PATTERNS OF MAINLY TOURISM SECTORS AT LOCAL LEVEL BY EMPLOYEE'S CHARACTERISTICS USING GIS MULTIVARIATE CLUSTERING ANALYSIS – ROMANIA CASE STUDY

Cristina LINCARU
Dr, FeRSA, Department of Labour Market, National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania ORCID ID: 0000-0001-6596-1820
cristina.lincaru@yahoo.de

Speranța PÎRCIOG
Dr, Scientific Director, National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania ORCID ID: 0000-0003-0215-038X
pirciog@incsmps.ro

Draga ATANASIU
Senior Researcher, National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania ORCID ID: 0000-0002-9695-8592
incsmps1@incsmps.ro

Cristina STROE
Senior Researcher, Department of Social Policies, National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania ORCID ID: 0000-0001-8384-6084
cristinaradu@incsmps.ro

Vasilica CIUCĂ
Dr, Dr, General Director, National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania ORCID ID: 0000-0003-4687-6377
silviaciuca@incsmps.ro

Adriana GRIGORESCU
Dr., Department of Public Management, National University of Political Studies and Public Administration, Correspondent Member of Academy of Romanian Scientists, Bucharest, Romania ORCID ID: 0000-0003-4212-6974
adrianagrigorescu11@gmail.com

Abstract

The tourism sector, before the Corona Strikes, works as a inclusive development engine for many countries' economies and labour markets. In a global world, with increasing travel opportunities, tourism offers both labours intensive and knowledge-intensive activities, across many economic sectors. Tourism is a spatially dependent sector and also a tradable one. The Methodology for tourism statistics (Eurostat 2014), Tourism Satellite Accounts (TSA 2010) and The International Recommendations for Tourism Statistics 2008 (IRTS 2008) differentiate the "mainly tourism" industries at four digits. We identify the natural cluster by number and pattern, at 3189 local spatial units (NUTS 5) by eight attribute variable employees: gender (male, female), age (youth, adult and aged) and education detained level (low, medium and high). Sectors are detailed at two digits only (H51- Air transport, I55 - Hotels and other accommodation facilities and N79-Activities of tourist agencies and tour operators; other reservation services and tourist assistance). Romanian National Institute of Statistics provides 2011 Census data. We apply the Multivariate Clustering Analysis with K Means algorithm as a Spatial Statistical Tool in Arc Gis Pro 2.3, an unsupervised machine learning an Artificial Intelligence technique, appropriate for Big Data. Clusters resulted illustrates natural hidden patterns of local labour markets pooling in the sense of Urban& Jacobian economies, but also some insight regarding the Morettian externalities sources. These results are useful for Regions Smart Specialisation Strategies development of human resources & talents to increase innovation capabilities and inclusive job creation, but also for a prompt recovery post-Covid Pandemic.

Keywords: tourism, labour force characteristics, Multivariate Clustering Analysis, local labour markets, regional specialisation, education level, age and gender analysis

JEL classification: J210, C38, R23

1. Introduction

1.1. Importance

The economic importance of Travel & Tourism sector as a driver to employment by the latest World Travels Tourism Council's (WTTC) 2020 Economic Impact report, which shows it supported in 2019 more than 330 million jobs, or 1 in 10 jobs around the world, explaining the creation 10.3% of global GDP. WTTC (2020b) "latest annual research, in conjunction with Oxford Economics, shows the Travel & Tourism sector experienced 3.5% growth in 2019, outpacing that of the global economy (2.5%) for the ninth consecutive year. Over the past five years, one in four new jobs were created by the sector, **making Travel & Tourism the best partner for governments to generate employment**". (WTTC 2020a) The increasing importance of tourism for growth and jobs is signalled by Turner (2018) as the sector that "generates prosperity across the globalised world". Future indicates an increase in growing to 2030 at 425 million jobs, or 11.1 jobs around the world, creation 1.47% of global GDP. The growth rate is 2.4%, close to the 2.5% global average. (WTTC 2020b) For Europe Union tourism and travel sector supported in 2019 more than 22.6 million jobs, or 11.2% from total EU employment, explaining the creation of 10.3% of global GDP.(WTTC 2020b)

Bănilă (2018) cites WTTC 2018 for Romania's tourism performance in 2017 with 1.4% of the country's gross domestic product (GDP) and 208,500 jobs. Romania OECD *Tourism Trends and Policies* (2020) me ntion that „according to the Tourism Satellite Account, the direct contribution of tourism to GDP in 2017 was RON 23.9 billion, 2.8% of the total GDP, and the tourism sector directly supported 373 074 jobs". The 2018 National Tourism Development Strategy assumes among its four operational objectives to support the private innovation and increase digitisation. In the Project of The National Strategy for Ecotourism Development is planned that until 2030 creating the conditions for ecotourism development in natural protected areas. Human resources development is one key area of focus.

1.2. The new global problem is Corona Pandemic

In 2020 the first semester its effect on tourism sector at global scale is dramatic. WTTC estimate in April 2020 that the Global economic impact of COVID-19 is five times more than 2008 Global Financial Crises, that put 75 million jobs at risk, accounting decreases to GDP around 2.1 trillion $. In relative terms, the COVID-19 impact is about 23% of jobs to be destroyed from total sector jobs with direct impact in the global unemployment rate increasing with 2.1 pp. (WTTC 2020c)

Petrova (2020) cites OECD 2020 that among South Eastern European countries Romania and Serbia are seriously hit by COVID-19. This negative economic impact is the result of value chain disruptions, especially in the manufacturing and tourism sectors.

Starkov (2020) points out that Corona "though unprecedented in scale and is not the first global crisis to hit hard the hospitality industry". The reaction should be similar to Zika or hurricane season manifest before at regional ora global level.

WTTC's launched on 21th of April 2020 in London, UK, four principles to ensure swift recovery for the Travel & Tourism sector and the global economy following the end of the COVID-19 outbreak. Human capital calls in the forth principle, focused on assuring the "Financial support for workers, businesses and promotion for a prompt recovery". Hartenstein and Waugh (1994) and Schlossberg (1995) considers the loss of work a critical event for the individual exposed to a significant risk of social exclusion.

1.3. Disruptive innovations in the tourism sector announced by megatrends

Especially after the Corona Pandemic is visible, the harder time for the hotels industry of 2020 very difficult to predict. (SiteMinder 2020) Hoteliers need to imagine and adopt new technology developments to meet customer demand. SiteMinder, Revinate, IDeaS and Dr Peter O'Connor found that critical management systems (revenue, customer relationship, property and channel) will add to customer-focused systems the data-focused network. Accommodation is in an intense digital transformation process. The success of this process requests **new strategies and objectives as well as new business model adoption**. Bradley (2019) lists in the top must-have IoT five web-wired tech trends in hospitality today:

Automated Guestrooms, Engagement via Mobile, Interfaces Incorporated Naturall y, Better Online Branding and Personalized Room Setups.

