a payoff value of 14;
- *Linear Regression* with a payoff value of 16;
- *Polynomial Regression* with a payoff value of 26
- *Simple Regression Tree* with a payoff value of 32;
- *Gradient Boosted Tree Regression* with a payoff value of 36;
- *Tree Ensemble Regression* with a payoff value of 37;
- *Random Forest Regression* with a payoff value of 44.

However, since normalization is expected to be used in the prediction with the ANN machine learning algorithms, before realizing the actual data prediction it is necessary to proceed with the denormalization of the data. In fact, the data have been normalized in the range between 0 and 1. To carry out the denormalization and calculate the variations between the historical series and the predicted value, it is necessary to proceed with the de-normalization which was carried out using the following formula:

$$\begin{cases} \boldsymbol{y = y' * StandardDeviation(y) + Mean(y)} \\ \boldsymbol{y = Denormalized\ Data\ Before\ Prediction} \\ \boldsymbol{y' = Normalized\ Data} \end{cases}$$

Once the denormalization was carried out, it was possible to calculate the variations in the value of ICT Specialists in European countries. Following the choice of the best algorithm for the prediction or the ANN-Artificial Neural Network, the following predictions are made:

- *Austria* with a positive change from an amount of 14.33 up to a value of 15.933 units equal to a value of 1.60 units equal to a value of 11.15%;
- *Belgium* with a decrease from an amount of 16.66 units up to a value of 16.438 units or equal to a variation of -0.23 units equal to a value of -1.37%;
- *Bulgaria* with an increase from an amount of 10.33 units up to a value of 14.695 units or equal to a variation of 4.36 units equal to a value of 42.21%;
- *Croatia* with an increase from an amount of 10.66 up to a value of 14.87 units equal to a change of 4.21 units equal to an amount of 39.44%;
- *Cyprus* with an increase from an amount of 9.00 units up to a value of 14.56 units equal to a value of 5.56 units equal to an amount of 61.81%;
- *Czechia* with an increase from an amount of 13.33 units up to a value of 15.42 units equal to a value of 2.09 units equal to 15.68%;
- *Denmark* with a decrease from an amount of 17.33 units up to a value of 16.65 units equal to a variation of -0.67 units equal to a variation of -3.89%;
- *Estonia* with a variation from an amount of 20.00 units up to a value of 17.34 units or equal to a variation of -2.66 units equal to an amount of -13.30%;
- *European Union* with a variation from an amount of 13.00 units up to a value of 15.45 units equal to a variation of 2.46 units equal to an amount of 18.91%;
- *Finland* with a decrease from an amount of 22.66 to a value of 17.58 units equal to a variation of -5.08 units equal to -22.43%;
- *France* with an increase from an amount of 14.00 units up to a value of 15.47 units or equal to a variation of 1.48 units equal to an amount of 10.54%;
- *Germany* with an increase from a value of 13.33 units up to a value of 15.62 units equal to an amount of 2.29 units equivalent to a value of 17.15%;
- *Greece* with an increase from an amount of 7.00 units up to a variation of 14, 125 units or equal to a variation of 7.13 units equal to an amount of 101.79%;
- *Hungary* with an increase from an amount of 11.33 units up to a value of 15.16 units or equal to a variation of 3.83 units equal to 33.84%;
- *Ireland* with a variation from an amount of 16.33 units up to a value of 16.87 units equal to a variation of 0.54 units equal to a variation of 3.30 units;
- *Italy* with an increase from an amount of 11.66 units up to a value of 14.89 units equal to a variation of 3.23 units equal to a variation of 27.68%;
- *Latvia* with a variation from an amount of 10.33 units up to a value of 14.66 units equal to a variation of 4.33 equal to a variation of 41.93%;
- *Lithuania* with a variation of 10.33 units equal to a variation of 14,440 units equal to a variation of 4.11 equivalent units from 39.74%;

- *Luxembourg* with a decrease from an amount of 20.33 units up to a variation of 17.098 units equal to a variation of -3.24 units equal to an amount of 15.91%;
- *Malta* with an increase from a value of 15.33 units up to a value of 15.671 units equal to an amount of 0.34 units equal to an amount of 2.20%;
- *Netherlands* with a decrease from a value of 18.667 with a variation of 16.95 units equal to a variation of -1.71 units equal to a value of -9.18%;
- *Poland* with a variation from an amount of 10.33 units equal to a variation of 14.636 units equal to a variation of 4.30 units equal to an amount of 41.64%;
- *Portugal* with a variation from 12.00 units up to a value of 15.011 units or equal to a variation equal to an amount of 3.01 units equal to a variation of 25.09%;
- *Romania* with an increase from an amount of 7.667 units up to a value of 14.17 units or equal to a value of 6.51 units equal to an amount of 84.90%;
- *Slovakia* with an increase from an amount of 12.33 units up to a value of 15.02 units or equal to a value of 2.69 units equal to an amount of 21.81%;
- *Slovenia* with an increase from a value of 13.00 units up to a value of 15.43 units equal to a value of 2.44 units equal to a value of 18.75%;
- *Spain* with an increase from an amount of 12.00 units up to a value of 15.006 units equal to a value of 3.01 units equal to a variation of 25.05%;
- *Sweden* with a decrease from an amount of 23.33 units equal to a variation of 15.88 equal to a variation of -7.45 units equal to a value of -31.94%.

On average, using the *ANN-Artificial Neural Network* algorithm, a growth of ICT Specialists is expected from an amount equal to 13.81 units up to a value of 15.54 units or equal to a variation of 1.73 units equal to an amount of 12.53%.

Comparison between historical values and predicted values using the ANN-Artificial Neural Network.				
Country	**2021**	**Prediction**	**Absolute variation**	**Percentage Variation**
Austria	14,333	15,931	1,60	11,15
Belgium	16,667	16,438	-0,23	-1,37
Bulgaria	10,333	14,695	4,36	42,21
Croatia	10,667	14,874	4,21	39,44
Cyprus	9,000	14,563	5,56	61,81
Czechia	13,333	15,424	2,09	15,68
Denmark	17,333	16,659	-0,67	-3,89
Estonia	20,000	17,341	-2,66	-13,30
European Union	13,000	15,458	2,46	18,91
Finland	22,667	17,583	-5,08	-22,43
France	14,000	15,475	1,48	10,54
Germany	13,333	15,620	2,29	17,15
Greece	7,000	14,125	7,13	101,79
Hungary	11,333	15,168	3,83	33,84
Ireland	16,333	16,872	0,54	3,30
Italy	11,667	14,896	3,23	27,68
Latvia	10,333	14,666	4,33	41,93
Lithuania	10,333	14,440	4,11	39,74
Luxembourg	20,333	17,098	-3,24	-15,91
Malta	15,333	15,671	0,34	2,20
Netherlands	18,667	16,953	-1,71	-9,18
Poland	10,333	14,636	4,30	41,64
Portugal	12,000	15,011	3,01	25,09
Romania	7,667	14,176	6,51	84,90
Slovakia	12,333	15,023	2,69	21,81
Slovenia	13,000	15,437	2,44	18,75
Spain	12,000	15,006	3,01	25,05
Sweden	23,333	15,880	-7,45	-31,94
Mean	13,810	15,540	1,73	12,53

Figure 17. Comparison between historical values and predicted values using the ANN-Artificial Neural Network.

6. Prediction with Augmented Data

A prediction was then made through augmented data. Augmented data were produced by adding to the historical series the predictions made using the ANN-Artificial Neural Network algorithm.

$$\boldsymbol{AugmentedData} = \boldsymbol{HistoricalSeries} + \boldsymbol{ANNPrediction}$$

Using augmented data, it was possible to make predictions by comparing eight different machine learning algorithms to predict the value of ICT Specialists. 70% of the data was used for learning while the remaining 30% was used for actual prediction. For the choice of the most performing algorithm, a comparison was made based on the ability to maximize the R-squared and minimize statistical errors. The statistical errors used are the $\boldsymbol{MAE(y,\hat{y})} = \frac{1}{N}\sum_{t-1}^{N} |y_i - \widehat{y_i}|$, $\boldsymbol{MSE(y,\hat{y_i})^2} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y_i})^2$, $\boldsymbol{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(y_k - \widehat{y_k})^2}$.

Through the analysis of the indicated parameters, it was possible to obtain the following ranking of the algorithms based on the predictive capacity, that is:

- *PNN-Probabilistic Neural Network* with a payoff value equal to an amount of 4;
- *Tree Ensemble Regression* with a payoff value of 10;
- *ANN-Artificial Neural Network* with a payoff value equal to a value of 11;
- *Random Forest Regression* with a payoff value equal to a value of 17;
- *Gradient Boosted Tree Regression* with a payoff value equal to a value of 18;
- *Simple Regression Tree* with a payoff value of 24;
- *Linear Regression* with a payoff value of 28;
- *Polynomial Regression* with a payoff value of 32.

Also, in this case it was necessary to carry out a denormalization of the errors to be able to compare the prediction made with the augmented data with the prediction made with historical data. The analysis carried out shows that the best predictor algorithm is the PNN-Probabilistic Neural Network which presents the best results in terms of maximization of R ^ 2 and statistical errors. Using the PNN it was possible to obtain the following results:

- *Bulgaria* with an increase in value from an amount equal to 14.70 units up to a value of 15.75 units or equal to a growth of 1.06 units equal to a value of 7.20%;
- *Estonia* with a decrease from an amount of 17.34 units up to a value of 16.39 units or equal to a variation of -0.95 units equal to a value of -5.51%;
- *Greece* with a variation from an amount of 14.13 units up to a value of 15.55 units equal to an amount of 1.43 units equal to a value of 10.12%;
- *Hungary* with an increase from 15.17 units up to a value of 15.76 units or equal to a variation of 0.59 units equal to an amount of 3.90%;

- *Ireland* with a decrease from an amount of 16.87 units up to a value of 16.26 units or equal to a variation of -0.61 units equal to a value of -3.62%;
- *Portugal* with an increase from a value of 15.01 units equal to a value of 15.80 units equivalent to a variation of 0.78 units equal to a value of 5.23%;
- *Slovenia* with an increase from an amount of 15.44 units up to a value of 15.91% units up to a value of 0.47 units equal to a value of 3.06%;
- *Spain* with an increase from an amount of 15.01 units equal to a value of 15.76 units equal to a change of 0.75 units equal to an amount of 5.02%.

Overall, the value of the incremental variation increased from a value of 15.46 units up to a value of 15.90 units equal to a variation of 0.44 units equal to a value of 3.18%.

Prediction made with augmented data with the best predictor PNN-Probabilistic Neural Network.

Countries	Prediction 1	Augmented Data	Absolute Variation	Percentage Variation
Bulgaria	14,70	15,75	1,06	7,20
Estonia	17,34	16,39	-0,95	-5,51
Greece	14,13	15,55	1,43	10,12
Hungary	15,17	15,76	0,59	3,90
Ireland	16,87	16,26	-0,61	-3,62
Portugal	15,01	15,80	0,78	5,23
Slovenia	15,44	15,91	0,47	3,06
Spain	15,01	15,76	0,75	5,02
Mean	15,46	15,90	0,44	3,18

Figure 18. Prediction made with augmented data with the best predictor PNN-Probabilistic Neural Network.

It is also possible to make a comparison between the R-squared and the statistical errors between the value of the prediction with historical data and the prediction with augmented data with the following results:

- *R^2*: with a reduction from a value of 0.98 units up to a value of 0.97 units or equal to a value of -0.01 units equal to a value of -1.02%;
- *Mean Absolute Error*: with an increase from a value of 0.03 units up to a value of 0.04 units or equal to an amount of 0.01 units equal to a value of 25.37%;
- *Mean Squared Error*: with a variation from an amount of 0.001 up to a value of 0.002 units or equal to a value of 0.001 units equal to a variation of 51.78%;
- *Root Mean Squared Error*: with a variation from a value of 0.037 units up to a value of 0.046 units or equal to a variation of 0.009 units equal to a value of 23.202%.

Therefore, it results that the value of the average of the statistical errors has grown from an amount of 0.023 to a value of 0.028 units or equal to a variation of 0.006 units equal to a value of 24.71%.

7. Conclusions

We have estimated the value of ICT Specialists in Europe between 2016 and 2021 for 28 European countries. In the first paragraph, an analysis of the literature is presented which highlights the role of training for ICT operators in promoting digitization and economic growth at the country level. In the third paragraph an econometric analysis was presented to estimate the value of the presence of ICT specialists in Europe. The data were analyzed using the following econometric techniques, namely: Panel Data with Fixed Effects, Panel Data with Random Effects, WLS and Pooled OLS. The results show that the value of ICT Specialists in Europe is positively associated with the following variables: "*Desi Index*", "*SMEs with at least a basic level of digital intensity*", "*At least 100 Mbps fixed BB take-up*" and negatively associated with the following variables: "*4G Coverage*","*5G Coverage*", "*5G Readiness*", "*Fixed broadband coverag*e", "*e-Government*", "*At least Basic Digital Skills*", "*Fixed broadband take-up*", "*Broadband price index*", "*Integration of Digital Technology*". Subsequently, two European clusters were found by value of "*ICTG Specialists*" using the k-Means clustering algorithm optimized by using the Silhouette coefficient. Clustering highlights the presence of a significant contrast between the values of Northern Europe and the values of Central and Southern Europe. There is therefore a significant gap in terms of the presence of ICT specialists who should be subject to specific economic policies. Moreover, eight different machine learning algorithms were compared to predict the value of "*ICT Specialists*" in Europe. The results show that the best prediction algorithm is ANN-Artificial Neural Network with an estimated growth value of 12.53%. Finally, "*augmented data*" were obtained using the ANN-Artificial Neural Network, through which a new prediction was made which estimated a growing value of the estimated variable equal to 3.18%.

