DOI 10.2478/sbe-2023-0028
SBE no. 18(2) 2023

THE APPLICATION OF SOCIAL NETWORK ANALYSIS TO ECONOMIC AND REGIONAL DEVELOPMENT: TERTIARY EDUCATIONAL ATTAINMENT BY SEX AND NUTS 2 REGIONS

FILENTA Pagona
Department of Economic Sciences, International Hellenic University, Serres, Greece

KYDROS Dimitrios
Department of Economic Sciences, International Hellenic University, Serres, Greece

Abstract:
Social Network Analysis (SNA) is a process of investigating the structure of a network. With the contribution of graph theory, it studies the relationships (links) that are established between different units, such as individuals, groups, or organizations (nodes), which are taken as interdependent rather than as autonomous. A social network can be compared to a web, which records the relationships or interactions of people. These networks include the social or professional interactions between units through which information and services are exchanged or contacts are recorded on a systematic basis. In this paper we attempt to introduce SNA in Economic and Regional Development, i.e. to examine the correlation of a regional education statistic indicator between certain regions. Specifically, we study the share of citizens who have completed higher education for the age group 25-64 years by gender and NUTS 2 regions. We present the visualisation of several networks, which are generated based on the degree of correlation of the data, and compute measures of centrality to compare the level of higher education by gender. The aim is to determine whether the application of Social Network Analysis could be an alternative research method in the study of Economic and Regional Development.

Key words: *Social Network Analysis (SNA), Economic and Regional Development, European Regions NUTS 2, Tertiary educational attainment, gender equality*

1. Introduction

The term "social network" depicts a social set consisting of some individuals, groups, or organizations (nodes) and some interactions (links) that are created between them. The study of these interactions is called Social Network Analysis (SNA) and attempts to describe the structure of the link-relationships or to identify the flow of information between individuals and organizations or groups. In SNA, social structures are explored using graphs. Examples of social structures commonly shown through SNA include social media networks, information circulation, friendship and acquaintance networks, business networks, knowledge networks, work relationships, disease transmission, and sexual relationships (Wasserman & Faust, 1994). The theory of SNA is not limited to the scientific field of sociology but rather finds application in different fields such as computer networks, biology, and economics. Also, its methods have been applied to cultural and literary works (Kydros & Anastasiadis, 2017).

The theory of Economic and Regional Development is about the geography of prosperity and its progress. While regions are in the process of development and adaptation, uneven economic, social, and environmental benefits are created as it is an unequal process. The mobility of people and capital, as well as inter-regional trade, foster the negative relationship between growth and regional imbalances (Petrakos, et al., 2005). Therefore, despite the existence of regional development programs for almost fifty years, the spatial distribution of wealth and economic activity leads to the formation of regional imbalances. Convergence-divergence goes beyond the scope of the academic issue when studying the policy that will lead to the economic development of a region and the decrease of inter-regional imbalances (Alexiadis, 2020). According to the concept of economic and regional development, convergence is the process by which a less developed region grows more quickly than a developed region (Koudoumakis, et al., 2019).

Economic and regional development in the European Union is a subject of study in international literature, which has been studied in depth by many scholars over the past decades (Petrakos & Saratsis, 2000). Many research use various quantitative and statistical analysis techniques to examine the spatial and structural variations across different locations.

This paper attempts to apply SNA to economic and regional analysis. Specifically, it uses the tertiary education attainment in the NUTS2 regions of the European Union from 2010 to 2021 by gender to construct correlation tables, which will be processed to obtain adjacency matrices. Then, through the analysis and visualization software package NodeXL Pro (Smith, et al., 2010) the construction of networks (graphs) is attempted. Using appropriate SNA techniques, i.e. appropriate metrics and centrality measures, we will conclude the interaction between the NUTS2 regions.

Consequently, we try to draw conclusions about the trend of interaction of tertiary graduation rates in the NUTS 2 regions of the EU by gender. In other words: How are the networks structured with Female and Male data? Do they exhibit the same homogeneity? Can regions be grouped (communities formed) based on the correlation between Female and Male education rates? What would such a grouping mean in terms of European homogeneity in education?

2. Social Network Analysis (SNA) and Economic and Regional Development

In this section we will focus on SNA, as well as Economic and Regional Analysis. We will discuss definitions, concepts and indicators, which we will use in the next section.

2.1 Social Network Analysis (SNA)

The term "social network" refers to a social structure consisting of several nodes, i.e. individuals, groups, or organizations and several links, i.e. social or professional relationships/contacts that develop between them. Examples include employees in a company, groups of individuals, such as football teams participating in a league, and even communities, regions or countries, such as a network of international trade. In a set of nodes, we can look at the different relationships that connect the nodes, taking a different network each time. Students in a class create one network in terms of "sports activities outside school" and a different network in terms of "friendly relationships between them". Actors (nodes) and their actions (links) are treated as interdependent rather than independent-autonomous units (Wasserman & Faust, 1994).

Networks are often depicted through matrices or graphs in which nodes are represented as points and links are represented as lines. A graph corresponds to a network in mathematics. In graph theory, pairwise relationships between objects are modeled and studied using the geometry of position (graphs) (Otto, et al., 2022). A graph G consists of a set of nodes (vertices) N= $\{n_1, n_2, n_3,\ldots,n_N\}$, with n=|N| the total number of nodes, and a set of links A= $\{l_1, l_2, l_3,\ldots, l_L\}$ represented as lines between the node. When an edge e ∈ E joins two nodes n_i, n_j ∈ N, then e = (n_i, n_j) and the nodes are called adjacent. A node's neighborhood, N_i (g), is the set of nodes that are related to it in such a way that N_i (g) = $\{n_1 \mid n_2 \subseteq g\}$ (Jackson, 2005).

While studying graphs there are some fundamental descriptive measures that provide important information about the typology of networks. Graph density (d) is calculated as the ratio of the total number of existing edges to the total number of edges that are likely to exist in a network:

$$d = \frac{2e}{N(N-1)}$$

When the density of a graph is close to 1, then this graph is dense, otherwise it is characterized as sparse.