The innovation changes rapidly and dynamically transform this sector. Technology developments needed to meet increasingly sophisticated customers demand. Next to actual COVID-19 shock, Bloomberg Media Group (2019) announced since 2019 some disruptive world, transformed megatrends and their implications for travel & tourism. The five Bloomberg Media Group (2019) megatrends are interlinked, and comprises: "reality enhanced; life restructured; data revolutionized; power redistributed and consumption reimagined". (Bloomberg Media Group 2019)

USA Deloitte report identified that customers overwhelmingly value shared experiences. (Reichheld et al. 2020) To rethink human experience, authors title five fundamental principles: "be obsessed with all things human, proactively identify and understand human needs before they are expressed, execute with humanity, be authentic and change the world".

1.4. Talents and human resources essential innovation capability in the tourism sector

Langford, Weissenberg, and Gasdia (2019) mark that "with all the focus on emerging technology, it's often easy to forget that people will remain one of travel and hospitality's biggest challenges throughout 2019. In an industry built on service excellence, people can be a brand's most powerful competitive asset." They also emphasized that "the travel industry can't grow without **talent:** Labor gaps are not new to travel, but the magnitude of the current workforce shortage certainly is. Rapid industry growth and an evolving workforce remain key drivers". Labour shortage marks both skilled and unskilled labour. Cameron (2017) found that the 2017 US labour shortage covers with immigrants: "even tho ugh immigrants comprise only 13 per cent of the US population—they account for 31 per cent of the workforce in the hotel and lodging industry and 22 per cent in restaurants". Vasilescu, Aparaschivei, and Roman (2012) signals for Romania, that the labour market "was experiencing large and increasing shortages of labour and skills, coupled with large migration abroad".

1.5. Adaptive capacity and adaption characteristics for the tourism sector

The tourism sector is vulnerable to many factors of variability and changes. Its adaptive capacity presents variate patterns. Parsons et al. (2018) explore the case of Samoa, a Pacific island nation highly dependent on beach tourism and already vulnerable to a variety of natural hazards. Authors found that tourism operators assure the adaptive capacity to climate change impact as a network and not an individual.

Perles-Ribes et al. (2017) re -evaluate the tourism-led growth hypothesis in the case of Spain, a leading country in the tourism industry, after the 2008 Global Financial and Economic Crisis and the 2010 Arab Spring shocks. Araújo-Vila, Fraiz-Brea, and de Araújo (2020) found that in 2020 Spain has not fully recove red after the financial crises, and worries its population continuously since then. Tourism, a key industry to the Spanish economy, next to the construction industry are still the most affected sectors. Supplementary, the "COVID-19 is an even bigger problem, which expected to have catastrophic consequences for many countries' economies, including Spain" (Araújo-Vila, Fraiz-Brea, and de Araújo 2020)

Las Vegas is the case of a successful overcoming of the 2008 Global Recession, which "benefitted from stabilized or even increasing international tourism demands". Lim and Won (2020) found that the success of thi s tourism destination & location "was possible due to its adaptive capability in Complex Adaptive System (CAS**). Highly specialized service-oriented regional economies** can enhance regional resilience by improving adaptive capacity towards within-sector related variety. **Diversification through the adaptive capability** of Las Vegas' tourism contributed to economic recovery and resilience".

Romão, Guerreiro, and Rodrigues (2017) point that tourism activities face unprecedented levels of competition in the global economy, territories as destinations are also in competition and not only product and services providers. Tourism is a space-dependent economic sector. Authors conclude that the "local resource management, promotional strategies, transport systems or accommodation provider can be more **efficiently planned** if there is some collaboration among clusters of regions with similar characteristics". The logic of this approach, according to authors, allow developing "a strategy of differentiation aiming at the provision of unique experiences based on the specific territorial resources could lead to a more sustainable form of tourism development" (Romão, Guerreiro, and Rodrigues 2017)

Frocrain and Giraud (2019) classify industries into "tradable and non-tradable categories using an index of geographic concentration since for tradable industries production tends to be geographically separated from consumption". Tourism is also tradable, where **its foreign consumers do the moving,** at the location of tourism service providing and consuming take place. Persons mobility high restrictions as a preventive measure for Coronavirus Pandemic control locked the tourism. The foreign consumer does not access any more any destination location. In consequence all the sectors that depend on tourism (accommodation, cuisines, beverages, cultural events,…etc) are locked almost in an entire share in many countries.

2. Theoretical framework - economy of agglomerations

In general, regardless of the crises is need to increase the capability to deliver smart, sustainable and inclusive growth, to find the path to create new jobs for all and better lives.

The purpose of the work is to identify human capital natural hidden patterns at locality level in Romania (NUTS 5), of employed in tourism mainly sectors, especially with a tertiary level of education, called **talents,** by age and by gender. The secondary purpose is to check if in the most vulnerable areas, with highest concentration of poor people are talents employed in the studied sectors.

Agglomeration of creative workers Moretti (2012) creates positive externaliteis. Talents employed in tradable sectors have high multiplicator effect, in terms of job creation in nontradable sectors. Frocrain and Giraud (2019) conclude that, during 2004-2013 period, in France, for every 100 new tradable jobs, including the jobs in tourism too, that emerged in an employment area, 64 additional non-tradable jobs were created in the same area". Frocrain and Giraud (2019) cites Le Garrec (2008) and justify that some shares from sectors like "*Food and beverage service" (56) \Accommodation" (55), \Travel agency and tour operator activities, and other reservation service and related activities" (79), \Creative, artistic and entertainment activities" (90), \Libraries, archives, museums and other cultural activities" (91), \Gambling and betting activities" (92), all could be included in the tradable group".*

*This investigation has i*mportant implications for public policy in view to identify the externalities needed for successful and adaptive development of mainly tourism industries in regions with a high propensity of smart specialization, regardless their level of development. Clusters defined by the high share of talents in mainly tourism sectors marks **the local labour markets specialised as well the presence of labour market pooling in the analysed sectors**. Measures and actions concentrate and precisely focused in these clusters could accelerate the regional development policies implementation, new knowledge diffusion, new skills development, related variety skills development, increasing productivity, increasing the innovation capacity to acquire and development of the sector, new jobs creation in the tourism sector. On the short term, this process, accordingly adjusted could, could assure the rapid recovery after the Corona Pandemic shock.

Systems, societies and individuals are increasingly interconnected and interdependent in a global world. Also, the tourism world is transformed by competitiveness. In the profit absence tourism sector is notsustainable, practically the incomes need to be higher than the costs. Industry rivalry includes according to Porter (1979): *the customers* who want to receive more and to pay less; bargaining power of *suppliers* who want to be paid more and offer less; the creation of new sectors, respectively rivals from other industries through *the substitution* effect that offers products/services that offer the same functionality at lower costs; *new entrants* to the market by attracting customers from the old producers by reducing costs. Space is at the core of decreasing costs, directly or indirectly, mainly through: innovation and economies of agglomeration.

2.1. Innovation and talents

Innovation is the way to make a profit in the monopolistic competition model, an aspect noted by Stiglitz since 1977. Llano (2000) found that creating the new (incremental or radical) eliminates competitors, rivals, lowers costs and thus maximizes temporary profit. The zero moment is the creation of the new in location, that becomes Center. Next, its neighbours adopt it through dispersion or contagion. Finally, if the case of the whole world. The last locations that adopt this new are the periphery. Fujita, Krugman, Venables (1999) define the Center-Periphery model within the new geographical economy. This model describes three localization forces: two agglomeration / centripetal forces given by the relationship between costs and demand and a dispersion / centrifugal force determined by the local competition.