Figure Index

Bibliography

[1] P. Polycronidou, V. Zoumpoulidis e S. Valsamidis, «Labour Digitalization in Europe,» *Intellectual Economics,* vol. 15, n. 2, 2021.

[2] S. Szilárd, A. Benedek e L. Ionel-Cioca, «Soft skills development needs and methods in micro-companies of ICT sector,» *Procedia-Social and Behavioral Sciences,* vol. 238, pp. 94-103, 2018.

[3] N. Abramova e N. Grishchenko, «ICTs, labour productivity and employment: Sustainability in industries in Russia,» *Procedia Manufacturing,* vol. 43, pp. 299-305, 2020.

[4] E. Herman, «The influence of ICT sector on the Romanian labour market in the European context,» *Procedia Manufacturing,* vol. 46, pp. 344-351, 2020.

[5] M. Gellerstedt, S. M. Babaheidari e L. Svensson, «A first step towards a model for teachers' adoption of ICT pedagogy in schools,» *Heliyon,* vol. 9, n. e00786, p. 4, 2018.

[6] V. Bilozubenko, O. Yatchuk, E. Wolanin e M. Korneyev, «Comparison of the digital economy development parameters in the EU countries in the context of bridging the digital divide,» *Problems and Perspectives in Management,* vol. 18, n. 2, p. 206, 2020.

[7] C. Ciurea e F. G. Filip, «Identifying business models for re-use of cultural objects by using modern ICT tools,» *Informatica Economica,* vol. 22, n. 1, pp. 68-75, 2018.

[8] K. Pažur Aničić, B. Divjak e K. Arbanas, «Prestige and Collaboration Among Researchers in the Field of Education and Career Development of ICT Graduates: Is There a Cross-Fertilization of Research and Knowledge?,» *Journal of Information and Organizational Sciences,* vol. 41, n. 2, pp. 231-262, 2017.

[9] I. Huđek, K. Širec e P. Tominc, «Digital skills in enterprises according to the European digital entrepreneurship sub-indices: Cross-country empirical evidence,» *Management: Journal of Contemporary Management Issues,* vol. 24, n. 2, pp. 107-119, 2019.

[10] R. Bologa e A. R. Lupu, «A Model for Predicting Future Demand for ICT Specialists in Romania,» *Economic Computation & Economic Cybernetics Studies & Research,* vol. 51, n. 1, 2017 .

[11] S. Marinković, I. Nikolić e J. Rakićević, «Selecting location for a new business unit in ICT industry,» *Zbornik Radova Ekonomski Fakultet u Rijeka,* vol. 36, n. 2, pp. 801-825, 2018.

[12] N. Trushkina, «Development of the information economy under the conditions of global economic transformations: features, factors and prospects,» *Virtual Economics,* vol. 2, n. 4, pp. 7-25, 2019.

[13] A. Zečević e J. Radović-Stojanović, «The use of information and communication technologies in enterprises in Serbia,» *Ekonomika preduzeća,* vol. 65, n. 5-6, pp. 393-402, 2017.

[14] A. I. Vodă e L. D. Radu, «Investigating Economic Factors of Sustainability in European Smart Cities,» *European Journal of Sustainable Development,* vol. 7, n. 1, pp. 107-107, 2018.

[15] Y. Bilan, H. Mishchuk, I. Roshchyk e I. Kmecova, «An Analysis of Intellecutal Potential and its Impact on the Social and Economic Development of European Countries,» *Journal of Competitiveness,* vol. 12, n. 1, p. 22–38, 2020.

[16] I. Yatskiv, «Why don't women choose stem? Gender equality in stem careers in Latvia,» *International Journal on Information Technologies & Security,* vol. 1, pp. 79-88, 2017.

[17] C. Y. Luk, «A Human Rights-Based Approach to Bridge Gender Digital Divide: The Case Study of India,» *Gender Gaps and the Social Inclusion Movement in ICT IGI Global,* pp. 24-44, 2019.

[18] A. Marks, S. Chillas, L. Galloway e G. Maclean, «Confusion and Collectivism in the ICT Sector: Is FLOSS the answer?,» *Economic and Industrial Democracy,* vol. 41, n. 1, pp. 167-188, 2020.

[19] M. A. Morozov, N. S. Morozova e M. Moldazhanov, «Innovative training in digital transformation,» *Advances in Economics, Business and Management Research. Proceedings of the International Conference on Economics, Management and Technologies ,* p. 499, 2020.

[20] M. D. Magomedov, O. V. Karabanova e V. V. Stroev, «Impact of Information and Communication Technologies Development on the Labor Market in the Russian Federation,» *2nd International Scientific and Practical Conference on Digital Economy,* pp. 297-300, 2020.

[21] A. Leogrande, A. Costantiello e L. Laureti, «The Broadband Penetration in Europe,» *Journal of Applied Economic Sciences (JAES),* vol. 7, pp. 324-349, 2021.

[22] A. Leogrande, N. Magaletti, G. Cosoli e A. Massaro, «Fixed Broadband Take-Up in Europe,» *SSRN 4034298.,* 2022.

[23] A. Leogrande, N. Magaletti, G. Cosoli e A. Massaro, «Broadband Price Index in Europe,» 2022.

[24] A. Leogrande, N. Magaletti, G. Cosoli e A. Massaro, «e-Government in Europe. A Machine Learning Approach,» 2022.

[25] A. Massaro, «Information Technology Infrastructures Supporting Industry 5.0 Facilities,» 2022.

8. Authors biography

Alessandro Massaro. Professor Alessandro Massaro (ING/INF/01, FIS/01, FIS/03) carried out scientific

research at the Polytechnic University of Marche, at CNR, and at Italian Institute of Technology (IIT) as Team Leader by activating laboratories for nanocomposite sensors for industrial robotics. He is in MIUR register as scientific expert in competitive Industrial Research and social development. He was the head of the Research and Development section and scientific director of MIUR Research Institute Dyrecta Lab S.r.l. Member of the International Scientific Committee of Measurers IMEKO and IEEE Senior member, he received an award from the National Council of Engineers as Best Engineer of Italy 2018 (Top Young Engineer 2018). He is currently researcher at LUM Enterprise S.r.l., and professor at LUM University Libera Università Mediterranea "*Giuseppe Degennaro*".

Nicola Magaletti, Business development manager graduated in mechanical engineering with over 30 years of work experience in structured companies offering IT solutions and consultancy services, working on the management of innovation processes and in the launch of new business initiatives. Since 2018 he has been part of the Lum Enterprise team as Operational Manager and Technical-Scientific Manager of the "Smart District 4.0" R&D project of which the company is the lead partner.

Gabriele Cosoli, Senior IT Specialist and Solution architect with a degree in Computer Science and over five years of previous experience, specialized in the analysis and design of ICT solutions in various application areas and technological frameworks. Certified on "*Machine Learning by Stanford University on Coursera*" Master on "*Agile and Digital Project Management - Advanced Course*" at 24ORE Business School.

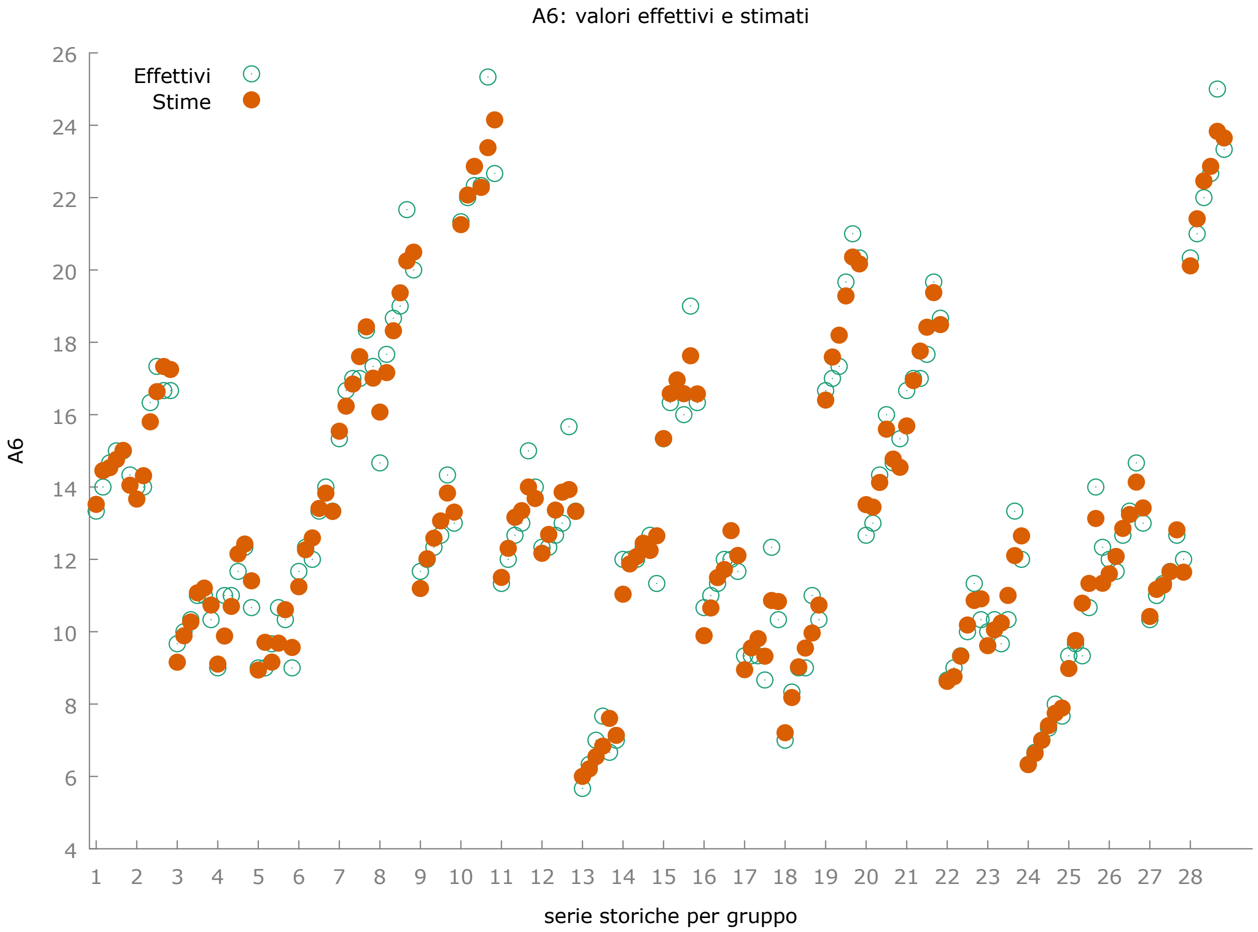
Giardinelli Vito O. M.. Project Manager with a degree in Economics and Management (2016, Lum University, Casamassima, Italy) and a Master in Business Strategy and Entrepreneurship (2017, Sda Bocconi, Milan, Italy). Experience: (i) 2014, staff audit, KPMG S.p.A.; (ii) from 2017 to 2020, project manager, Facile.it (Italian price comparison web-site); (iii) from 2021, Lum Enterprise, business developer.

Angelo Leogrande is assistant professor at the LUM University-Giuseppe Degennaro in Casamassima-Bari, where he is also researcher for the LUM Enterprise S.r.l. a spin off oriented to develop digitalization services for SMEs. He worked at Dyrecta Lab Research Institute, a research institute officially recognized by the Italian Minister of University and Research-MIUR, where he acquired professional hard skills on the economic consequences of Industry 4.0, Big Data, Machine Learning and Artificial Intelligence. His research interests include Cooperative Banking, Business Ethics, Innovation Technology, Knowledge, and R&D.