Other indicators that are important and examine the overall location, role, and activity of nodes in a network are centrality measures, which lead to conclusions about nodes that occupy important or central positions in the network. Betweenness centrality mentions to nodes that occupy a position "between" other nodes. A node with high betweenness centrality acts as an intermediary between other nodes in terms of their shortest path and is thus considered essential in the flow of information between pairs of nodes. Consequently, the removal of such a node would mean a break in the communication of the network (Newman, 2002). The clustering coefficient measures the local density. That is, it is an indicator that measures the degree to which the nodes of a graph tend to cluster together, i.e. whether the nodes of a network create triangles. This index detects whether certain nodes in a graph tend to create strongly connected regions. It has been shown that in many social networks, there are nodes that form densely connected communities, with high local densities. The clustering coefficient can be calculated for the overall graph to give an insight into the overall degree of network clustering and at the local level as an indication of the integration of individual nodes (Watts & Strogatz, 1998).

Node clustering refers to segmenting a graph in order to identify high link density between nodes within a group (cluster) and low link density between nodes in different groups. Such highly connected clusters are commonly referred to as communities and the process of identifying the optimal communities in a graph is known as community detection (McNulty, 2022). When analyzing different networks, it can be important to discover communities within them (Girvan & Newman, 2002).

2.2 Economic and Regional Development

Economic and regional development refers to the development process aimed at improving the economic, political and social well-being of a region. The spatial distribution of income, economic opportunities and activities, which is not equal, has been the subject of intense study in recent decades. These studies focus on spatial and structural regional imbalances, using various statistical analyses. It is important to mention that Economic and Regional Development is not about a static figure, but about the complex potential dynamics of regions.

In the EU in recent years, regional inequality has increased rapidly. This phenomenon is not only observed in the European area, but in many countries, both developed and developing (Ganong & Shoag, 2015). In many rural areas and medium-small metropolitan areas, which used to be prosperous, there is a combination of job loss or labor force decline and a decline in per capita income relative to the national average. In contrast, many large metropolitan areas have experienced a particularly high level of income and employment, having created high-paying jobs (Iammarino, et al., 2018).

The analysis of the cohesion of EU regions is particularly important as the Union is not a homogeneous economic area. The regions that make it up have different levels of growth and development potential (Monford, et al., 2013). It is estimated that growing inequalities harm the EU's economic growth and threaten social peace and political stability (Widuto, 2019). The formation of cohesion policy, which results in the elimination of disparities and regional development, is primarily aided by the collecting, analysis, and assessment of quantitative data (Annoni, et al., 2019; Cuadrado-Roura, 2001). The vast majority of studies have utilized various quantitative and statistical analysis techniques to concentrate on the geographical and structural inequalities of regions.

There are different approaches and methodologies when analyzing Economic and Regional Development. Examples could include exploratory spatial data analysis (ESDA), which describes spatial distributions (Annoni, et al., 2019; Dall;erba, 2005; Ertur & Koch, 2006), β-convergence and sigma-convergence, methods that analyze interregional data over time (Chocholata & Furkova, 2017). In addition, they include measures that study the spatial autocorrelation of regions, i.e. whether the growth of one region is directly related to the corresponding growth of a neighboring region. Statistical measures of spatial autocorrelation are Moran's I and Geary's C indices (Lopez-Bazo, et al., 1999). Other indices used when studying Economic and Regional Development are the Gini, Atkinson, and Theil indices, which calculate economic difference (Monfort, 2008).

Finally, several variables were used in the Economic and Regional Analysis. The most common index used is Gross Domestic Product (per capita). However, there is much research that uses alternatively other indexes like employment/unemployment rate, education level, population indicators, etc. All these indexes are often used in statistical analysis for the welfare of regions (Nijkamp & Abreu, 2009).

In this paper, the measure used is tertiary education attainment. Every country's growth, economy, and competitiveness are driven by the higher education industry. According to Grant (2017), education is a key driver of development, which contributes substantially to economic growth. Judson (1998) argues that there is a positive and substantial correlation between GDP growth and human capital accumulation when economic resources are allocated more efficiently to the education sector (Botoroga, et al., 2022). Governments support higher education in a variety of ways and systems due to these factors. The state's competitive edge in connection to the environment is ensured by investment in education, particularly higher education. In the EU, 40.7% of people aged 30-34 had a tertiary degree in 2018. In 2017, the EU-28 allocated an average of 4.8% of GDP to education, according to EUROSTAT. For each country to build a competitive advantage, a viable, developed, and supported higher education industry is a requirement. This competitive advantage is built on strengths in innovation, research, highly skilled labor force, higher percentage of white-collar professionals, etc. Different nations and cultures handle this problem in different ways (Karanović, et al., 2023). According to Grant (2017), education is a key driver of development, which contributes substantially to economic growth. Judson (1998) argues that there is a positive and substantial correlation between GDP growth and human capital accumulation when economic resources are allocated more efficiently to the education sector.

When it comes to gender equality in various economic criteria, there is still a long way to go, even if many countries have succeeded in reducing it. The number of female leaders (and women who are on their route to gaining the abilities to become leaders) is rising along with the growth of professions with increasingly diverse requisite skill sets. The current rate of development is thought to be slow, and more work still needs to be done to address the general issue of gender inequality and women's empowerment (Gharehgozli & Atal, 2021).

The indicator studied in this paper is defined as the percentage of the population aged 25-64 years old who have successfully completed tertiary education (e.g. university, higher technical institution, etc.). Educational attainment refers to ISCED (International Standard Classification of Education) 2011 level 5-8 for data from 2014 onwards and ISCED 1997 level 5-6 for data up to 2013. The indicator is based on the EU Labor Force Survey (Eurostat, 2023).

2.3 Methodology

Data collection is the systematic process of collecting and measuring information about the variables of interest to find answers to certain research questions, as well as conclusions. Usually, the amount of data already collected by others is large. In this paper, this is the case. In particular, we chose to use data, which have been published in the Eurostat database (https://ec.europa.eu/eurostat/data/database). The sample consists of data of regional statistics by NUTS classification. Specifically, we collected the tertiary education attainment for the age group 25-64 years by gender and NUTS 2 regions for the period of 2010-2021.

The NUTS classification can be modified, i.e. to accommodate changes in the territorial structure of the Member States. The time allowed for any modifications is at least three years. However, an amendment may be adopted in less than three years in the case of a substantial reorganization of a Member State. In this case, the Commission shall amend the classification at the end of the stability period following the rules of the NUTS classification. The fact of modifications may create some gaps in the data, i.e. some values in the time series, which are currently not available (Eurostat, methodologies and working papers, 2011). In this paper we use data from the NUTS 2021 classification which lists 242 regions at NUTS 2.