2019 European Innovation Score Board (EIS) treats the Innovation as input-output. Human resource, next to the attractiveness of research systems and innovation-friendly environment stand to the framework conditions for innovation, that are the critical innovation promoters of innovation performance, external for the enterprise. In EIS human resources cover: new doctorate graduates, population completed tertiary education and lifelong learning. In EIS Romania is ranked 28 as a modest innovator, still below the 2011 performance, but better than the last minimum form 2015. 2019 Regional Innovation Scoreboard (RIS) relative performance to EU in "2011" ranks Bucharest Ilfov as the pest performance with 54.1 2019 RIS in progress with 5.8 compared to 2017, but decreasing with -7.9 compared to 2011. All other regions present the innovation performance lower than 1/3 from the EU average in 2019, with the exception of Vest region with 34.3.

Other international systems that measure innovation and talent are the World Economic Forum (WEF)'s Global Competitiveness Index (GCI) and OECD Science, Technology and Industry Scoreboard (STIS).The GCI 2019 Report points that „countries must improve talent adaptability; that is, enable the ability of their workforces to contribute to the creative destruction process and cope with its disruptions. Talent adaptability also requires a well functioning labour market that protects workers rather than jobs." (Schwab 2019)

OECD STIS 2017 concludes that „Workers in digitally intensive industries exhibit both higher levels of cognitive skills (e.g. literacy, numeracy and problem solving), as well as non-cognitive and social skills (e.g. communication and creativity)". In this new context, tertiary education defines the best to „meet the rising demand for cognitive skills" .(OECD 2017)

Wei, Feng, and Zhang (2017) points, based on literature, that "innovation talents presence, as a most active and important resource in innovation activities; all that innovation talents make significant contributions to the improvement of innovation capability, but not vice versa".

2.2. The economies of agglomeration

The economies of agglomeration are one of the main determinants of the variation of productivity in the space. Frenken, Van Oort, and Verburg (2007, p.6) The central idea underlying the economics of agglomeration holds that clustering of economic activity occurs because firms experience some form of benefit from locating near one another. A broad definition of agglomeration economies is that it concerns economies from which a firm can benefit by being located at the same place as one or more other firms. They are related to economies of scale, which play an essential role in increasing productivity. The economies of agglomeration reduce the average costs of a long-term product, the result of an expansion of an activity (Bogart 1997). Fujita and Thisse (1996; 2013) address the "economic reasons for the existence of a large variety of agglomerations arising from the global to the local". They found that "the trade-off between various forms of increasing returns and different types of mobility costs are more fundamental" than natural features of geography. Sources of reduction of average costs are defined relative to the industrial unit frontier, as internal or external economies of scale.

Four sources of agglomeration economies (spillovers) are distinguished: one internal and three external - (Marshallian, Urbanisation and Jacobian). We add to these the fifth – the Morettian externalities or the Tertiary Human Capital Spillover (Moretti 2003). Under evolutionary trade theory and evolutionary growth theory (Saviotti & Pyka, 2004; Vernon, 1966), emphasises the distinction between the different sources of spillovers by qualitatively different types of benefits.

Internal economies of scale achives within an industry unit (a company, an industry, a city, a state, a region). The economies derive from the spatial expansion of the industry unit by lowering the costs from internal sources: technological, managerial, financial or risk mitigation. These may occur in a single firm due to production cost efficiencies realized by serving large markets (Krugman 1991), id global spread. There is nothing inherently spatial in this concept according to Frenken et al. (2007, p6), other factors than that the existence of a single large firm in space implies a large local concentration of factor employment;

External economies of scale or externalities achives within the outside of the industry unit. The spatial closeness of firms could lead to positive externalities, called agglomeration economies. These externalities could bepositive or negative (ex congestion, pollution).

Marshallian economies (technical externalities) or localization economies achieve benefits by neighbourhood with companies from the same sector/industry, respectively by specialization mechanisms. This type of external economies is available to all local firms within the same sector. Feser (2002) and Henderson (2003) points that Marshallian externalities arise from three sources: **labour market pooling,** creation of specialized suppliers, and the emergence of knowledge spillovers. This externality explains the productivity heterogeneity in space, whereby the productivity of labour in a given sector in a given city is assumed to increase with total employment in that sector. In short, Frenken, Van Oort, and Verburg (2007) prove that the expected benefit of the Marshallian externality is **productivity increase**. ***This productivity-****increasing is specific to the location where similar firms produce similar products. The spatial proximity expecte to "****spur incremental innovation and process innovation****"*. Consoli & Sanchez-Barrioluengo (2019) found that "**local industrial specialisation** shaped by an agglomeration of a high□skilled workers on the **demand for no tradable service jobs".** This tertiary educated human capital agglomeration has a positive local multiplier effect on the expansion of low□skill service employment and shortage of the "routine" mid skill jobs". The agglomeration of the firms, from the same sector, in some areas give the specialized character to its region. These companies realize mutual benefits, according to Beaudry and Schiffauerova (2009) through access **to highly specialized workforce pools**. The decrease of the costs of the demand is realized by the direct access to the labour force, the search costs of the labour force are reduced. The costs of the labour supply are also reduced by the existence of various employment/employment opportunities, reduced costs of intermediate inputs, communication and innovation spread. High specialisation in a sector vulnerable to global shocks could **hinder economic growth of the region.**

Urbanisation economies are external economies available to all local firms irrespective of sector and arising from urban size and density. These economies are the result of savings from the large-scale operation of the agglomeration or city as a whole and independent from industry structure. Relatively more populous localities are also more likely to house universities, industry research laboratories, trade associations and other knowledge-generating organizations. It is the dense presence of these organizations (not solely economic, but also social, political and cultural) that supports the production and absorption of know-how, stimulating innovative behaviour, and contributes to differential rates of interregional growth. Harrison, Kelley, and Gant (1996). In this case (urbanization economies), the sectors are unrelated; therefore, according to Frenken et al. (2007, p24) the expected benefit is the unemployment decrease. Raspe and Van Oort (2006) recommend evaluating the **economic potential of cities** based on three knowledge factors: "R&D", "innovation" and "knowledge workers". The R&D is input in practice in spatial economic strategies to build regions specialization. Raspe and Van Oort (2006) found that the "the successful introduction of new products and services to the market ("innovation") and indicators of skills of employees ("knowledge workers") explains better and more profoundly the urban employment and productivity growth".

Jacobian externalities are external economies stemming from a variety of sectors. Jacobs (1961; 1969) In a region, the variety of sectors are related or unrelated (Frenken et al., 2007, p24) define the related variety is „a source of regional knowledge spillovers, called Jacobs externalities". Unrelated variety is "a portfolio protecting a region from external shocks". The expected benefits, according to Frenken et al. (2007, p24), in Jacobian externalities, is employment growth as a result of **radical innovation processed finalized with new products that are associated with new skills, new employment and new markets.** Authors point that **"related variety in cities is responsible for job creation and not urban density in itself".** Since 1998, Quigley found that "the functional specialization of firms in heterogeneous industries, near of each other, is supposed to generate spatial interdependencies and generates benefits (and costs such as congestion) for everyone in that specific location". Quigley (1998) **Thus, variety in itself may be a new source of knowledge spillovers and innovation. Frenken, Van Oort, and Verburg (2007) explains** that "the diverse industry mix in an urbanized locality also improves the opportunities to interact, copy, modify and recombine ideas, practices and technologies across industries giving rise to Jacobs externalities. Important innovations stem from the recombination of knowledge present in different industries. Geographical proximity between firms in different industries renders such recombination more likely to occur, in particular, if firms also operate under similar institutional conditions".