9. Appendix

Modello 35: Effetti fissi, usando 168 osservazioni					
Incluse 28 unità cross section					
Lunghezza serie storiche = 6					
Variabile dipendente: A6					
	Coefficiente	*Errore Std.*	*rapporto t*	*p-value*	
const	6,13492	1,83114	3,350	0,0011	***
A1	−0,594517	0,120448	−4,936	<0,0001	***
A2	−0,349722	0,0656706	−5,325	<0,0001	***
A4	−0,704928	0,186265	−3,785	0,0002	***
A5	−0,366473	0,0830139	−4,415	<0,0001	***
A17	−1,76117	0,296586	−5,938	<0,0001	***
A19	1,39380	0,236717	5,888	<0,0001	***
A38	0,108615	0,0455436	2,385	0,0186	**
A40	−0,113020	0,0334273	−3,381	0,0010	***
A41	−0,185417	0,0259188	−7,154	<0,0001	***
A42	−0,174146	0,0281350	−6,190	<0,0001	***
A44	−0,394876	0,0804158	−4,910	<0,0001	***
A47	0,162786	0,0471308	3,454	0,0007	***
Media var. dipendente	13,32936	SQM var. dipendente		4,180713	

Somma quadr. residui	56,58072	E.S. della regressione	0,664859
R-quadro LSDV	0,980616	R-quadro intra-gruppi	0,734731
LSDV F(39, 128)	166,0325	P-value(F)	2,15e-92
Log-verosimiglianza	−146,9648	Criterio di Akaike	373,9297
Criterio di Schwarz	498,8882	Hannan-Quinn	424,6439
rho	−0,182317	Durbin-Watson	1,977891
Test congiunto sui regressori -			
Statistica test: F(12, 128) = 29,5441			
con p-value = P(F(12, 128) > 29,5441) = 3,23155e-031			
Test per la differenza delle intercette di gruppo -			
Ipotesi nulla: i gruppi hanno un'intercetta comune			
Statistica test: F(27, 128) = 26,1773			
con p-value = P(F(27, 128) > 26,1773) = 7,47784e-040			



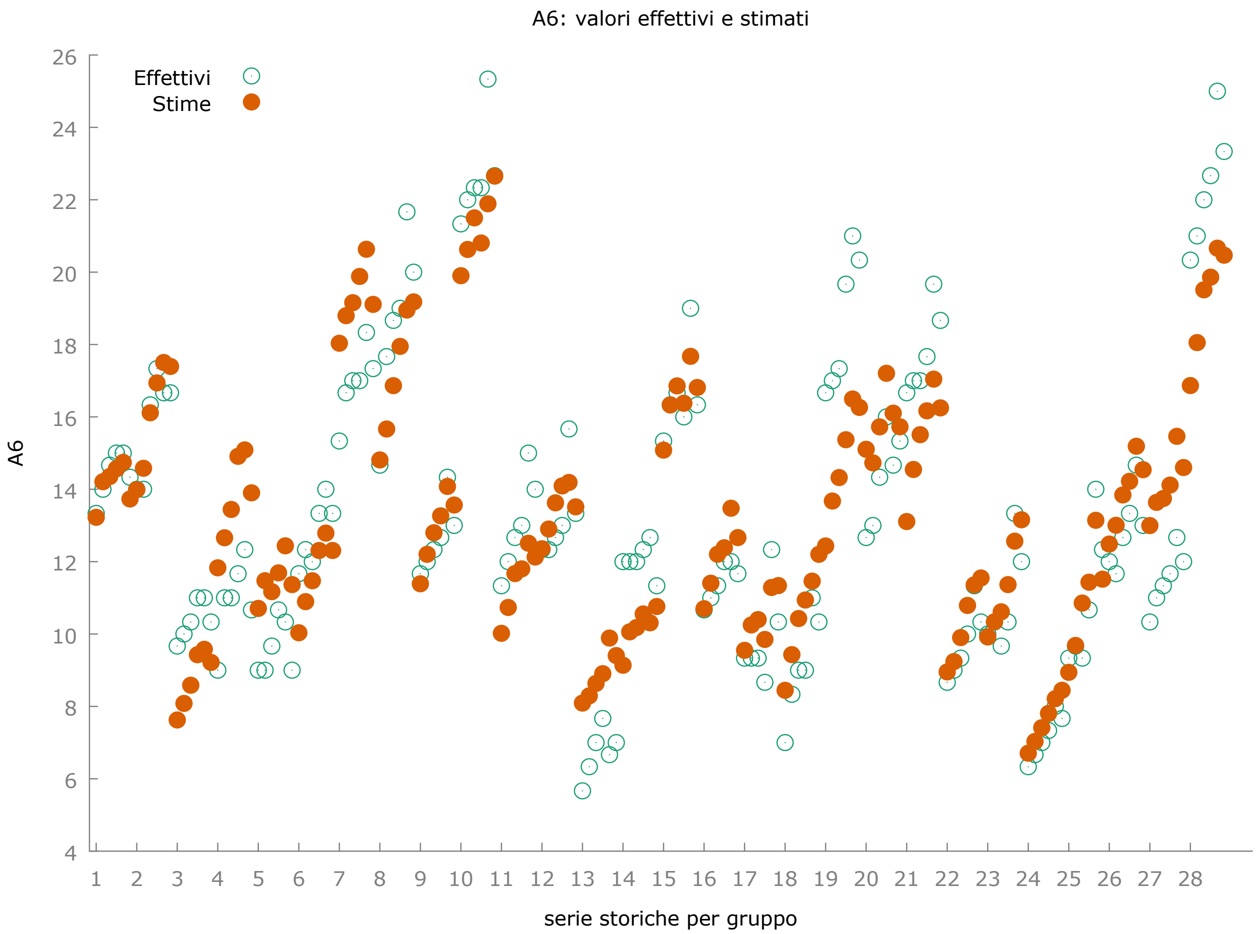
Modello 36: Effetti casuali (GLS), usando 168 osservazioni
Incluse 28 unità cross section
Lunghezza serie storiche = 6
Variabile dipendente: A6

	Coefficiente	*Errore Std.*	*z*	*p-value*	
const	3,13848	1,56794	2,002	0,0453	**
A1	−0,612614	0,111128	−5,513	<0,0001	***
A2	−0,384705	0,0592778	−6,490	<0,0001	***
A4	−0,622457	0,149989	−4,150	<0,0001	***
A5	−0,372053	0,0656354	−5,668	<0,0001	***
A17	−1,92543	0,271870	−7,082	<0,0001	***
A19	1,47561	0,207544	7,110	<0,0001	***
A38	0,111806	0,0425637	2,627	0,0086	***
A40	−0,130245	0,0316477	−4,115	<0,0001	***
A41	−0,198733	0,0240824	−8,252	<0,0001	***
A42	−0,181368	0,0256019	−7,084	<0,0001	***
A44	−0,336020	0,0683836	−4,914	<0,0001	***
A47	0,160244	0,0326328	4,911	<0,0001	***

Media var. dipendente	13,32936	SQM var. dipendente	4,180713
Somma quadr. residui	529,4704	E.S. della regressione	1,842292
Log-verosimiglianza	−334,8064	Criterio di Akaike	695,6128
Criterio di Schwarz	736,2243	Hannan-Quinn	712,0949
rho	−0,182317	Durbin-Watson	1,977891

Varianza 'between' = 2,71852
Varianza 'within' = 0,442037
Theta usato per la trasformazione = 0,837564
Test congiunto sui regressori -
Statistica test asintotica: Chi-quadro(12) = 468,917
con p-value = 9,04723e-093
Test Breusch-Pagan -
Ipotesi nulla: varianza dell'errore specifico all'unità = 0
Statistica test asintotica: Chi-quadro(1) = 203,247
con p-value = 4,08594e-046

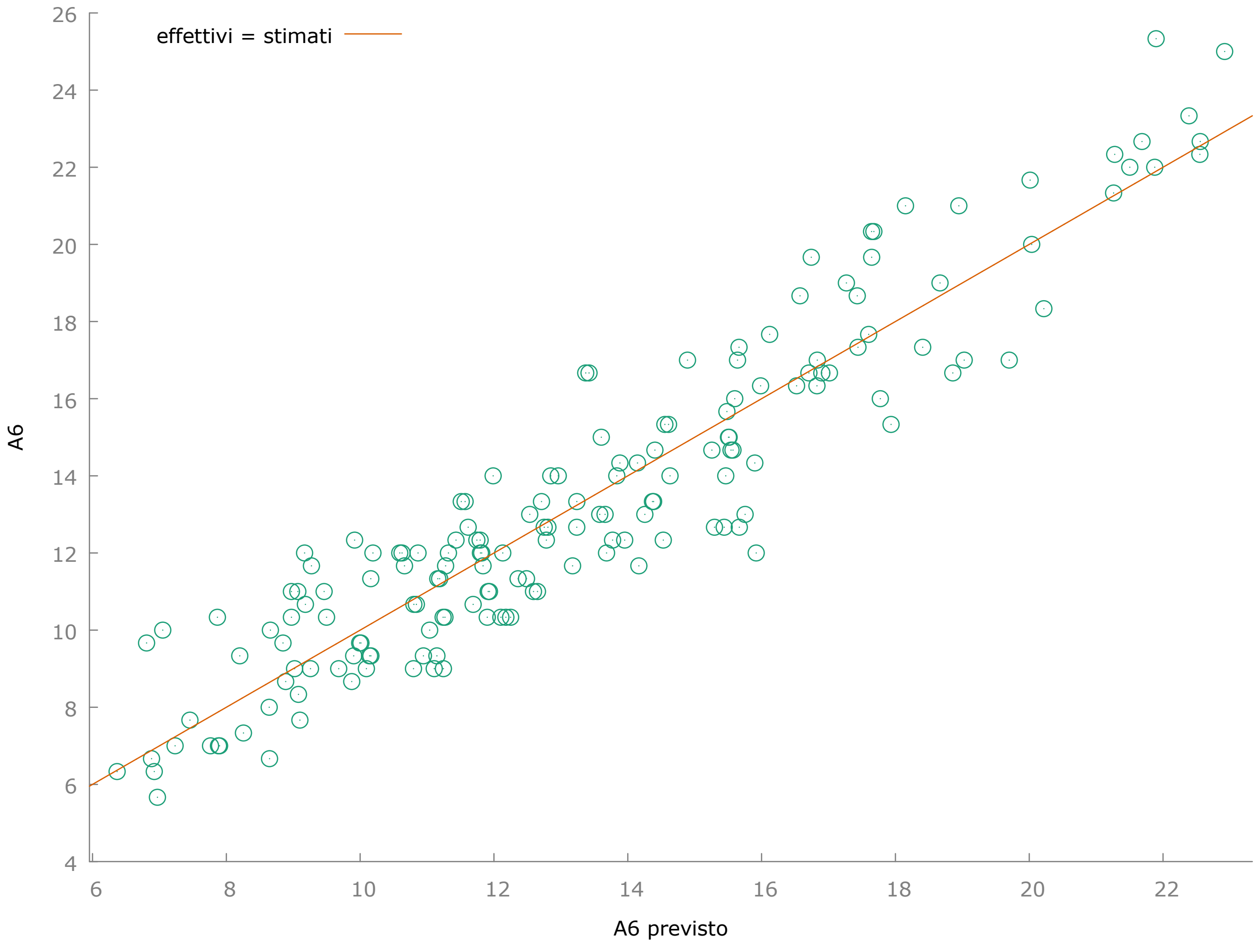
Test di Hausman -
Ipotesi nulla: le stime GLS sono consistenti
Statistica test asintotica: Chi-quadro(12) = 19,9674
con p-value = 0,0677047



Modello 37: Pooled OLS, usando 168 osservazioni
Incluse 28 unità cross section
Lunghezza serie storiche = 6
Variabile dipendente: A6

	Coefficiente	Errore Std.	rapporto t	p-value	
const	−1,94986	1,29462	−1,506	0,1341	
A1	−0,573264	0,106922	−5,362	<0,0001	***
A2	−0,614085	0,0615603	−9,975	<0,0001	***
A4	−0,463327	0,110125	−4,207	<0,0001	***
A5	−0,464343	0,0487181	−9,531	<0,0001	***
A17	−2,65698	0,260098	−10,22	<0,0001	***
A19	1,93037	0,169431	11,39	<0,0001	***
A38	0,209739	0,0458798	4,571	<0,0001	***
A40	−0,170921	0,0520026	−3,287	0,0013	***
A41	−0,275466	0,0314993	−8,745	<0,0001	***
A42	−0,224362	0,0361862	−6,200	<0,0001	***
A44	−0,359648	0,0572243	−6,285	<0,0001	***
A47	0,143272	0,0197081	7,270	<0,0001	***

Media var. dipendente	13,32936	SQM var. dipendente	4,180713
Somma quadr. residui	369,0064	E.S. della regressione	1,542947
R-quadro	0,873580	R-quadro corretto	0,863792
F(12, 155)	89,25578	P-value(F)	3,55e-63
Log-verosimiglianza	−304,4771	Criterio di Akaike	634,9542
Criterio di Schwarz	675,5657	Hannan-Quinn	651,4363
rho	0,728077	Durbin-Watson	0,381458



Modello 38: WLS, usando 168 osservazioni					
Incluse 28 unità cross section					
Variabile dipendente: A6					
Pesi basati sulle varianze degli errori per unità					
	Coefficiente	*Errore Std.*	*rapporto t*	*p-value*	
const	−1,92196	0,903345	−2,128	0,0350	**
A1	−0,616825	0,0839079	−7,351	<0,0001	***
A2	−0,601152	0,0392591	−15,31	<0,0001	***
A4	−0,455733	0,0744274	−6,123	<0,0001	***
A5	−0,462679	0,0370624	−12,48	<0,0001	***
A17	−2,67555	0,209245	−12,79	<0,0001	***
A19	1,93650	0,127415	15,20	<0,0001	***
A38	0,190809	0,0270234	7,061	<0,0001	***
A40	−0,200905	0,0341203	−5,888	<0,0001	***