To fill in the data of the tertiary education attainment rate, which were not available, we made some computations; these are described in the Appendix.

The overall process resulted in the final sample of the study, which consists of 239 NUTS 2 regions. Therefore the data, i.e. the tertiary education attainment of the 239 regions for the period 2010-2021 was compiled in a spreadsheet. Next, correlation tables were constructed with the tertiary education attainment data, based on the following correlation coefficient:

$p_{i,j} = \frac{cov_{i,j}}{\sigma_i \, \sigma_j}$, where

cov_{ij} = the covariance of regions i and j, where $cov_{ij} = \frac{\lfloor P_i - E(P_i) \rfloor \lfloor P_j - E(P_j \rfloor}{n-1}$,

E(P) = the average of each region, n = the count of the regions,

σ_i = the standard deviation of region i and σ_j = the standard deviation of region j

The correlation coefficient for each pair of regions i, j was calculated and a cross-correlation, rectangular, one diagonal and symmetric matrix was created for each data set, with dimensions 239 X 239, which includes values between -1 to 1. When the correlation coefficient tends to 1, then the two regions move in the same direction, i.e. when the tertiary education attainment of one region increases, the share of the second region also increases; when the correlation coefficient tends to -1, the two regions move in the opposite direction, i.e. when the tertiary education attainment of one region increases, the share of the second region decreases and vice versa. If the coefficient is zero, no prediction can be made because this means that the two regions are moving randomly.

The correlation coefficient matrices were then used to construct a correlation network, where each node represents a region. A certain threshold (θ) was also set, which can take values from -1 to 1. In the case where a correlation coefficient of the matrix is greater than θ, a link is created that connects the two regions. The result of this process is an NXN adjacency matrix where:

$$A_{i,j} = \begin{cases} 1, & if\ n_i\ is\ connected\ to\ n_j \\ 0, & otherwise \end{cases}$$

The process was repeated, setting different values to θ, resulting in different adjacency matrices, which have the same number of nodes (regions), however, they have a different number of links (Huang, et al., 2009). The next step was to construct graphs, i.e. converting the adjacency matrices into graphs, so that a visualization of the nodes and their links is possible. This was achieved through the NodeXL Pro software package (Smith, et al., 2010). Finally, using the appropriate tools in NodeXL Pro, metrics were then calculated to conclude the community structure and the interaction of the regions.

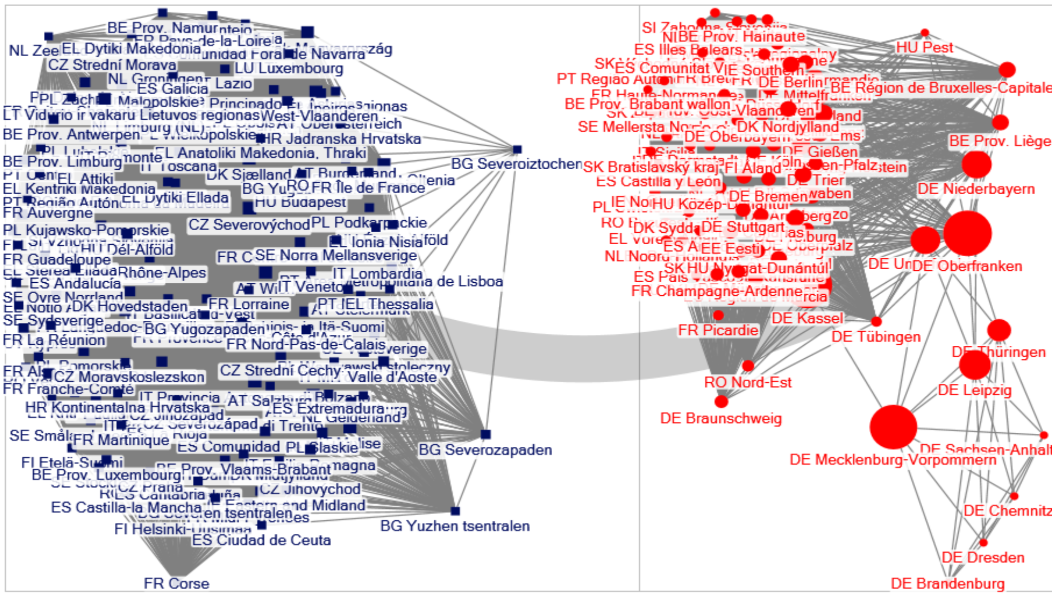
3. Data Analysis

In the next section, we will attempt to study the networks constructed for different values of θ (threshold), as discussed previously. First, the whole procedure was applied for the Female and Male data. To study the correlation of the tertiary education attainment rate by gender between NUTS 2 regions, we set threshold values from θ > 0,6 to θ > 0,9. That is, we tried to find out whether the rate of citizens who have completed tertiary education between regions moves in the same or opposite direction at different degrees of correlation (for different values of θ). In addition, we calculated centralities for each constructed network and the graph density. The size of the nodes was based on betweenness centrality, i.e. the higher the centrality, the larger the size of each node (Brandes, 2001). Follow the networks constructed by gender.