Beaudry and Schiffauerova (2009) calls **urbanization and Jacobian externalities in the same time.** In this case, works the **diversity mechanisms of sectors, which support the economic growth of the region**. These cases are specific to companies located in a big city, where the neighbouring companies do not belong to the same industry. The sources of benefits are:

- access to a large market, consumption can also be local, lower transport costs;
- access to a variety of specialized services available only in large cities;
- potential for the dissemination of knowledge and technology from different industries, including access to a much more diverse technological spectrum offered by research. Urban agglomeration offers benefits of lower Jacobian costs, according to the previous context. The peripheral centre structure reflects spatial models of optimizing the use of resources by **reducing the costs of transport, energy and knowledge/information - also called connectivity costs.**

The last sources of externalities are the Morettian externalities or the Tertiary Human Capital Spillover. These refer to the situation when high human capital have a spillover effect on a city economy (Moretti 2003). The magnitude of human capital spillover is tremendously significant for education policy, for new talents attraction in the region, to increase the productivity, voter participation increasing, crime and violence behaviour decreasing, etc. Moretti (2004) demonstrated that supplementary to agglomeration this effect is observable. This last externality presents interest, especially in the context of the need to accelerate the digital transformation of the tourism sector.

2.3. Tourism - concept and measurement approaches

Regulation 692/2011 concerning European statistics on tourism (and repealing Council Directive 95/57/EC) revises and updates Council Directive 95/57/EC and takes into account the internationally recommended methodology which is provided in the IRTS 2008 define 'Tourism' as "the activity of visitors taking a trip to a main destination outside the usual environment, for less than a year, for any main purpose, including business, leisure or other personal purpose, other than to be employed by a resident entity in the place visited."(Eurostat 2014)

Based on the *Methodology for tourism statistics* (Eurostat 2014), *Tourism Satellite Accounts* (TSA 2010) and The International Recommendations for Tourism Statistics 2008 (IRTS 2008) the "**mainly tourism**" industries includes the four digits detailed activities: H5110 Passenger air transport, I5510 Hotels and similar accommodation, I5520 Holiday and other short-stay accommodation, I5530 Camping grounds, recreational vehicle parks and trailer parks, N7910 Travel agency and tour operator activities. (Table 6 and 7)

We intend to explore patterns of employment in Romania's mainly tourism sectors at two digits, at the NUTS 5 level. NIS - National Institute of statistics provide this data granulation by 2011 Census data provided. (H51Air transport, I55 Hotels and other accommodation facilities and N79 Activities of tourist agencies and tour operators; other reservation services and tourist assistance)

2.3.1. Travel agency and tour operator is a tourism engine is of strategical importance to support jobs and inclusive growth in all regions

Travel agency, tour operator reservation service and related activities (N79) is a high human capital concentrator, included by Eurostat in the **Knowledge Intensive Activities (KIA).** According to Moretti (2010), the agglomerations of high human capital predict the success of locations and provide a multiplier effect. N79 create demand for Transportation and Storage (H, especially H59 – air transport)) and Accommodation and Food Service Activities (I, mostly accommodation (NACE I55), food and beverage service activities (NACE I56)). Together with knowledge, transport and accommodation shape the tourism industries, strongly linked with culture and spirituality. N79 is the highest high human capital concentrator among all tourist industries, the fact that allows it to benefit from technological innovation OECD (2007) fully. **N79 is a Less Knowledge-Intensive Market Services (LKIMS)** from the perspective of Employment classification by the intensity of technology and knowledge usage. That means this sector share the KIS sectors characteristics: generate high value-added (over 21% according to OECD 2007), have a high-intensity information use, based on ICT and digitisation adoption is increasing its efficiency and quality of services, services that provide in highly competitive markets, N79 becoming an innovation creator not only adopter (Bryson and Daniels, 2007.p 179) N79 as KIS presents a multidimensional profile: (i) knowledge – generate and exploit; (ii) innovation and (iii) spatial proximity /regional dimension. Muller and Doloreux (2007). On the background that "Europe continues to stand as the most visited region, welcoming half of the world's international tourist arrivals" World Tourism Organization (UNWTO, 2018). N79 as tourism engine is of strategical importance to support jobs and inclusive growth in all regions. EU28 at regional level (NUTS2) presents high spatial heterogeneity, with different specialisation degrees and competitive performance levels.

2.4. Romanian tourism industries performance

2.4.1. Romanian tourism industries short profile in the context of the economy

Following the structures of (Eurostat. 2020) and using Eurostat data, Table 1 and 2 present the brief pattern of Romanian tourism industry in the context of an economy in 2017. The 690 thousand enterprises made a turnover of 345 million euros, respectively, a value added at factor cost of 88.8 million Euros and employed 5.35 million of persons. (Table 1) In relative terms, total tourism industries cover 6.2% of enterprises, 2.1% from turnover, 2.6% from value added and 4.7% from employment. Mainly tourism industries comprise 1.3% of enterprises, 0.8% from turnover, 0.9% from value-added and 1.2% from employment. Mainly the tourism sector has some gross multiplication effects over partially tourism across the key indicators ratio indicate, as follows: 3.8 for enterprise number, 1.5 for turnover, 1.9 for value-added and 2.8 for employment. These ratios are comparable with the proportions at EU27 level (Eurostat 2020), respectively of 4.9 for enterprise number, 1.4 for turnover, 1.9 for value-added and 2.8 for employment.

Before the Corona Pandemic, at one mainly tourism employee was espected to have 3 employees in partially tourism sectors.

Table 1. Romanian tourism industries short profile in the context of economy in 2017 (absolute values)

	Number of enterprises	Turnover (million EUR)	Value added at factor cost (million EUR)	Number of persons employed
Business economy other than services (1)	485215	292990	66909	4020121
Services other than tourism industries (2)	162306	44682	19154	1079148
Total tourism industries (3)	42603	7395	2311.6	249744
Tourism industries (mainly tourism) (4)	**8960**	**2912.9**	**800**	**65941**
Tourism industries (partially tourism) (5)	33643	4482.1	1511.6	183803
Total	690124	345067	88374	5349013

Note: Due to unreliable data at country level and rounding, deviations can occur between total and subtotal
(1) B-N_S95_X_K; (2) NACE sections H, I, J, L, M, N, S95; (3) NACE classes H491, H4932, H4939, H501, H503, H511, I551, I552, I553, I561, I563, N771, N7721 and division N79; (4) NACE classes (H51+I55+N79) RIS data – Romanian Institute of Statistics 2011 Census data; (5) NACE classes H491, H4932, H4939, H501, H503, I561, I563, N771, N7721 and N799; (6) H49+H50+H51; Source: Eurostat (online data code: sbs_na_sca_r2, sbs_na_1a_se_r2)

Table 2. Romanian tourism industries short profile in the context of economy in 2017 (relative values)