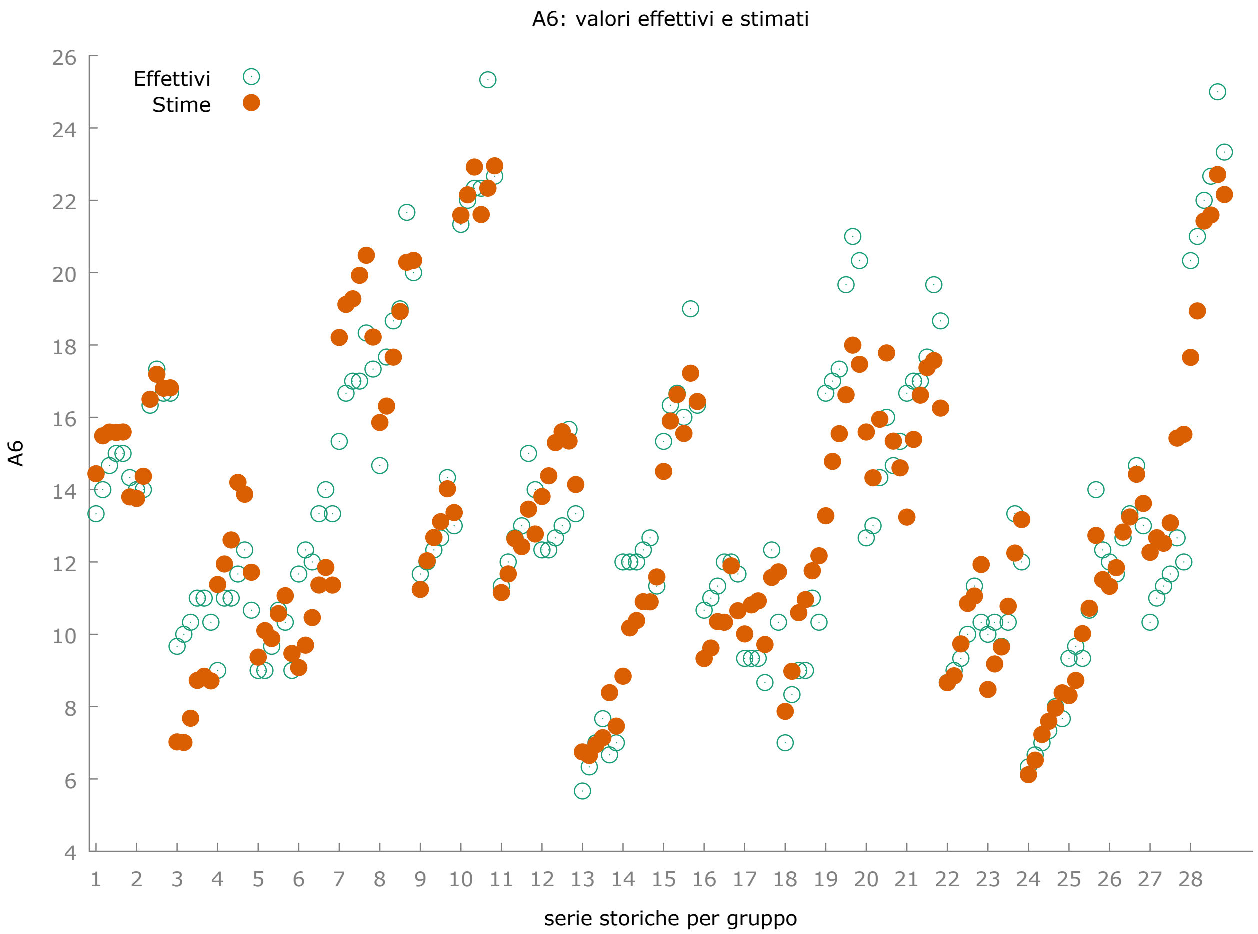
A41	−0,279457	0,0201039	−13,90	<0,0001	***
A42	−0,235747	0,0311228	−7,575	<0,0001	***
A44	−0,359840	0,0407181	−8,837	<0,0001	***
A47	0,153883	0,0146571	10,50	<0,0001	***

Statistiche basate sui dati ponderati:

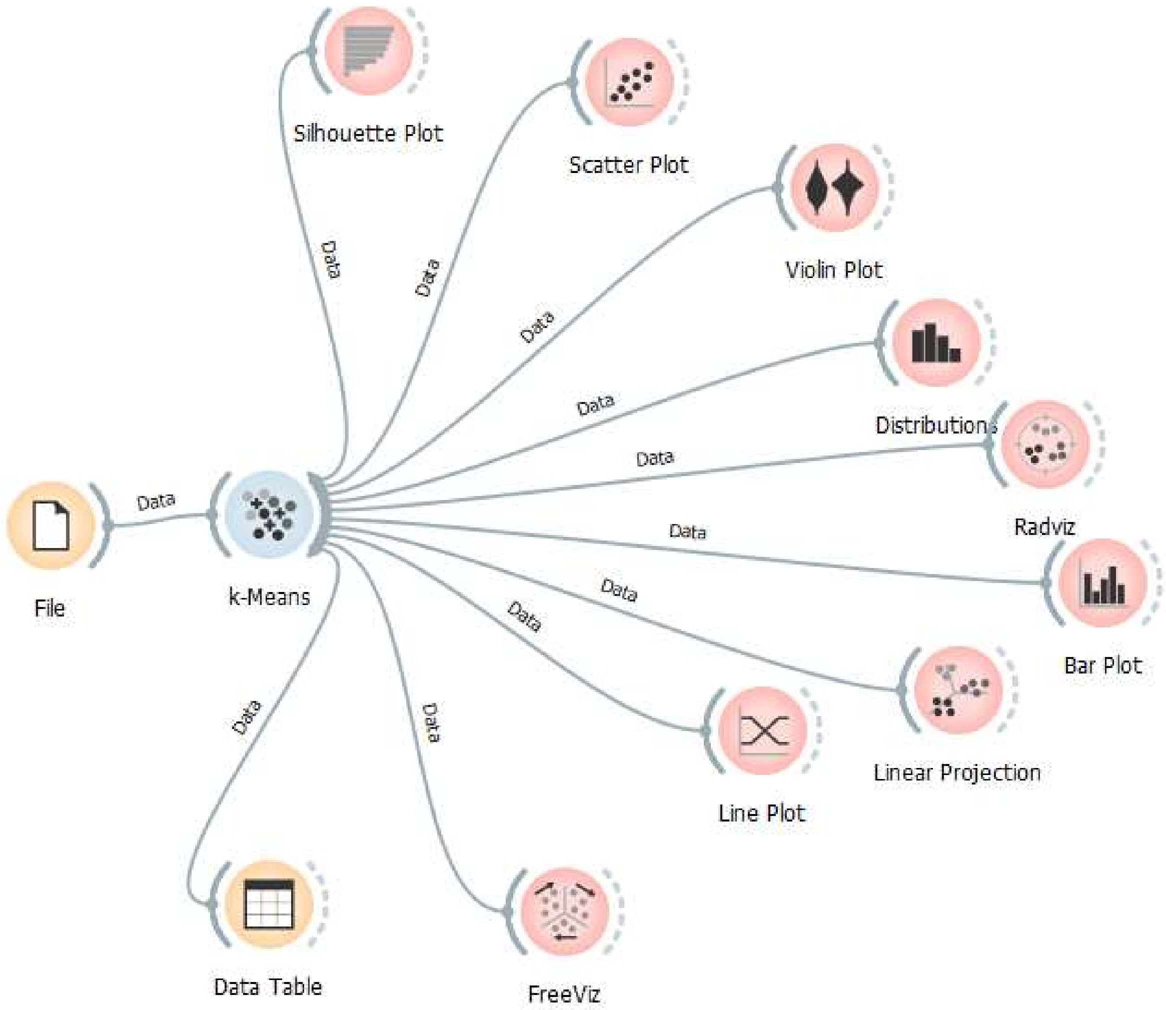
Somma quadr. residui	157,9333	E.S. della regressione	1,009418
R-quadro	0,931990	R-quadro corretto	0,926725
F(12, 155)	177,0074	P-value(F)	6,65e-84
Log-verosimiglianza	−233,1912	Criterio di Akaike	492,3824
Criterio di Schwarz	532,9939	Hannan-Quinn	508,8645

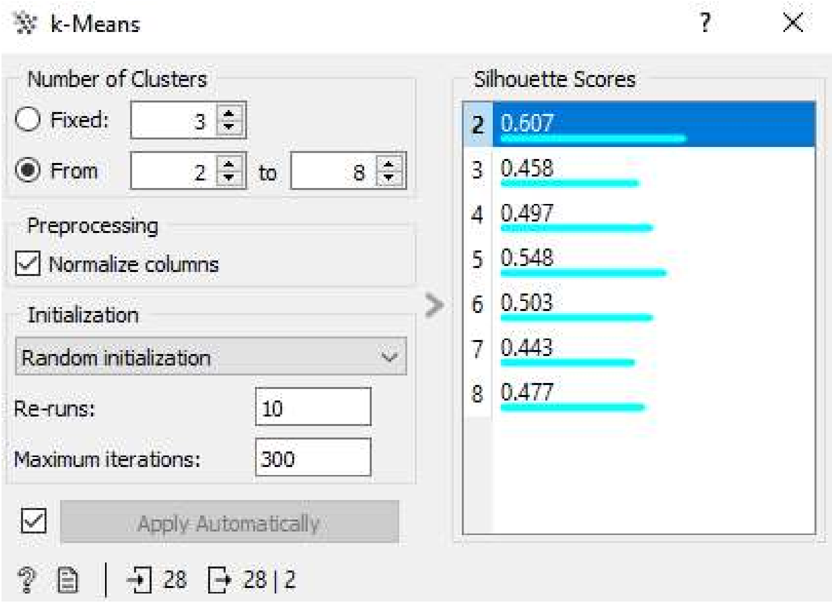
Statistiche basate sui dati originali:

Media var. dipendente	13,32936	SQM var. dipendente	4,180713
Somma quadr. residui	378,9162	E.S. della regressione	1,563528