3.1. Network of Female Data

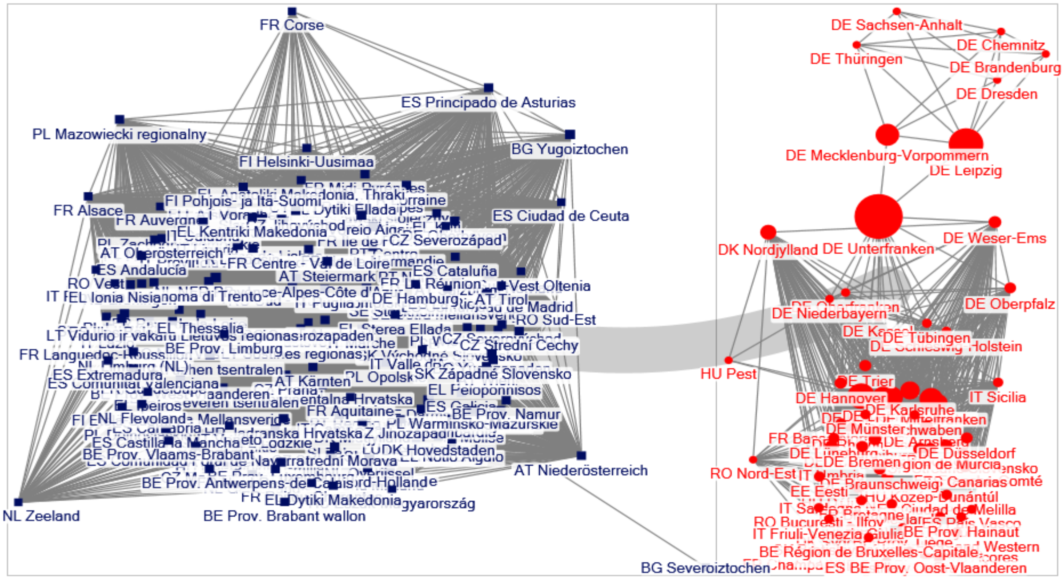
The first graph created for the Female data is the graph with threshold θ>0,6. This network (Figure 1) consists of two communities of regions, and only one isolated node. In the first community 144 nodes exist, which are connected to each other with 10.034 links, while in the second one 94 nodes exist with 3.533 links. The network is quite dense (graph density = 0,85), so no important conclusions can be drawn. For this reason, the next network was created with an increased threshold of θ > 0,70 (Figure 2).

Figure 1: Female Network, Betweenness Centrality, Correlation θ>0,6



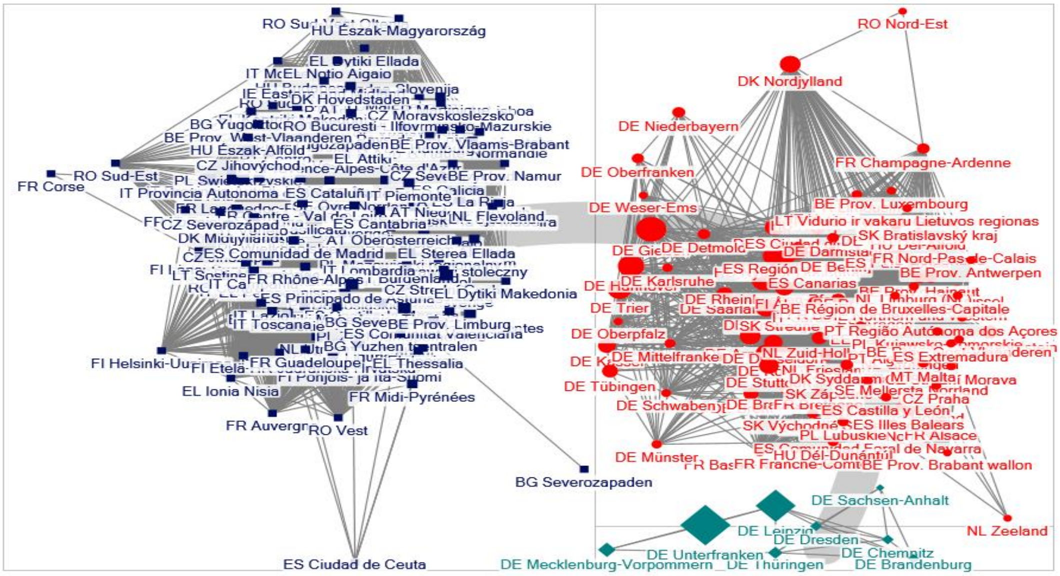
In the following network for θ > 0,70 (Figure 2) the resulting communities are also two. The density is reduced to 0,79, as increased θ means that the number of edges between the nodes is smaller. Also, the number of the isolated node is still one, i.e. it is the same region (FR Guyane). The first community consists of 161 nodes, connected with 12.224 links, while the second community consists of 77 nodes with 2.053 links.

Figure 2: Female Network, Betweenness Centrality, Correlation θ>0,7



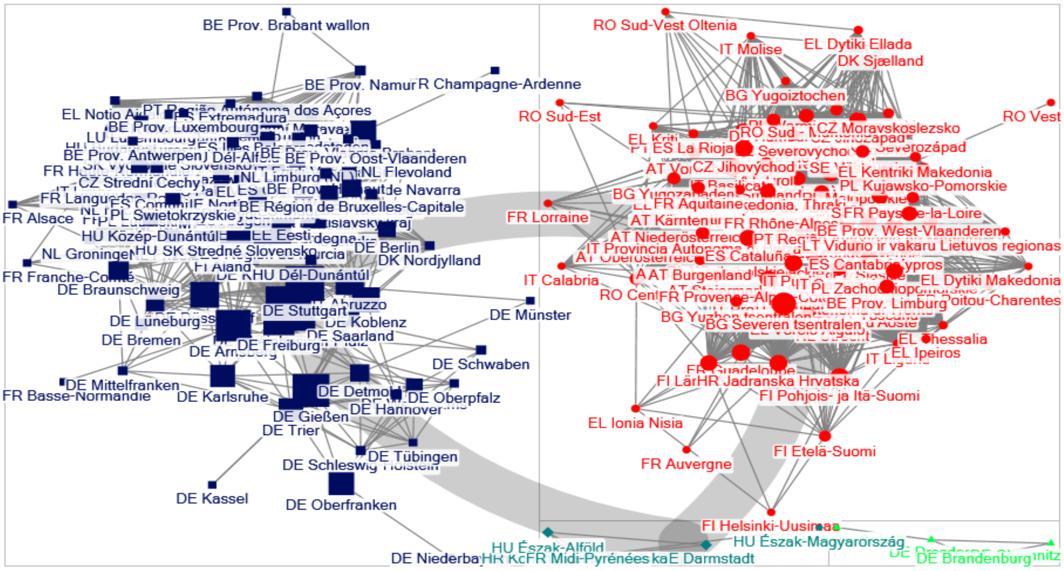
Increasing the threshold θ to 0,80 (Figure 3) results in a new network, consisting of three communities: the first one consists of 133 nodes, connected by 7.909 links, the second one consists of 95 nodes with 3.215 links, and the third one includes 8 nodes and 18 links,. The graph's density was reduced to 0,68, while the isolated nodes were increased to 3.