	Number of enterprises, %	Turnover (million EUR), %	Value added at factor cost (million EUR), %	Number of persons employed (hundred), %
Business economy other than services (1)	70.3	84.9	75.7	75.2
Services other than tourism industries (2)	23.5	12.9	21.7	20.2
Total tourism industries (3)	6.2	2.1	2.6	4.7
Tourism industries (mainly tourism) (4)	**1.3**	**0.8**	**0.9**	**1.2**
Tourism industries (partially tourism) (5)	4.9	1.3	1.7	3.4
Total	100	100	100	100

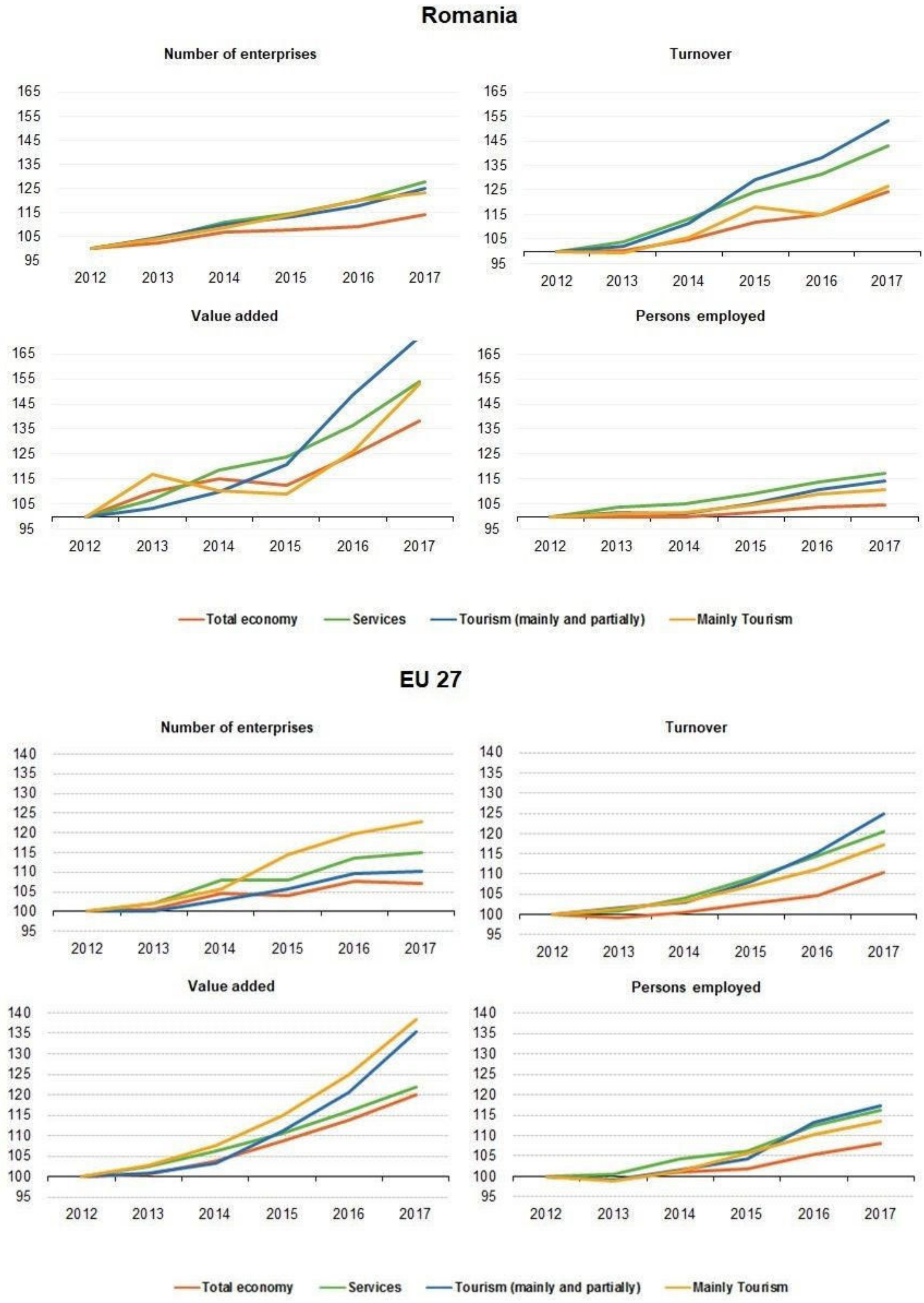
Note: Due to unreliable data at country level and rounding, deviations can occur between total and subtotal
(1) B-N_S95_X_K; (2) NACE sections H, I, J, L, M, N, S95; (3) NACE classes H491, H4932, H4939, H501, H503, H511, I551, I552, I553, I561, I563, N771, N7721 and division N79; (4) NACE classes (H51+I55+N79) RIS data – Romanian Institute of Statistics 2011 Census data; (5) NACE classes H491, H4932, H4939, H501, H503, I561, I563, N771, N7721 and N799; (6) H49+H50+H51; Source: Eurostat (online data code: sbs_na_sca_r2, sbs_na_1a_se_r2)

2.4.2. Tourism and economy evolution by key indicators of the economy, Romania compared with EU27, during 2012-2017

Evolution of tourism (mainly and partially) present a positive trend, overpassing the total economic performance, in all the four key indicators, in Romania as well as at EU27 level (on average). (Figure 1)

Mainly tourism is performing as the best performances, over services and tourism for enterprises number and value-added at EU27 level. This fact indicates the engine role of mainly tourism sector in the economy. In Romania, even if the industry of mainly tourism is increasing with great pace is still below the service and tourism performances. Comparing the level of 2012 fixed base key indicators for Romania with EU27 average is visible that mainly tourism increase in Romania is higher for value-added (153% for Ro and 138% for EU27) and turnover (127% for Ro and 117% for EU27), equal for enterprise number (123%) and slightly below for employment (111% for Ro and 113% for EU27).

Figure 1: Tourism and economy evolution by key indicators of the economy, Romania and UE27, 2012-2017 (index 2012=100)



Source: Eurostat (online data code: sbs_na_sca_r2, sbs_na_1a_se_r2), (Eurostat 2020)

3. Method

Aydin (2019) announced that GIS & Spatial Machine Learning: Transforming Our Planet's Pulse to Action. Artificial intelligence includes Machine Learning, and Machine Learning included Depp Learning, in a Matrioska relationship. **The new logic is to learn rules and patterns from data, while explicit regulations and knowledge do not exist; only data represented explicitly.** Aydin (2019) **The result is knowledge and rules inferred from data.** Mitchell (1997) **define machine learning as a** "computer program is said to learn from experience E concerning some class of tasks T and performance measure P, if its performance … improves with experience E". AMchine Learning in GIS uses data -driven algorithms and techniques that automate tasks as "prediction, classification and clustering". The novelty of ML in Arc GIS is the incorporation of geography in compution id. shape, density, contiguity, spatial distribution, or proximity". (Aydin 2019)

The method is **Multivariate Clustering Analysis** (MCA), the non-spatial (constrained) version, applied in Arc Gis Pro 2.3. (Box 1). Since the release of Arc GIs Pro is available a new set of Statistical Spatial Tools, including MCA K Means. (Aydin 2019) classify Multivariate Clustering among the Machine Learning Tools in Arc GIs, next to Spatially Constrained Multivariate Clustering, Density-based Clustering, Image Segmentation, Hot Spot Analysis, Cluster and Outlier Analysis, Space-Time Pattern Mining and Time Series Clustering.