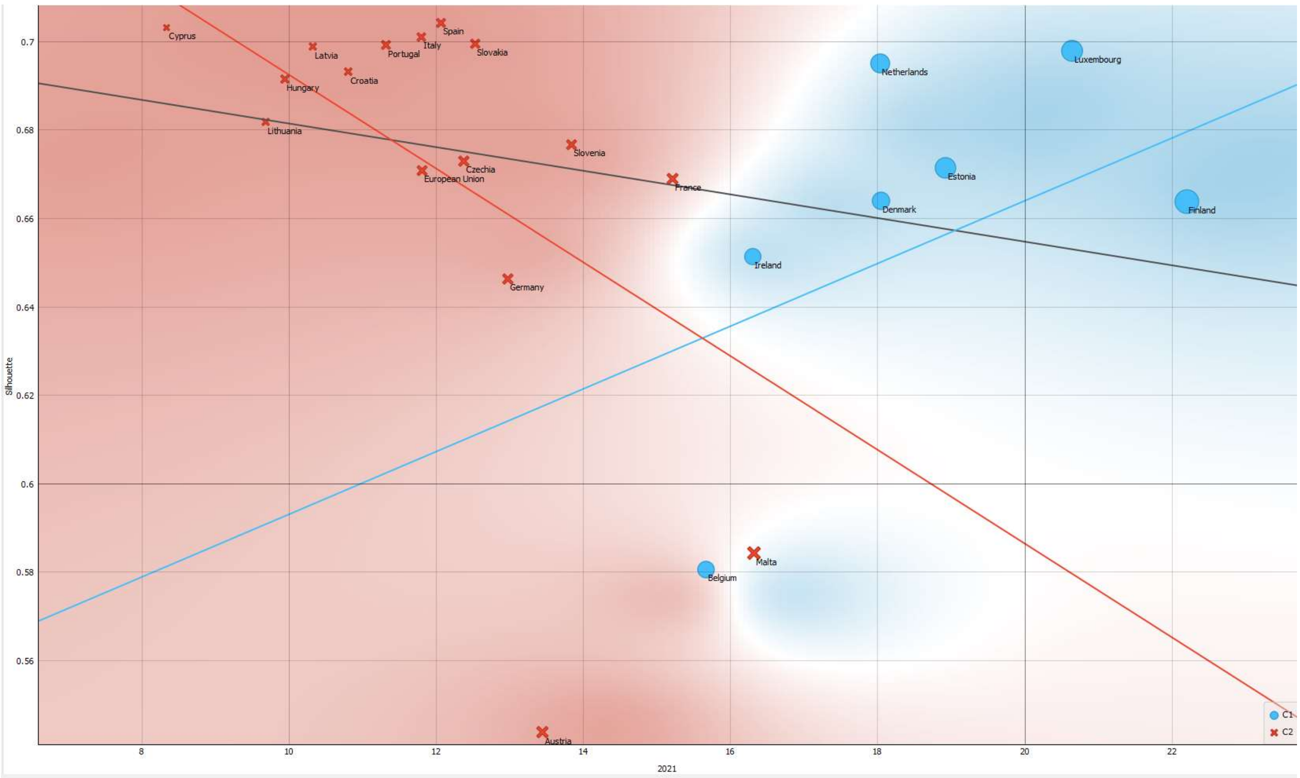


8.1 Clusterization

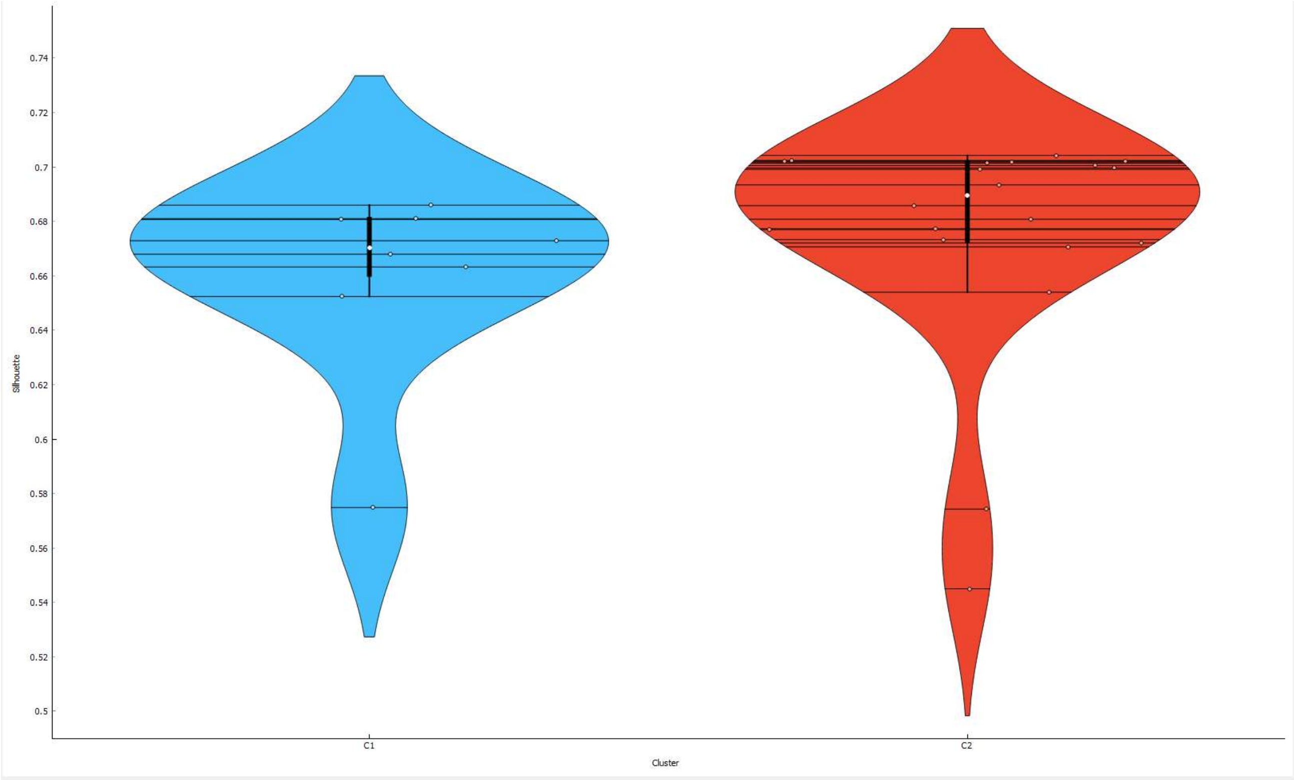

Silhouette Plot
Scatter Plot
Violin Plot
Distributions
Radviz
Bar Plot
Linear Projection
Line Plot
FreeViz
Data Table
File
k-Means
Data

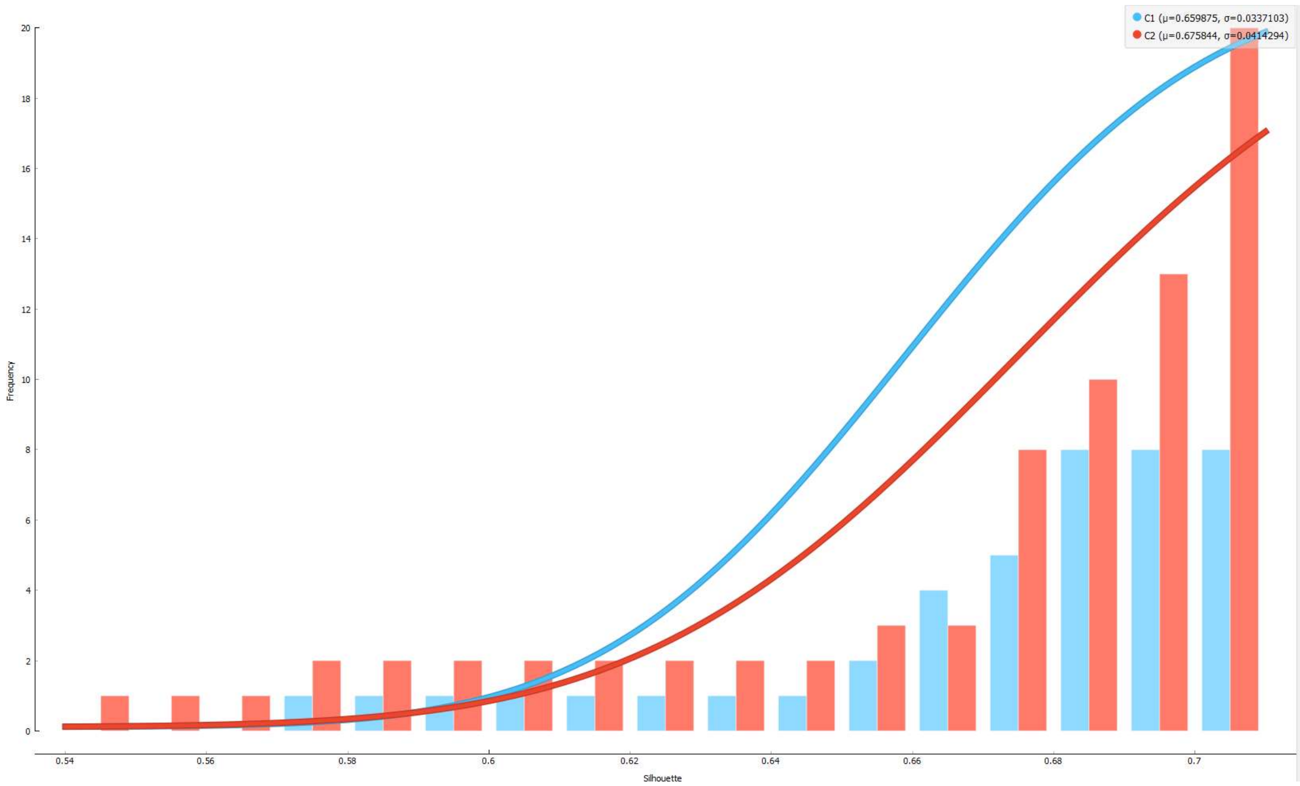


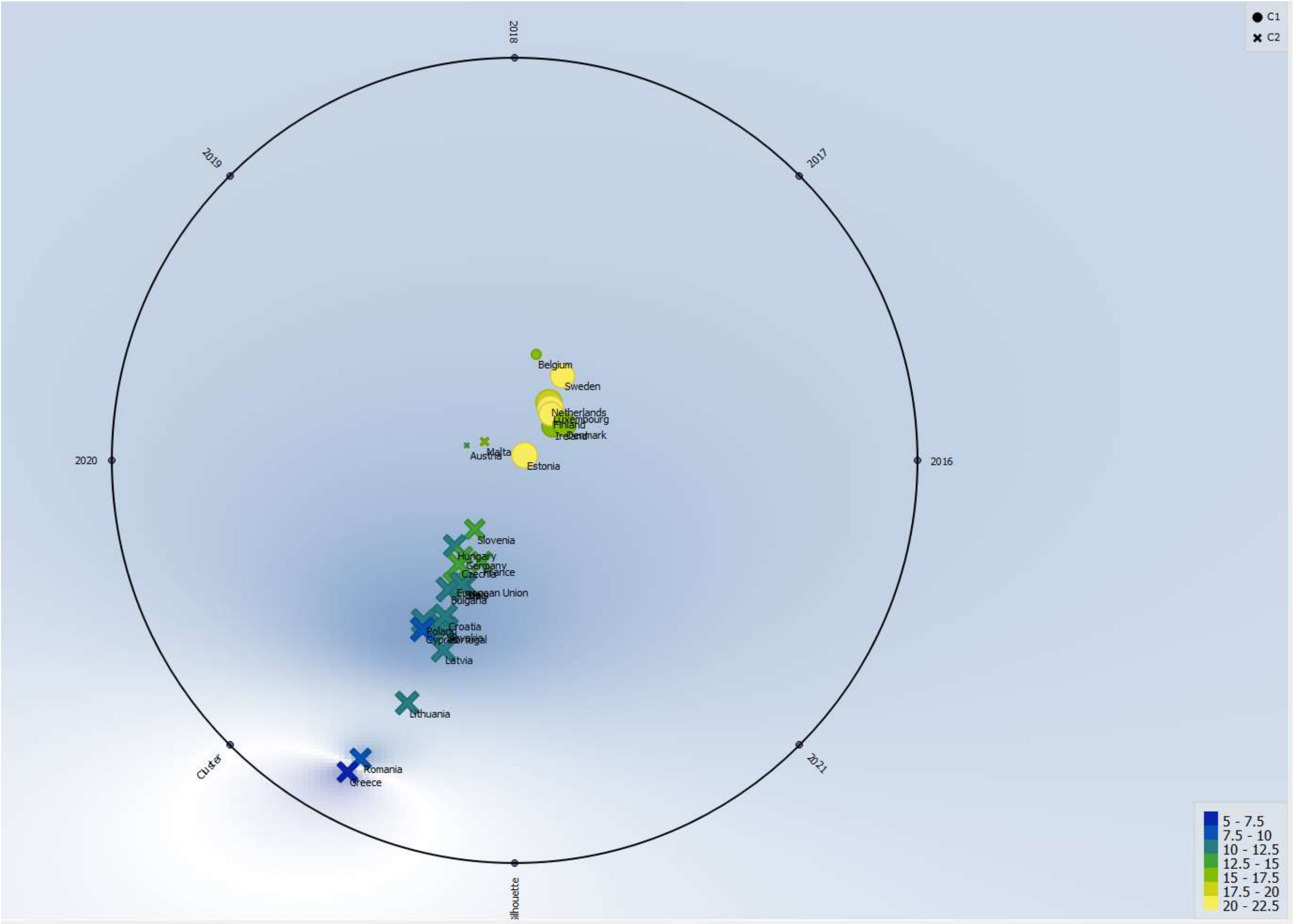
k-Means
Number of Clusters
Fixed: 3
From 2 to 8
Preprocessing
Normalize columns
Initialization
Random initialization
Re-runs: 10
Maximum iterations: 300
Apply Automatically
28
28 | 2
Silhouette Scores
2 0.607
3 0.458
4 0.497
5 0.548
6 0.503
7 0.443
8 0.477