Figure 3: Female Network, Betweenness Centrality, Correlation θ>0,8



In the last network created for the Female data the threshold of θ > 0,9 was set (Figure 4), i.e. the highest correlation. The number of isolated nodes in this network was increased to 12. These nodes belong to 2 regions of Bulgaria, 4 regions of Germany, 1 region of Spain, 2 regions of France, 1 region of Hungary, 1 region of Netherlands, and 1 region of Romania. This network is by far the sparsest, with a density of 0,41 and an increase in the number of communities, equal to 4.

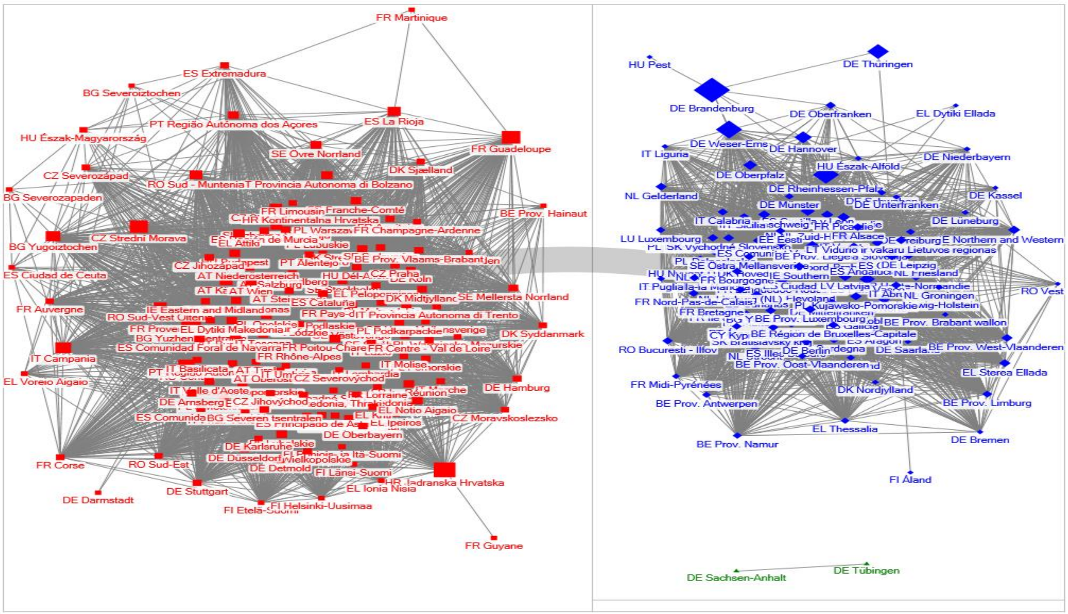
Figure 4: Female Network, Betweenness Centrality, Correlation θ>0,9



3.2. Network of Male Data

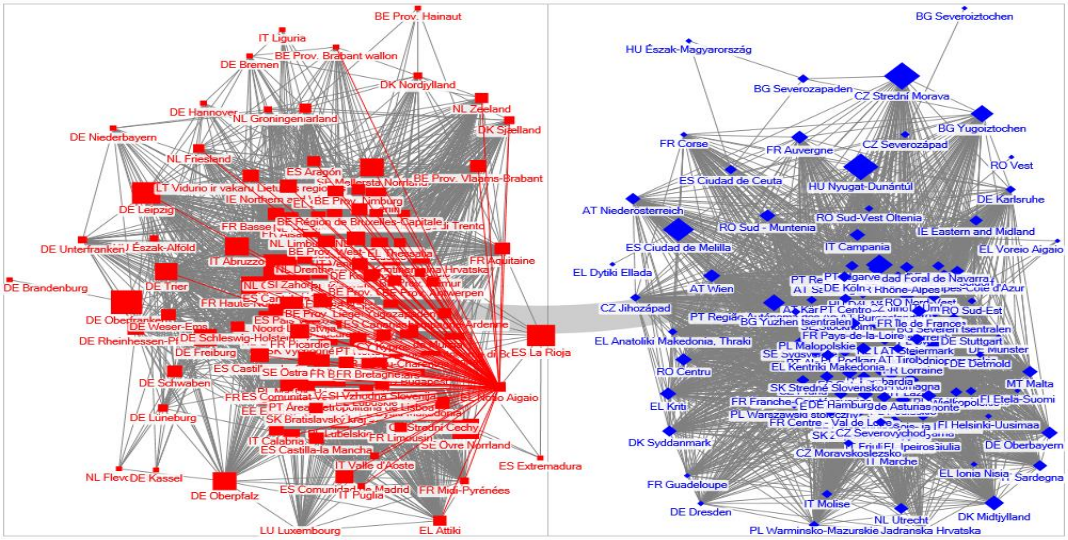
The first graph created for the Male data is the graph when the threshold θ>0,6. In this network (Figure 5) three communities of regions can be distinguished. The first community consists of 130 nodes, which are connected to each other with 6.801 links, while the second and third group consist of 102 nodes with 3.836 links and 2 nodes and 1 link respectively. An important observation is the fact that in the network exist 5 isolated nodes. The network is quite dense (graph density = 0,66), so no substantial conclusions can be drawn. For this reason, the next network was created with an increased threshold of θ > 0,70 (Figure 6).