Ruthartr (2018) points out that:

> "**Clustering algorithms are a type of unsupervised machine learning**, meaning you don't have to define what it means to be a cluster up front (often referred to as training the model). Instead, the algorithm does that for you by evaluating the data and **finding natural patterns that exist**".

MCA algorithm finds "natural subsets or groupings of features based on either location (spatial component only), values (attributes only) or a combination of both location and values".

K-means algorithm developed by (Steinhaus 1957) and named by (MacQueen 1967). (Ruthartr 2018)

Arc Gis Pro software uses MCA with K-means in the idea shaped by (Jain 2010):

> "Organizing data into sensible groupings is one of the most fundamental modes of understanding and learning. Cluster analysis is the formal study of methods and algorithms for grouping or clustering, objects according to measured or perceived intrinsic characteristics or similarity. Cluster analysis does not use category labels that tag objects with prior identifiers, i.e., class labels. The absence of category information distinguishes data clustering (unsupervised learning) from classification or discriminant analysis (supervised learning). The aim of clustering is to find structure in data and is therefore exploratory in nature. Clustering has a long and rich history in a variety of scientific fields. One of the most popular and simple clustering algorithms, K-means, was first published in 1955." (Jain 2010)
>
> This tool is appropriate for big data sets, Big Data. It has applied successfully to market segmentation (Wikipedia, 2020). Also, For this reason, we exploit fully the Census data, in a new manner. We apply MCA on polygon data geocodate, as the spatial component, represented by Local Administrative Units, respectively the NUTS 5 level. The employment data by the 8 characteristics, splitted in 8 density variable are the attributed data associated with the spatial location."

Our objective is to reveal natural **hidden patterns which are very difficult to see just by looking at the data points on the map.**

3.1. Multivariate Clustering Analysis

MCA tool is relatively recently used in literature. Moral et al. (2016) made a GIS-based multivariate clustering for characterization and ecoregion mapping from a viticultural perspective. Authors delineate homogeneous zones by climate and big topographical data with high variability, for Extremadura (southwestern Spain), an outstanding wine region. Romão, Guerreiro, and Rodrigues (2017) concludes that spatial analysis "is a useful contribution of in tourism studies with a clear impact on the goodness of the of the econometric model and the identification of spatial patterns in tourism activities and its determinants"." Michaelides, Economakis, and Lagos (2006) uses Multivariate Clustering Analysis for employment and regional planning in Greece. Bena et al. (2013) apply MCA to analyse the job tenure and work injuries in relation to previous experience and difference by age. Tatarczak and Boichuk (2017; 2018) apply multivariate methods and explore the nature of youth unemployment and unemployment in Poland in more precise detail using dendrograms.

We use **MCA, K Means algorithm (K Means++), to identify seeds used to grow the cluster.** The result is to "organise, group, differentiate and catalogue" mainly tourism employment by age, education and gender characteristics by eigh characteristics. The feature is the NUTS 5 / local administrative unit, a spatial unit area. The employment data are the attribute data Pierre (2014), in our case described by eight variables. For this purpose, we use the following notations the analysed groups, defined by employee characteristics:

a) Level of education

1 – Low (ISCED 0-2, which corresponds to a level of education at best equivalent to lower secondary education);

2 – Medium (ISCED 3, which corresponds to a level equivalent to upper secondary education);

3 – High (ISCED 5-7, which corresponds to a level equivalent to tertiary education);

where we use the notations:

PP_51_1+ PP_51_2 +PP_51_3=100% (1)

PP_55_1+ PP_55_2 +PP_55_3=100% (2)

PP_79_1+ PP_79_2 +PP_79_3=100% (3)

Economic sectors:

H51 Air transport,

I55 Hotels and other accommodation facilities and

N79 Activities of tourist agencies and tour operators; other reservation services and tourist assistance,

b) Age intervals:

T – Youth (15-24 years old);

A - Adults (25-54 years old);

V – Old (55-64 years old);

where we use the notations:

PP_51_T+ PP_51_A +PP_51_V=100% (4)

PP_55_T+ PP_55_A +PP_55_V=100% (5)

PP_79_T+ PP_79_A +PP_79_V=100% (6)

c) Gender- M (males); F - (females).

where we use the notations:

PP_51_F+ PP_51_M =100% (7)

PP_55_F+ PP_55_M =100% (8)

PP_79_F+ PP_79_M =100% (9)

Result 8 groups, iterated in the following codification by share and by main tourism sector:

Digits 1 to 2 are the share symbol;

Digits 4 to 5 are the sector code;

Digits higher than 6 are the employee's characteristics;

Limit of the method: we made a tradeoff between the best spatial granularity and conceptual concordance with the sector definition. As it is visible in table 8 and 9 from Annex result in an overestimation of employment is mainly the tourism sector with 7.9 thousand persons, respective with 13.6%. Also, data are from different sources with limits of comparability. Our data are RIS data – Romanian Institute of Statistics 2011 Census and the Tourism data matrix are from SBS Eurostat.

The objective is to **create clusters (N) / or groups, as similar as possible** of local labour market classified by the similarity of the features for the eight groups mentioned [(1-Low, 2-Medium and 3- high), age (T- youth, A – adults and V – aged) and gender (M- males, F-Females)], across NUTS 5 locations from national territory of Romania (called here features). The resulted number of clusters/groups follow the rule that the "**features within each output group are as similar as possible while groups are as different as possible"** (Pierre 2014). The total number of local administrative units is 3189, higher than the 30, the minimum number of features for MCA. Each of the eight groups is analysed by their similarities that reflect differences in each characteristic share, resulting in three spatial patterns typologies, for each sector (H51, I55 and N79).

Each group is analysed by the eight employees characteristics variable.

> "The values of the Analysis Fields are standardized by the tool because variables with large variances (where data values are very spread out around the mean) tend to have a larger influence on the clusters than variables with small variances. Standardization of the attribute values involves a z-transform, where the mean for all values is subtracted from each value and divided by the standard deviation for all values. Standardization puts all the attributes on the same scale." (ESRI ArcGis Pro, 2020)

Spatial unit is NUTS 5, respectively country level, and the software is ARC GIS Pro. Multivariate clustering analysis (MCA) tool uses the K-Means algorithm.(ESRI ArcGis Pro, 2020) The natural clusters are identified directly from the data, and MCA is an unsupervised machine learning method. Data are grouped in clusters. Here "*all the features within each cluster are as similar as possible, and all the clusters themselves are as different as possible*". (ESRI ArcGis Pro, 2020) MCA is not a spatial tool but produces a spatial pattern of transitional labour markets by gender and age.

The locality level is the lowest administrative unit codified as NUTS5 according to EUROSTAT, equivalent to LAU2 Local Administrative Unit, classified according to SIRUTA. The Information System of Administrative-Territorial Units Register - SIRUTA, is a fundamental tool in automatic data processing in the territorial both statistical system and the economic system – general Financial in Romania. It works based on the legal framework: Law no. 2/1968, Decree-Law no. 38/1990 as a legal trustee's regional organization of Romania, HG no. 575 bis / 1992, CNS President Order 817/1994.

We use 2011 Census data provided by the Romanian National Institute of Statistics.