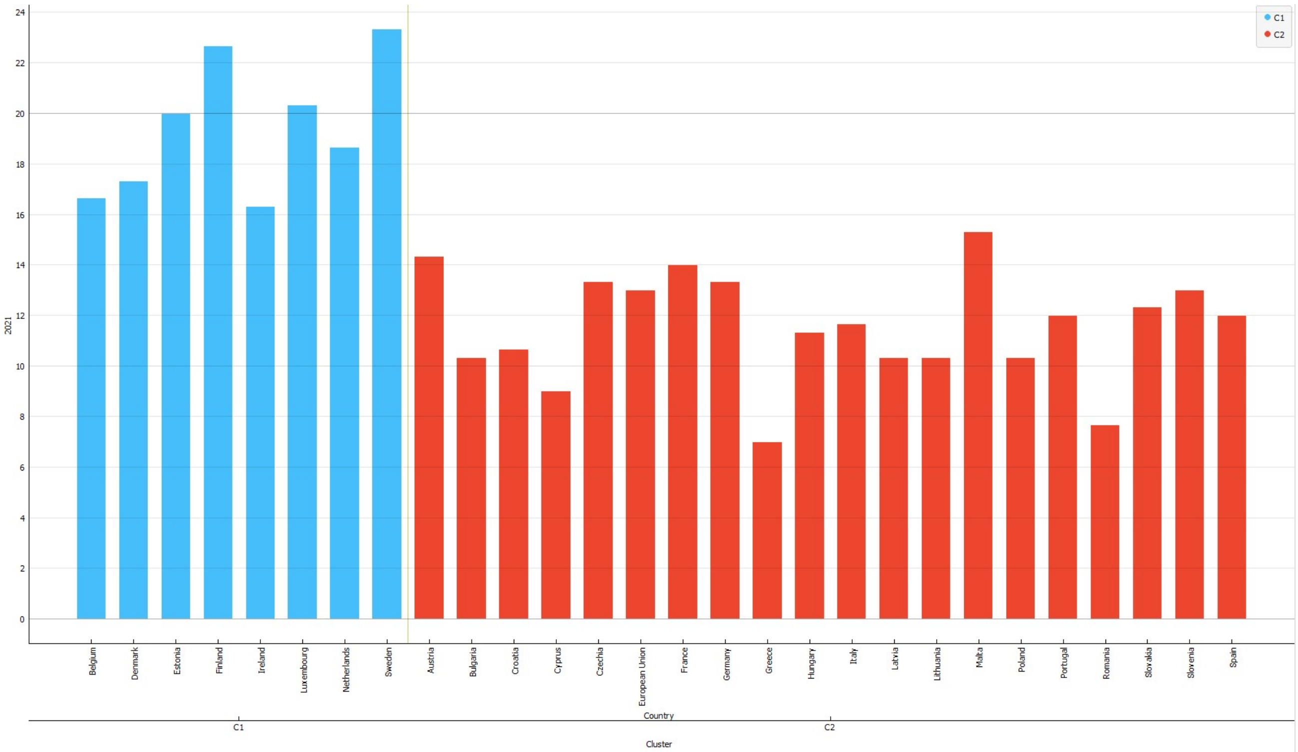


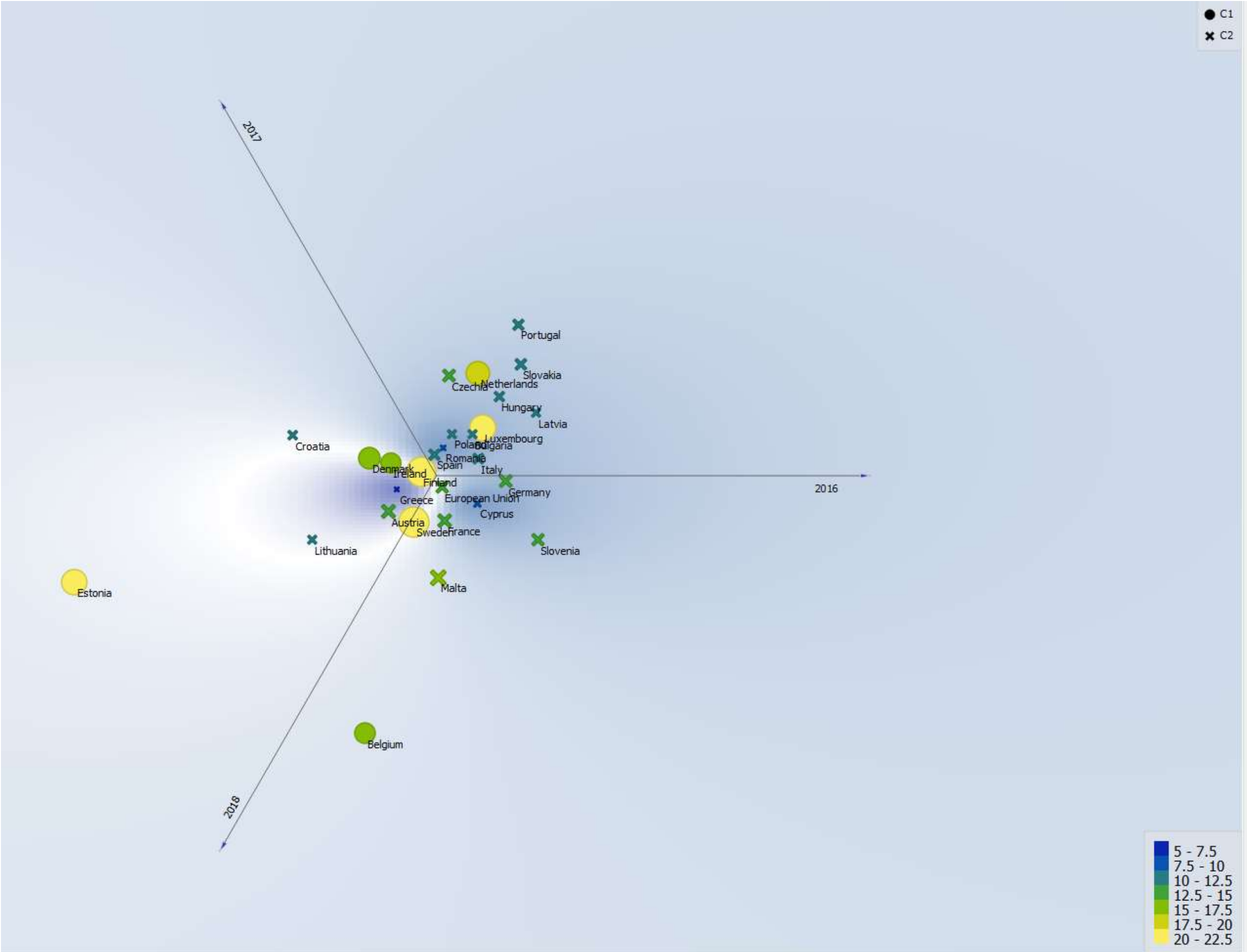
Cyprus
Latvia
Portugal
Italy
Spain
Slovakia
Hungary
Croatia
Lithuania
Czechia
European Union
Slovenia
France
Germany
Netherlands
Luxembourg
Estonia
Denmark
Finland
Ireland
Belgium
Malta
Austria
C1
C2
Silhouette
2021

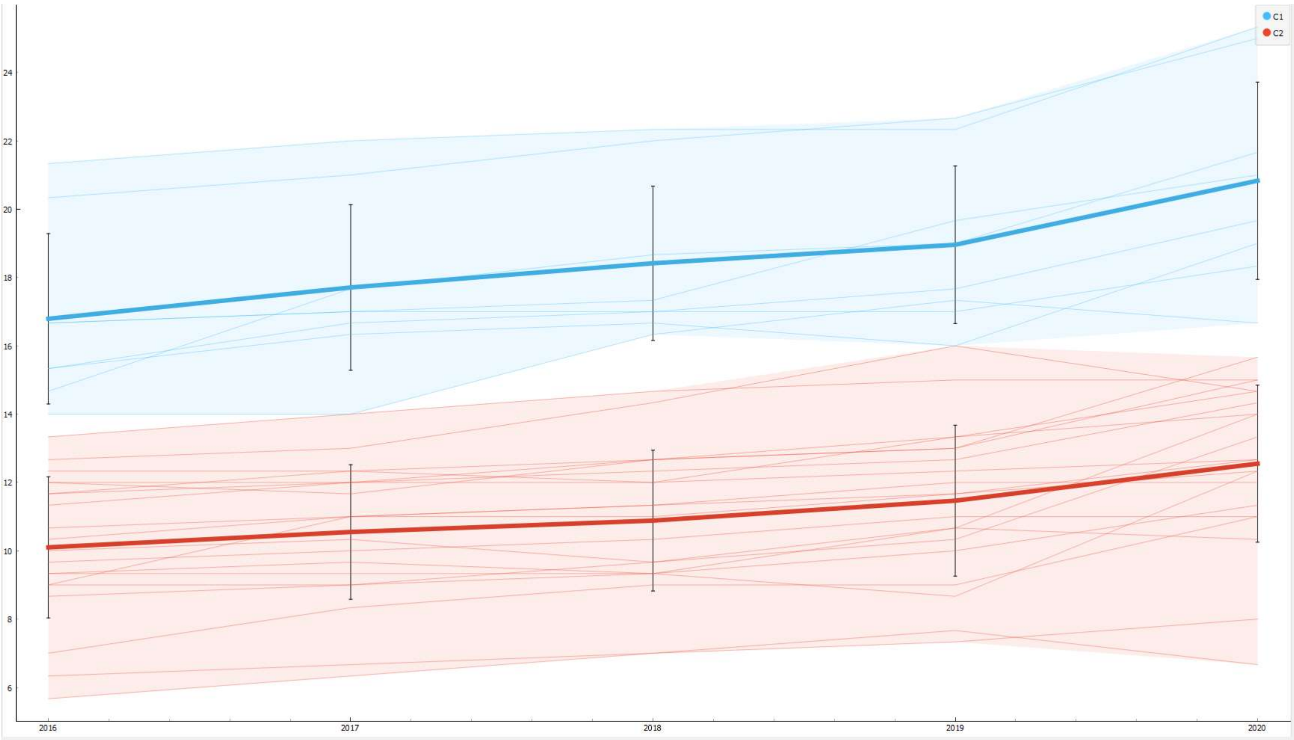

0.74
0.72
0.7
0.68
0.66
0.64
0.62
0.6
0.58
0.56
0.54
0.52
0.5
Silhouette
C1
C2
Cluster

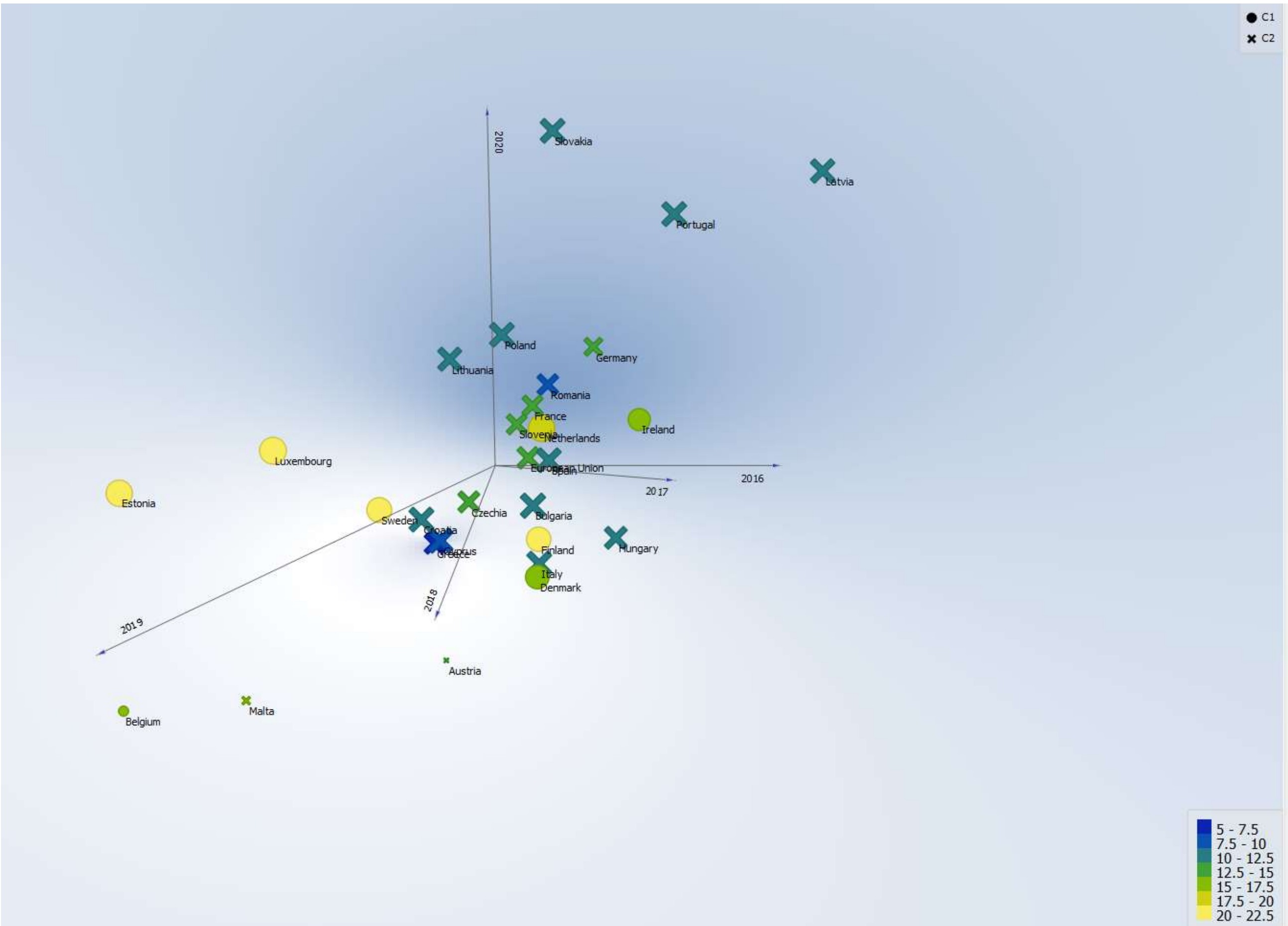

C1 (μ=0.659875, σ=0.0337103)
C2 (μ=0.675844, σ=0.0414294)
20
18
16
14
12
10
8
6
4
2
0
Frequency
0.54
0.56
0.58
0.6
0.62
0.64
0.66
0.68
0.7
Silhouette