Figure 5: Male Network, Betweenness Centrality, Correlation θ>0,6



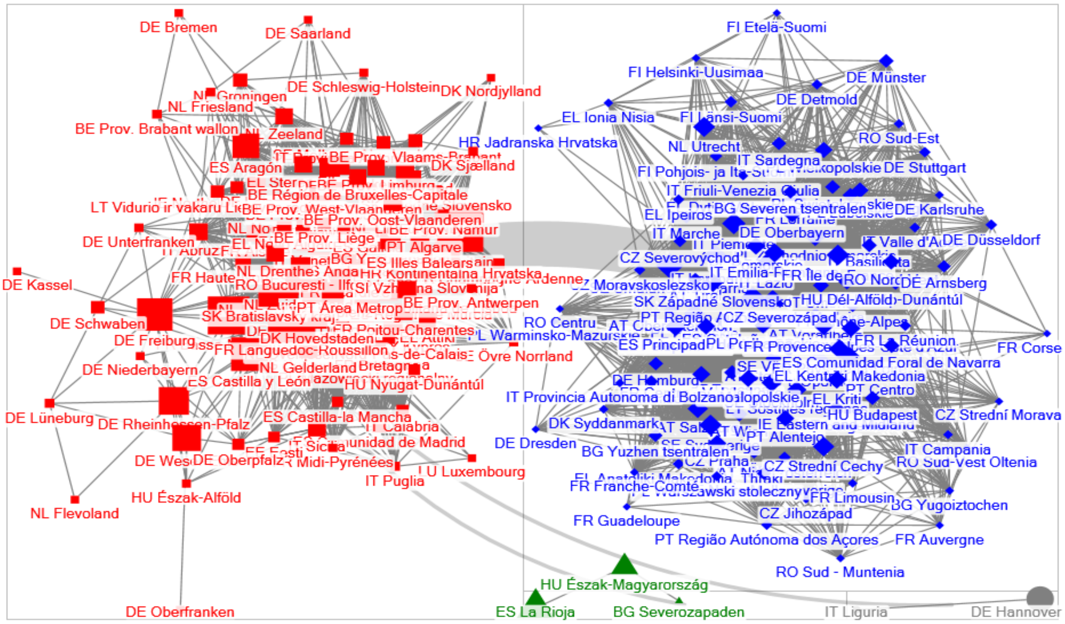
In the following network for θ > 0,70 (Figure 6) the resulting communities are two. The density is reduced to 0,53 as increased θ means that the number of edges between the nodes is smaller. Also the number of the isolated nodes increased to 13. The first community consists of 116 nodes, which are connected with 4.302 links, while the second consists of 110 nodes with 4.372 links.

Figure 6: Male Network, Betweenness Centrality, Correlation θ>0,7



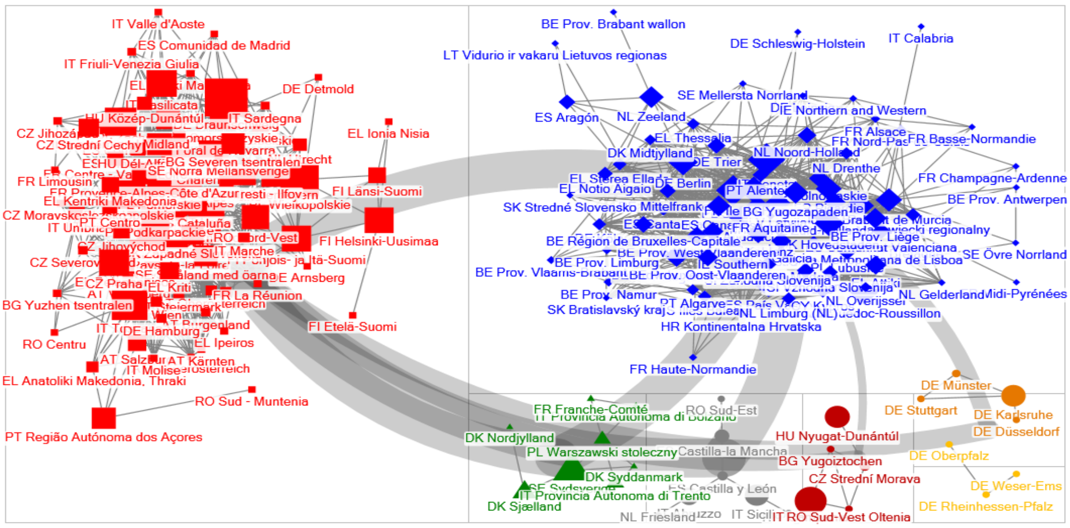
Increasing the threshold θ to 0,80 (Figure 7) results in a new network, consisting of four communities: the first one consists of 107 nodes, connected by 2.929 links, the second one consists of 105 nodes with 3.283 links, the third one includes 3 nodes and 2 links, and the last consists of only 2 nodes and 1 link between them. The density of the graph was reduced to 0,35, while the isolated nodes were increased to 22.

Figure 7: Male Network, Betweenness Centrality, Correlation θ>0,8



In the last network created for the Male data we set θ > 0,9 (Figure 8). The number of isolated nodes in this network increased to 48. This network is by far the sparsest, with a density of 0,19, and an increase in the number of communities, equal to 7.

Figure 8: Male Network, Betweenness Centrality, Correlation θ>0,9



3.3 Discussion

Having completed the analysis by network, it was considered useful to make some observations on the structure of the networks based on the resulting communities. A Synoptic topology of the created networks is shown in Table 1 and Table 2 below, from which some conclusions can be drawn. The higher the threshold, the lower the correlation of tertiary education attainment between regions. As a result, the number of isolated nodes increases, the number of edges connecting nodes decreases, i.e., the network is less dense, and also the number of communities in each network increases.

Table 1: Typology of Female Data

Correlation	> 0,6	> 0,7	> 0,8	> 0,85	> 0,9
Nodes	238	238	236	233	227
Edges	24.268	22.486	19.403	16.529	11.774
Isolated Nodes	1	1	3	6	12
Graph Density	0,85	0,79	0,7	0,46	0,21
Communities	2	2	3	3	4

Table 2: Typology of Male data

Correlation	> 0,6	> 0,7	> 0,8	> 0,85	> 0,9
Nodes	234	226	217	206	191
Edges	18.799	15.319	10.139	6.851	3.458
Isolated Nodes	5	13	22	33	48
Graph Density	0,66	0,53	0,35	0,32	0,19
Communities	3	2	4	3	7

Analysis by gender

Regarding the differences found between the Female and Male networks, it can be observed that the Female network is overall more homogeneous than the Male network. In networks with the same threshold θ the isolated nodes of Male are far more than the corresponding network of Female, e.g. when θ> 0,6 the isolated nodes in the Male data equals 5, while in Female it equals 1. Also as the threshold θ increases to 0,7, 0,8 and 0,9 the isolated nodes increase for Male data to 13, 22 and 48, while in Female data initially it is fixed at 1, and then increases to 6 and 12.

The number of communities also indicates the homogeneity of a network. The more communities are formed in a graph, the higher the heterogeneity. Comparing the number of communities in the networks, we find that the Male data have a larger number of communities for each threshold. The greatest divergence in the number of communities is found for θ>0,9, where the Male data created 7 communities, while the Female data created 4 communities.