Method (Box 1)

The R2 value reflects how **much of the variation** in the original TestScores data was retained after the clustering process, so the larger the R2 value is for a particular variable, the better that variable is at discriminating among your features. (ESRI ArcGis Pro, 2020)

$$R2 = (TSS - ESS) / TSS \tag{10}$$

TSS is the total sum of squares
ESS is the explained sum of squares

Number of cluster k

MCA clustering effeteness is measured "*using the Calinski-Harabasz pseudo F-statistic, a ratio reflecting within-group similarity and between-group difference*:

$$F = \frac{\frac{R^2}{n_c-1}}{\frac{1-R^2}{n-n_c}} \tag{11}$$

Where:

$$R^2 = \frac{\mathrm{SST} - \mathrm{SSE}}{\mathrm{SSE}} \tag{12}$$

SST is a refelction of between-cluster differences
SSE reflects within-cluster similarity

$$SST = \sum_{i=1}^{n_c} \sum_{j=1}^{n_i} \sum_{k=1}^{n_v} (V_{ij}^k - \overline{V^k})^2 \tag{13}$$

$$SSE = \sum_{i=1}^{n_c} \sum_{j=1}^{n_i} \sum_{k=1}^{n_v} (V_{ij}^k - \overline{V_t^k})^2 \tag{14}$$

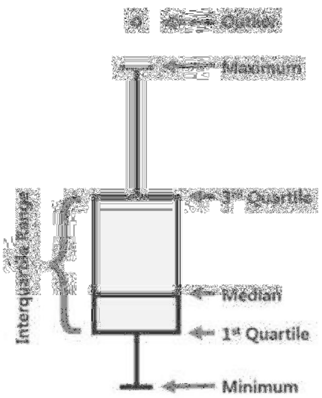
n = the number of features
n_i=the number of features in cluster i
n_c=the number of classes (cluster)
n_V=the number of variables used to cluster features
V_{ij}^k= the value of the k^{th} variable of the j^{th} feature in the i^{th} cluster
$\overline{V^k}$= the mean value of the k^{th} variable
$\overline{V_t^k}$= the value of the k^{th} variable in cluster i

Evaluating Number Clusters, a chart will be created showing the pseudo F-statistic values calculated. The highest peak on the graph is the largest F-statistic, indicating how many clusters will be most effective at distinguishing the features and variables you specified. Based on pseudo F-statistic chart we select the best k and run again the tool. Among results, nexst to the ***cluster map are the number of features per clusters****.*

Clustering Method

The Multivariate Clustering tool uses the K Means algorithm by default. The goal of the K Means algorithm is to partition features so the differences among the features in a cluster, over all clusters, are minimized. Because the algorithm is NP-hard, a greedy heuristic is employed to cluster features. The greedy algorithm will always converge to a local minimum, but will not always find the global (most optimal) minimum. The K Means algorithm works by first identifying seeds used to grow each cluster. Consequently, the number of seeds will always match the Number of Clusters. The first seed is selected randomly. Selection of remaining seeds, however, while still employing a random component, applies a weighting that favours selection of subsequent seeds farthest in data space from the existing set of seed features (this part of the algorithm is called K Means ++). Because of the random component in finding seeds whenever you select Optimized seed *locations or Random seed locations for the Initialization Method, you might get variations in clustering results from one run of the tool to the next.*

Outputs



Box plots *are used to show information about both the characteristics of each cluster as well as characteristics of each variable used in the analysis. The graphic below shows you how to interpret box plots and their summary values for each Analysis Field and cluster created: minimum data value, 1st quartile, global median, 3rd quartile, maximum data value, and data outliers (values smaller or larger than 1.5 times the interquartile range). Hover over the box plot on the chart to see these values as well as the interquartile range value. Any point marks falling outside the minimum or maximum (upper or lower whisker) represent data outliers. The default parallel box plot chart summarizes both the clusters and the variables within them.Each node of the mean lines points the cluster's average value for each Analysis Field.*

Source: (ESRI ArcGis Pro, 2020)

4. Findings

Our main findings are the cluster patterns of Employment in Romania for each sector of the mainly tourism activities (H51, I55 and N79) at NUTS 5 level. **The clusters group features (local administrative units) with similar characteristics described by eight variables**. The statistics of variables characteristic by feature coupled with the number of clusters (N). selected N are generate statistics by the eight variable characteristics of employment associated with, the degree of goodness (R^2) that variable is discriminating among the features. **Final results are the cluster maps and the number of features per clusters** (Bennett, Vale, and d'Acosta 2015).

All three clusters profile do not have outliers. We apply the K Means algorithm .

4.1. Clusters profiles for H51 Air transport

The H51 sector employees mainly males adults with a secondary level of education. The average share of employees with tertiary level of education accounts 6% mean at NUTS 5 level.

Table 3. Statistics for employment in H51 at NUTS5 level, by shares in total for the level of education (1-Low, 2- Medium and 3- high), age (T- youth, A – adults and V – aged) and gender (M- males, F- females) and R^2

Variable	Mean	Std. Dev.	Min	Max	R^2 N=20	R^2 N=6
PP_51_A	13.9745	33.4090	0.00	100.00	0.984839	0.961982
PP_51_2	9.4201	26.6869	0.00	100.00	0.966076	0.855151
PP_51_M	11.9032	30.5678	0.00	100.00	0.962084	0.380300
PP_51_3	**6.0748**	**21.2066**	**0.00**	**100.00**	**0.957928**	**0.788985**
PP_51_T	1.7547	11.4439	0.00	100.00	0.955494	0.796854
PP_51_1	1.7833	11.5223	0.00	100.00	0.921003	0.834791
PP_51_F	5.3750	19.8586	0.00	100.00	0.910164	0.738453
PP_51_V	1.5490	9.8298	0.00	100.00	0.89477	0.818607

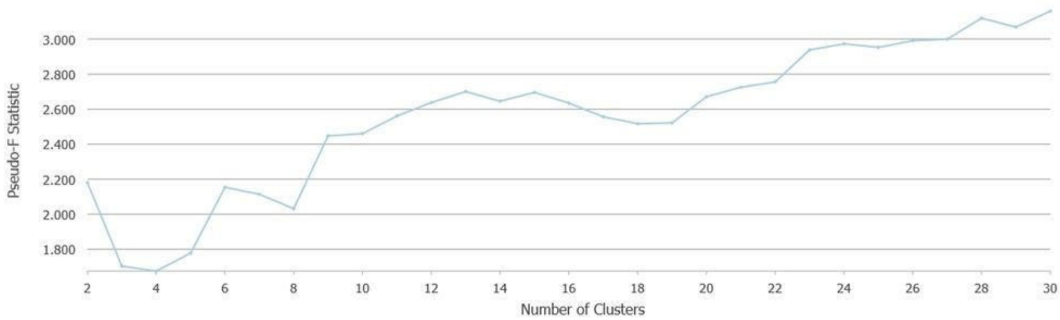
Note: N number of clusters

Source: data calculated by authors

The optimal number of cluster is 30. (Image 1) The clustering effectiveness is measured using the Calinski-Harabasz pseudo-F-statistic, which is a ratio of between-cluster variance to within-cluster variance**.** (ESRI ArcGis Pro, 2020) For the policies purpose, this number of clusters is too large. The criteria selection is given by the rule “the highest peak is the largest F-statistic, indicating how many clusters will be most effective at distinguishing the features and variables you specified.”