C1
C2
2018
2017
2016
2021
Silhouette
Cluster
2020
2019
Belgium
Sweden
Netherlands
Luxembourg
Finland
Ireland
Denmark
Malta
Austria
Estonia
Slovenia
Hungary
Germany
Czechia
France
European Union
Bulgaria
Croatia
Poland
Cyprus
Portugal
Latvia
Lithuania
Romania
Greece
5 - 7.5
7.5 - 10
10 - 12.5
12.5 - 15
15 - 17.5
17.5 - 20
20 - 22.5

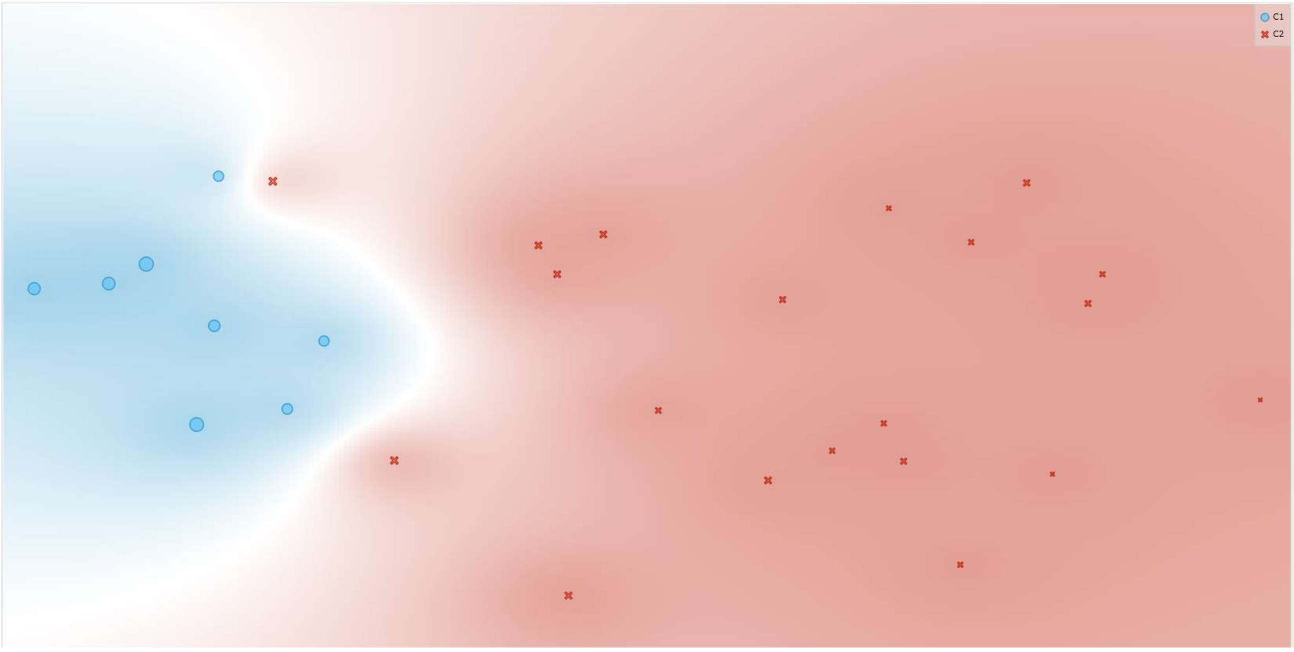

C1
C2
2021
Belgium
Denmark
Estonia
Finland
Ireland
Luxembourg
Netherlands
Sweden
Austria
Bulgaria
Croatia
Cyprus
Czechia
European Union
France
Germany
Greece
Hungary
Italy
Latvia
Lithuania
Malta
Poland
Portugal
Romania
Slovakia
Slovenia
Spain
Country
C1
C2
Cluster

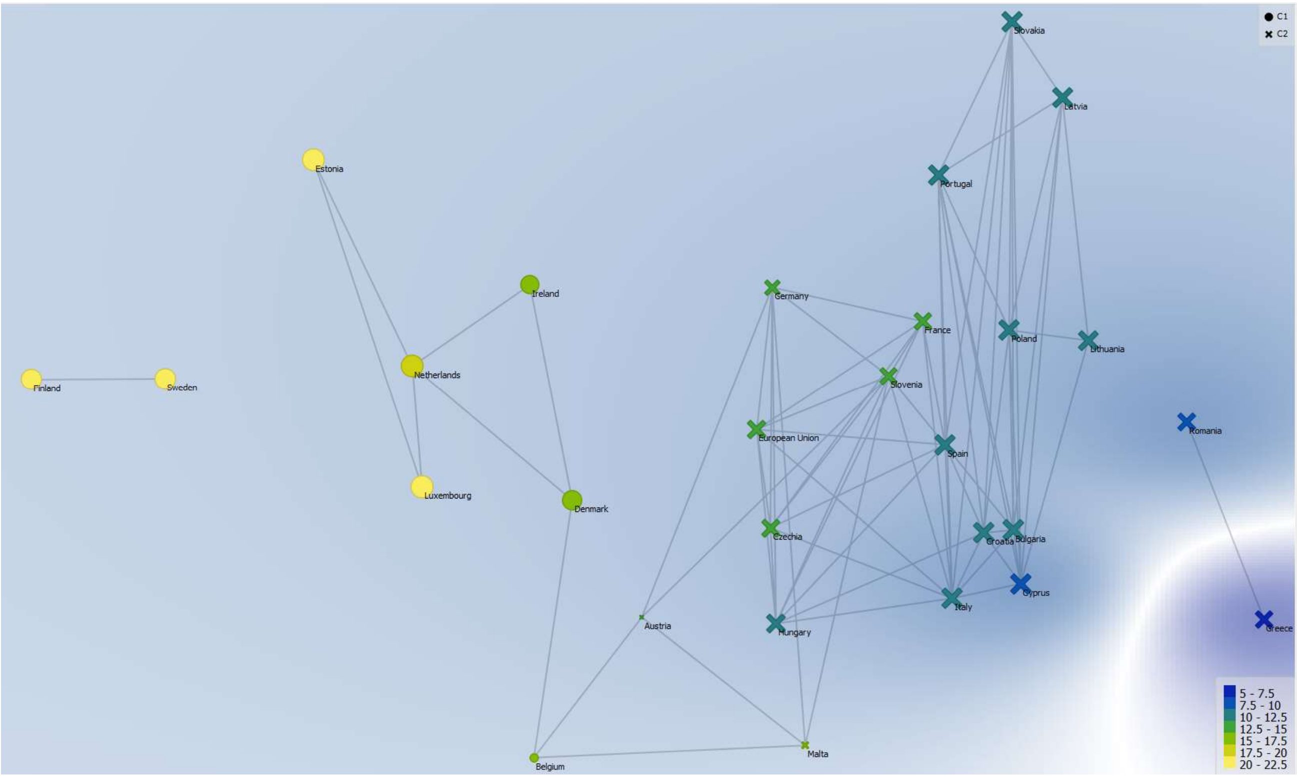

C1
C2
2017
2016
2018
Portugal
Slovakia
Czechia
Netherlands
Hungary
Latvia
Croatia
Luxembourg
Poland
Bulgaria
Romania
Spain
Italy
Denmark
Ireland
Finland
Germany
Greece
European Union
Cyprus
Austria
Sweden
France
Lithuania
Slovenia
Estonia
Malta
Belgium
5 - 7.5
7.5 - 10
10 - 12.5
12.5 - 15
15 - 17.5
17.5 - 20
20 - 22.5


C1
C2
24
22
20
18
16
14
12
10
8
6
2016
2017
2018
2019
2020

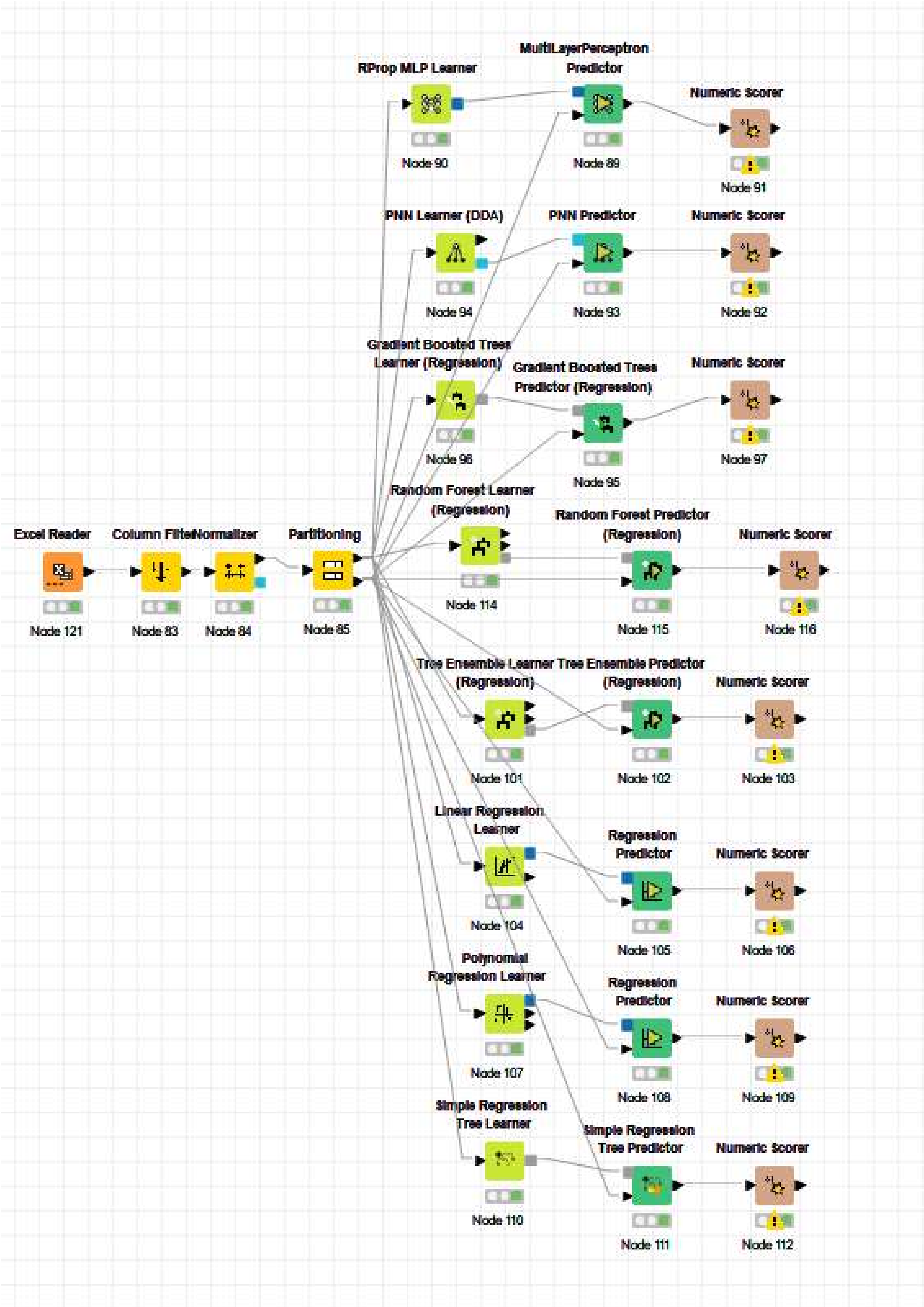

C1
C2
2020
2019
2018
2017
2016
Slovakia
Latvia
Portugal
Poland
Germany
Lithuania
Romania
France
Slovenia
Netherlands
Ireland
Luxembourg
European Union
Spain
Estonia
Czechia
Bulgaria
Sweden
Croatia
Cyprus
Greece
Finland
Hungary
Italy
Denmark
Austria
Malta
Belgium
5 - 7.5
7.5 - 10
10 - 12.5
12.5 - 15
15 - 17.5
17.5 - 20
20 - 22.5

	2021	Country	Cluster
2	16.6667	Belgium	C1
7	17.3333	Denmark	C1
8	20	Estonia	C1
10	22.6667	Finland	C1
15	16.3333	Ireland	C1
19	20.3333	Luxembourg	C1
21	18.6667	Netherlands	C1
28	23.3333	Sweden	C1
1	14.3333	Austria	C2
3	10.3333	Bulgaria	C2
4	10.6667	Croatia	C2
5	9	Cyprus	C2
6	13.3333	Czechia	C2
9	13	European Union	C2
11	14	France	C2
12	13.3333	Germany	C2
13	7	Greece	C2
14	11.3333	Hungary	C2
16	11.6667	Italy	C2
17	10.3333	Latvia	C2
18	10.3333	Lithuania	C2
20	15.3333	Malta	C2
22	10.3333	Poland	C2
23	12	Portugal	C2
24	7.66667	Romania	C2
25	12.3333	Slovakia	C2
26	13	Slovenia	C2
27	12	Spain	C2