Finally, the networks' density is a measure that allows us to perform some results on homogeneity. Looking at the network density for each threshold θ, we conclude that Female networks are denser compared to Male networks. This is essentially reflected in the number of links between nodes. Again, there are more links between nodes in each graph in the female data than in the corresponding graphs created from the male data. For example, when the threshold θ > 0,85, the edges between nodes in the overall network for the male data amount to 6.851, while the links in the corresponding network for the female data amount to 16.529. Moreover, the density of each network reflects the same result. Therefore, it could be concluded that the Female networks are, overall, more homogeneous compared to the Male networks.

The fact that Female networks are more homogeneous than Male means that in many regions, the rates for Female data tend to move in the same direction. In other words, it is observed that in several regions, the rates of women completing tertiary education are increasing or decreasing at a similar rate. For each level of correlation, there are fewer communities of regions in the Female networks, which means that the data are moving in the same direction, i.e., increasing or decreasing in parallel.

In contrast, the same cannot be said for the Male data. Indeed, when the correlation coefficient is high, the networks generated show that the proportions of men completing tertiary education over time are not homogeneous, i.e. they do not move in parallel but create different communities, i.e. groups of regions. In conclusion, many regions do not show the same trend in male tertiary education.

The above conclusion regarding the strong presence of women in higher education is confirmed by the Gender Equality Index 2020: digitalization and the future of work of the European Institute for Gender Equality. In recent years, the rates of women and men completing higher education have been steadily increasing in EU Member States, with the gap between them gradually reversing in favor of women. In 2010, only 20 % of women and 21 % of men had graduated from university, while in 2018, the percentage increased more for women than for men. Specifically, 26% of women and 25% of men had completed higher education in the age group of 15 years and above. The largest gender differences in tertiary education attainment in favor of women were found in Estonia, Latvia, and Sweden. Significant differences were also found in nine other Member States (Bulgaria, Denmark, Ireland, Cyprus, Lithuania, Poland, Portugal, Finland and Slovenia). Finally, in four countries, men were more likely to have a tertiary degree than women: Germany, Luxembourg, the Netherlands, and Austria (European Institut for Gender Equality, 2023).

Despite the fact that there are still certain gaps between the genders, the EU has made great progress toward gender equality in recent years. This is due to laws requiring equal treatment, the inclusion of the gender perspective in all other policies, and specific initiatives for the advancement of women. The increase in the proportion of women in the labor force and their success in obtaining better education and training are encouraging trends. However, there are still gender inequalities in the workforce, with women still outnumbering men in the lowest-paying industries and underrepresenting them in positions of power (European Commission, 2023).

4. Conclusions

In this paper, we have tried to introduce the SNA in Economic and Regional Development. Many academics who concentrate on the spatial and structural disparities between areas have been interested in the occurrence of imbalances in the spatial distribution of income, economic opportunities, and activities using various quantitative and statistical analysis. We have made an effort to provide a response to the topic of whether SNA may be used for regional and economic analysis. For this purpose, we used the percentage of citizens who have completed higher education. In particular, the analysis was based on the tertiary education attainment by gender and NUTS 2 regions for the period 2010–2021. SNA techniques were used to certify certain results. We calculate metrics, such as degree centrality, betweenness centrality, and the number of created communities to conclude about the interaction between the regions and compare the Female and Male data.

The network results show that as θ increases, the more communities (subgraphs) a network contains. It should be mentioned that regions belonging to the same community are highly correlated with each other, i.e., the tertiary education attainment for the period under consideration is moving in the same direction. We found that SNA locates the interaction between regions and identifies the regional communities. The paper shows that the networks created from the female data are more homogeneous than the male data. This highlights that the presence of women in higher education is strong for the majority of European regions. In other words, for each level of correlation there are fewer communities of regions in the female networks, which means that the data are moving in the same direction, i.e. increasing or decreasing in parallel.

In contrast, the networks created from the male data show less homogeneity, i.e. creating more communities of regions. This practically means that the regions are not moving in parallel with each other. Also, for different levels of correlation of the tertiary education rate, there are regions that stand out from the whole, as they show a different trend in the data, without moving like the rest of the European regions.

Finally, one comment that could be added is about the homogeneity of the data within Member States. In most Member States it is observed that there is no homogeneity, as the regions for the different levels of correlation of the data belong to different communities. The exceptions are Austria, the Czech Republic and the Netherlands, whose regions are not split but belong to the same community, i.e. the tertiary education rate has the same trend for their regions.

The introduction of SNA in Economic and Regional development could be extended to other variables such as GDP per capita, population density, the income of households, etc. Taking into account that the EU policy is to smooth out imbalances and differences and ensure coherence between regions, one realizes that the ideal network would consist of a single community, where no isolated nodes exist, i.e. a homogeneous network. This would practically mean that the macroeconomic indicators would move in the same direction for all regions, with no regions that stand out. In a next step, an attempt will be made to analyze the policy and strategy options that could be applied to specific regions in order to smooth out the imbalances.

The present study has been presented at the 15th International Conference "Economies of the Balkan and Eastern European Countries", EBEEC 2023, Chios, Greece (http://ebeec.ihu.gr/)

5. References

Alexiadis, S., 2020. Regional Convergence: Theory and Empirics. *Regional Science Inquiry*, XII(1), pp. 245-252.

Annoni, P., de Dominicis, L. & Khabirpour, N., 2019. Location matters: A spatial econometric analysis of regional resilience in the European Union. *Growth and Change*, pp. 1-32.

Botoroga, C.-A.et al., 2022. Assesing the Nexus Between Education, Economic Growth, and Innovation: An Empirical Analysis. *Studies in Business and Economics*, 17(3), pp. 18-34.

Brandes, U., 2001. A faster algorithm for betweeness centrality. *The journal of Mathematical Sociology*, 25(2), pp. 163-177.

Chocholata, M. & Furkova, A., 2017. Does the location and the institutional backround matter in convergence modelling of the EU regions?. *Central European Journal of Opereations Research*, 25(3), pp. 679-697.

Cuadrado-Roura, J. R., 2001. Regional convergence in the European Union: From hypothesis to the actual trends. *The Annals of Regional Science*, December, Issue 35, pp. 333-356.

Dall;erba, S., 2005. Distribution of regional income and regional funds in Europe 1989-1999: An exploratory spatial data analysis. *The Annals of Regional Science*, pp. 121-148.

Ertur, C. & Koch, W., 2006. Regional disparities in the European Union and the enlargement process: an exploratory spatial data analysis, 1995–2000. *The Annals of Regional Science*, 40(4), pp. 723-745.

European Commission, 2023. *Gender Equality Strategy.* Available onlline at: https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/gender-equality/gender-equality-strategy_en

European Institut for Gender Equality, 2023. *Gender Equality Index 2020: Digitalisation and the future of work,* s.l.: EIGE's publications.

Eurostat, methodologies and working papers, 2011. *Statistical books, Regions in the European Union, KS-RA-11-011-EN,* Luxembourg: Publications Office of the European Union.

Eurostat, 2023. *International Standart Classification of Education.* Available online at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=International_Standard_Classification_of_Education_(ISCED)

Ganong, P. & Shoag, D., 2015. *Why has regional income convergence in the US declined?,* Cambridge: Harvard Kennedy School Working Paper.

Gharehgozli, O. & Atal, V., 2021. Gender Wage Gap: Within and Across Borders. *Studies in Business and Econpmics*, 16(2), pp. 74-83.

Girvan, M. & Newman, M., 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 11 June, 99(12), pp. 7821-7826.

Huang, W.-Q., Zhuang, X.-T. & Ya, S., 2009. A network analysis of the Chinese stock market. *Physic A: Statistical Mechanics and its Applications*, 388(14), pp. 2956-2964.

Iammarino, S., Rodriguez-Pose, A. & Storper, M., 2018. Regional inequality in Europe: evidence, theory and policy implications. *Journal of Economic Geography*, pp. 1-23.

Jackson O. Matthew, 2005. The Economics of Social Networks. August.

Karanović, G., Dragičević, D. & Draženović, O., 2023. Financing Higher Education and Moral Hazard: A Systematic Literature Review. KnE Social Sciences. 8(1), pp. 233-255.

Koudoumakis, P., Botzoris, G., Protopapas, A. & Profilidis, V., 2019. The Impact of the Economic Crisis In the Process of Convergence of the Greek Regions. *Regional Science Inquiry*, XI(1), pp. 25-32.

Kydros, D. & Anastasiadis, A., 2017. Greek Political Language during the Economic Crisis- A Network Analytic Approach. *Social Networking*, 6(02), pp. 164-180.
Lopez-Bazo, E., Vaya, E. & Mora, A. J., 1999. Regional economic dynamics and convergence in the European Union. *The Annals of Regional Science*, 33(3), pp. 343-370.
McNulty, K., 2022. *Handbook of Graphs and Networks in People Analytics: With Examples in R and Python.* New York: Taylor & Francis Ltd.
Monford, M., Cuestas, J. C. & Ordonez, J., 2013. Real convergence in Europe: A cluster analysis. *Economic Modelling*, Volume 33, pp. 689-694.
Monfort, P., 2008. *Convergence of EU regions- Measures and evolution.* Available online at: https://ec.europa.eu/regional_policy/sources/docgener/work/200801_convergence.pdf
Newman, M. E. J., 2002. *The mathematics of networks,* Michigan: Ann Arbor: University of Michigan.
Nijkamp, P. & Abreu, M., 2009. Regional Development Theory. *International Encyclopedia of Human Geography*, December, pp. 202-207.
Otto, J., Chaodong, H. & Jian-Yu Fisher, K., 2022. Using Graph Theory to Optimize Career Transitions. *Studies in Business and Economics*, 17(3).
Petrakos, G., Rodriguez-Pose, A. & Rovolis, A., 2005. Growth, inmtegration, and regional disparities in the European Union. *Environment and Planning A*, Volume 37, pp. 1837-1855.
Petrakos, G. & Saratsis, Y., 2000. Regional inequalities in Greece. *Papers in Regional Science*, January, 79(1), pp. 57-74.
Smith, M. et al., 2010. *NodeXL: a free and open network overview, discovery and exploration add-in for Excel 2007/2010/2013/2016 from the Social Media Research Foundation.* Available online at: https://www.smrfoundation.org
Wasserman, S. & Faust, K., 1994. *Social Network Analysis: Methods and Applications.* Cambridge: Cambridge University Press.
Watts, D. J. & Strogatz, S. H., 1998. Collective dynamics of ‘small-world’ networks. *Nature*, June, Volume 393, pp. 440-442.
Widuto Agnieszda, 2019. *European Parliament Research Service, ERPS.* Available online at: https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637951/EPRS_BRI(2019)637951_EN.pdf

APPENDIX

Linear regression was applied to make projections for the following regions: Warszawski stoleczny (PL) Mazowiecki regionalny (PL) for 2010-2012, Guadeloupe (FR), Martinique (FR), Guyane (FR) and Reunion (FR) for 2010-2013. Furthermore, for the Male the same strategy was applied for Aland (FI) for 2010-2011 and for 2016-2017.

The two Irish regions Border, Midland and Western, and Southern and Eastern (NUTS 2013) were amended by NUTS 2016 classification to three regions: Northern and Western, Southern and Eastern, and Midland. As a result, there was missing data in the time series for 2010 and 2011. So the average of Border, Midland and Western, and Southern and Eastern were used to fill in the missing data in Northern and Western, Southern and Eastern and Midland.

The regions of Croatia Grad Zagreb and Sjeverna Hrvatska had data gaps from 2010 to 2020, while the Kontinentalna Hrvatska region (NUTS 2016) had a gap in 2021. It was decided to use the data of the first two regions for 2021 and average them to fill the gap in the latter region.

The Hungarian region of Közép-Magyarország existed in the 2013 NUTS classification, but the next classification amended the regions to include the Pest and Budapest. However the data for these regions were not available for 2010-2019. So it was decided that the data of Közép-Magyarország for the years 2010-2012 was used to fill the respective gaps of the two new regions. The same procedure was followed in the case of the Lietuva region (NUTS 2013) of Lithuania. In particular, the gaps for the years 2010-2012 of Sostines regionas and Vidurio ir vakaru Lietuvos regionas were filled with the data of the first region, i.e. Lietuva. Furthermore, the data of the region Mayotte of France was completely empty.

The procedure was applied to Female and Male Data.