8.3 Machine Learning and Predictions


RProp MLP Learner
MultiLayerPerceptron Predictor
Numeric Scorer
Node 90
Node 89
Node 91
PNN Learner (DDA)
PNN Predictor
Numeric Scorer
Node 94
Node 93
Node 92
Gradient Boosted Trees Learner (Regression)
Gradient Boosted Trees Predictor (Regression)
Numeric Scorer
Node 96
Node 95
Node 97
Random Forest Learner (Regression)
Random Forest Predictor (Regression)
Numeric Scorer
Excel Reader
Column Filter
Normalizer
Partitioning
Node 121
Node 83
Node 84
Node 85
Node 114
Node 115
Node 116
Tree Ensemble Learner (Regression)
Tree Ensemble Predictor (Regression)
Numeric Scorer
Node 101
Node 102
Node 103
Linear Regression Learner
Regression Predictor
Numeric Scorer
Node 104
Node 105
Node 106
Polynomial Regression Learner
Regression Predictor
Numeric Scorer
Node 107
Node 108
Node 109
Simple Regression Tree Learner
Simple Regression Tree Predictor
Numeric Scorer
Node 110
Node 111
Node 112

Performance of the algorithms in the prediction of ICT Specialists in Europe.				
	ANN	*PNN*	*Gradient Boosted Tree Regression*	*Random Forest Regression*
R^2	0,983015	0,981870	0,933903	0,931014
mean absolute error	0,028884	0,027210	0,063801	0,066713
mean squared error	0,001387	0,001481	0,005398	0,005634
root mean squared error	0,037246	0,038481	0,073474	0,075062
mean signed difference	0,007711	0,000000	0,015572	0,016010
mean absolute percentage error	0,091555	0,105584	0,266465	0,282925
	Tree Ensemble Regression	*Linear Rergession*	*Polynomial Regression*	*Simple Regression Treee*
R^2	0,934476	0,988741	0,980342	0,942206
mean absolute error	0,064332	0,020823	0,031035	0,054422
mean squared error	0,005352	0,030000	0,001606	0,004720
root mean squared error	0,073155	0,030325	0,040070	0,068705
mean signed difference	0,019026	0,002678	0,024548	0,031747
mean absolute percentage error	0,280708	0,107706	0,126414	0,250038

Ranking of algorithms by performance value with historical data.							
	R^2	*Mean absolute error*	*Mean squared error*	*Root mean squared error*	*Mean signed difference*	*Mean absolute percentage error*	*Total*
ANN	2	2	1	2	3	1	11
PNN	3	3	2	3	1	2	14
Linear Rergession	1	1	8	1	2	3	16
Polynomial Regression	4	4	3	4	7	4	26
Simple Regression Treee	5	5	4	5	8	5	32
Gradient Boosted Tree Regression	7	6	6	7	4	6	36
Tree Ensemble Regression	6	7	5	6	6	7	37
Random Forest Regression	8	8	7	8	5	8	44

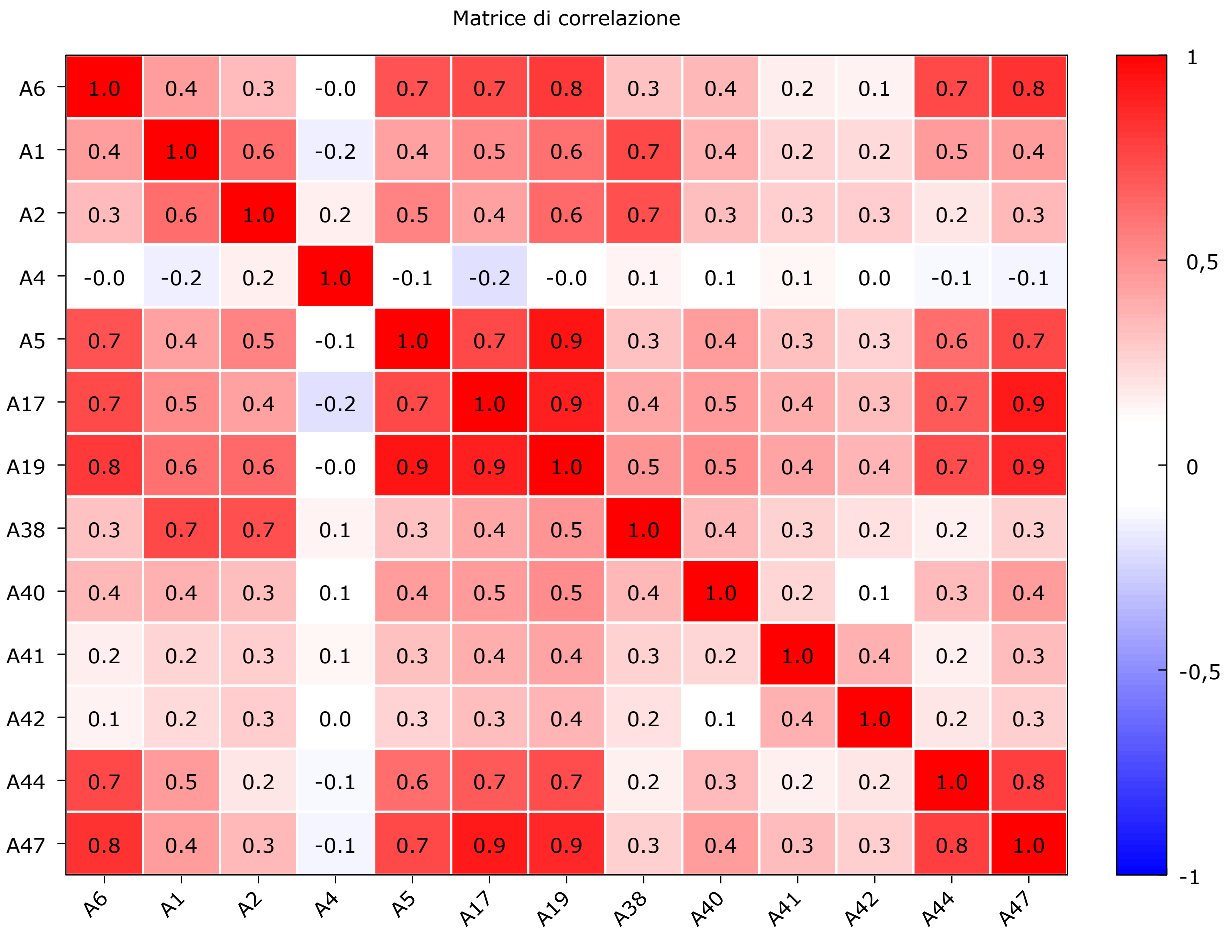
Comparison between historical values and predicted values through the use of the ANN-Artificial Neural Network.				
Country	**2021**	**Prediction**	**Absolute variation**	**Percentage Variation**
Austria	14,333	15,931	1,60	11,15
Belgium	16,667	16,438	-0,23	-1,37
Bulgaria	10,333	14,695	4,36	42,21
Croatia	10,667	14,874	4,21	39,44
Cyprus	9,000	14,563	5,56	61,81
Czechia	13,333	15,424	2,09	15,68
Denmark	17,333	16,659	-0,67	-3,89
Estonia	20,000	17,341	-2,66	-13,30
European Union	13,000	15,458	2,46	18,91
Finland	22,667	17,583	-5,08	-22,43
France	14,000	15,475	1,48	10,54
Germany	13,333	15,620	2,29	17,15
Greece	7,000	14,125	7,13	101,79
Hungary	11,333	15,168	3,83	33,84
Ireland	16,333	16,872	0,54	3,30
Italy	11,667	14,896	3,23	27,68
Latvia	10,333	14,666	4,33	41,93
Lithuania	10,333	14,440	4,11	39,74
Luxembourg	20,333	17,098	-3,24	-15,91
Malta	15,333	15,671	0,34	2,20
Netherlands	18,667	16,953	-1,71	-9,18
Poland	10,333	14,636	4,30	41,64
Portugal	12,000	15,011	3,01	25,09
Romania	7,667	14,176	6,51	84,90
Slovakia	12,333	15,023	2,69	21,81
Slovenia	13,000	15,437	2,44	18,75
Spain	12,000	15,006	3,01	25,05
Sweden	23,333	15,880	-7,45	-31,94
Mean	13,810	15,540	1,73	12,53

Results of the algorithms for predictive capacity through the use of augmented data.				
Metrics	*ANN*	*PNN*	*Gradient Boosted Tree Regression*	*Random Forest Regression*
R^2	0,952569406	0,97299908	0,936989251	0,949668858
Mean absolute error	0,036643809	0,036212326	0,038795097	0,047404676
Mean squared error	0,003698898	0,00210568	0,004913924	0,003925098
Root mean squared error	0,060818564	0,045887687	0,070099385	0,062650606
Mean signed difference	-0,025403426	-0,020660626	-0,034233852	-0,014404548
Metrics	*Tree Ensemble Regression*	*Linear Rergession*	*Polynomial Regression*	*Simple Regression Tree*
R^2	0,954591928	0,659888755	0,186231621	0,916620913
Mean absolute error	0,045541846	0,113392662	0,155563993	0,054591607
Mean squared error	0,00354117	0,02652374	0,06346212	0,006502358
Root mean squared error	0,059507734	0,162861106	0,251916891	0,080637202
Mean signed difference	-0,014367007	-0,045885553	-0,12256834	-0,013013302

Algorithm performance results with augmented data					
Algorithms	**R^2**	**Mean absolute error**	**Mean squared error**	**Root mean squared error**	**Total**
PNN	1	1	1	1	4
Tree Ensemble Regression	2	4	2	2	10
ANN	3	2	3	3	11
Random Forest Regression	4	5	4	4	17
Gradient Boosted Tree Regression	5	3	5	5	18
Simple Regression Treee	6	6	6	6	24
Linear Rergession	7	7	7	7	28
Polynomial Regression	8	8	8	8	32

Prediction made with augmented data with the best predictor PNN-Probabilistic Neural Network.				
Countries	**Prediction 1**	**Augmented Data**	**Absolute Variation**	**Percentage Variation**
Bulgaria	14,70	15,75	1,06	7,20
Estonia	17,34	16,39	-0,95	-5,51
Greece	14,13	15,55	1,43	10,12
Hungary	15,17	15,76	0,59	3,90
Ireland	16,87	16,26	-0,61	-3,62
Portugal	15,01	15,80	0,78	5,23
Slovenia	15,44	15,91	0,47	3,06
Spain	15,01	15,76	0,75	5,02
Mean	15,46	15,90	0,44	3,18

Differences between statistical errors of prediction with historical data versus prediction with augmented data				
Metrics	*Prediction 1*	*Augmented Data*	*Augmented Data-Prediction 1*	
	ANN	*PNN*	*Absolute Variation*	*Percentage Variation*
R^2	0,98	0,97	-0,01	-1,02
Mean absolute error	0,03	0,04	0,01	25,37
Mean squared error	0,00	0,00	0,00	51,79
Root mean squared error	0,04	0,05	0,01	23,20
Mean signed difference	0,01			
Mean absolute percentage error	0,09			
Mean	0,03	0,03	-0,01	-15,85


Matrice di correlazione
A6 1.0 0.4 0.3 -0.0 0.7 0.7 0.8 0.3 0.4 0.2 0.1 0.7 0.8
A1 0.4 1.0 0.6 -0.2 0.4 0.5 0.6 0.7 0.4 0.2 0.2 0.5 0.4
A2 0.3 0.6 1.0 0.2 0.5 0.4 0.6 0.7 0.3 0.3 0.3 0.2 0.3
A4 -0.0 -0.2 0.2 1.0 -0.1 -0.2 -0.0 0.1 0.1 0.1 0.0 -0.1 -0.1
A5 0.7 0.4 0.5 -0.1 1.0 0.7 0.9 0.3 0.4 0.3 0.3 0.6 0.7
A17 0.7 0.5 0.4 -0.2 0.7 1.0 0.9 0.4 0.5 0.4 0.3 0.7 0.9
A19 0.8 0.6 0.6 -0.0 0.9 0.9 1.0 0.5 0.5 0.4 0.4 0.7 0.9
A38 0.3 0.7 0.7 0.1 0.3 0.4 0.5 1.0 0.4 0.3 0.2 0.2 0.3
A40 0.4 0.4 0.3 0.1 0.4 0.5 0.5 0.4 1.0 0.2 0.1 0.3 0.4
A41 0.2 0.2 0.3 0.1 0.3 0.4 0.4 0.3 0.2 1.0 0.4 0.2 0.3
A42 0.1 0.2 0.3 0.0 0.3 0.3 0.4 0.2 0.1 0.4 1.0 0.2 0.3
A44 0.7 0.5 0.2 -0.1 0.6 0.7 0.7 0.2 0.3 0.2 0.2 1.0 0.8
A47 0.8 0.4 0.3 -0.1 0.7 0.9 0.9 0.3 0.4 0.3 0.3 0.8 1.0
A6 A1 A2 A4 A5 A17 A19 A38 A40 A41 A42 A44 A47
1
0,5
0
-0,5
-1