We run the MCA a number times, selecting the results for N 20 and 6. While N=6 is the first maximum in the pseudo F statistic share, distinguishes feature similarities and differences in an acceptable level. F level indicates that this variable divides the employees into clusters most effectively. The R2 value for N=6 reflects that the Male variable does not discriminate successful among our features. While out interest variable is the proportion of employees with a tertiary level of education, we consider acceptable this cluster division.

Image 1. Optimized Pseudo-F Statistic Chart for H51, N Optimal = 30 clusters

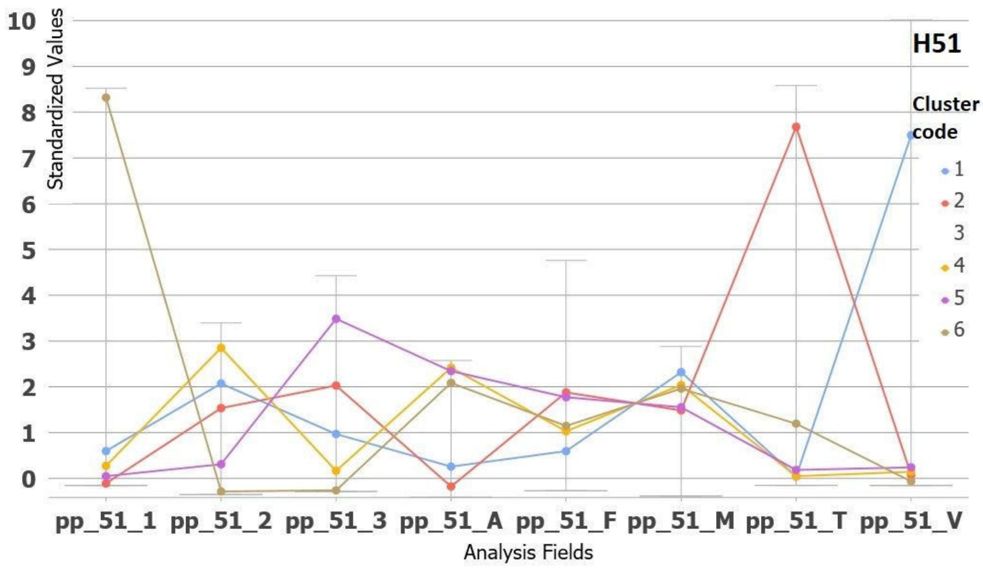


Source: graphic made by authors in Arc Gis Pro

The results of MCA are of 3 types: box plots, features per cluster chart and the MCA cluster map. All these results are linked to each other, each cluster identified is differentiated by a unique colour. In our case, MCA run on census tract to create 6 clusters.

The Box-Plots (Image 2) show information about “both the characteristics of each cluster as well as characteristics of each variable used in the analysis”. Our cluster of interest is Cluster number 5 (magenta colour). (Image 2) The cluster with pp_51_3 is the cluster with the highest proportion of employees with tertiary education. In opposition, Cluster 6 reflect the tract of employees with the low level of education. Cluster number 2 includes the locations with the highest share of youth from the same sector. Cluster number 1 offer the image for tracts of the place with the highest percentage of aged employees. Cluster 3 (white) do not present activity in the sector H51, regardless of the eight characteristics, there is no employment visible.

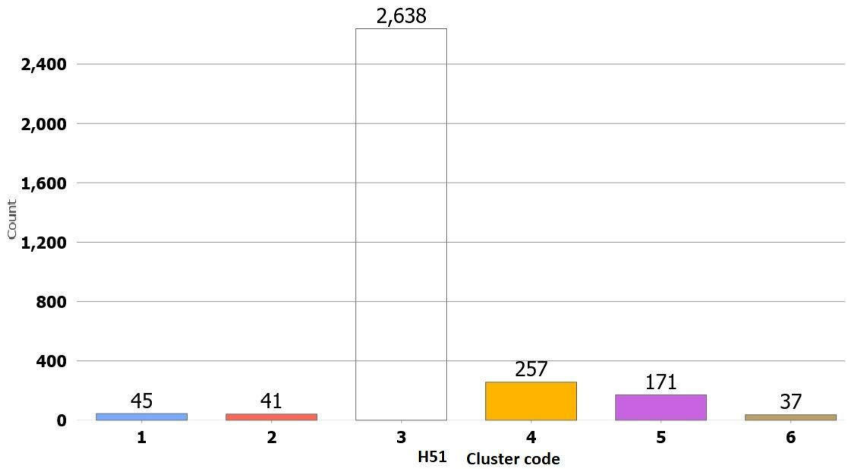
Image 2. Multivariate clustering Box-Plots, for H51, N=6



Source: graphic made by authors in Arc Gis Pro

The high concentration of H51 employees with tertiary education is in 171 features (NUTS5). (Image3)

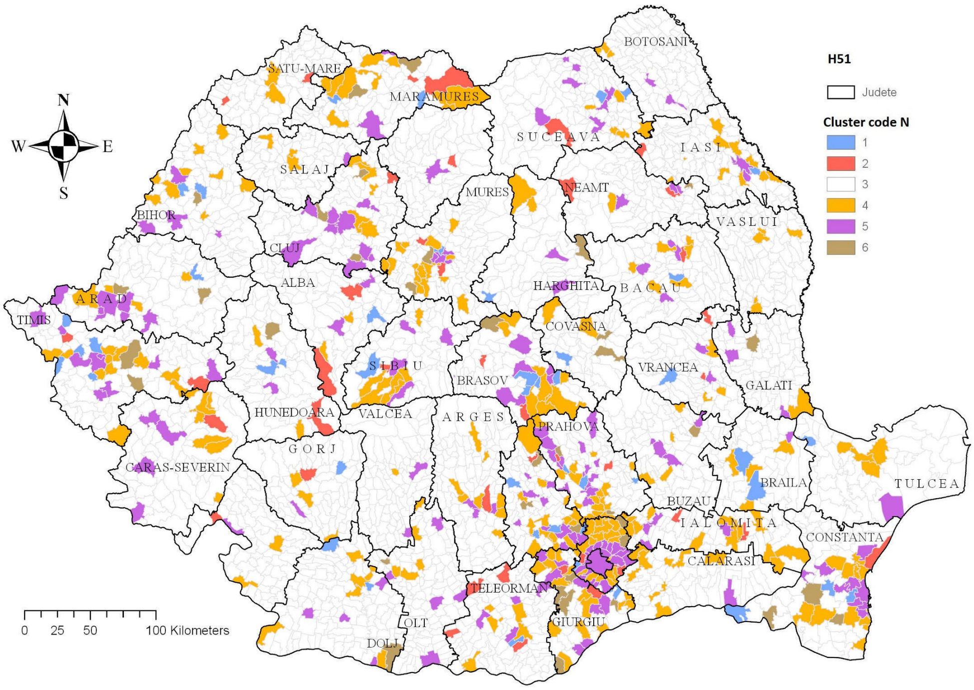
Image 3. Features per Cluster Chart, for H51, N=6



Source: graphic made by authors in Arc Gis Pro

Features with high tertiary employment in H51 are visible in Image 4 in Cluster 5 (magenta), an *indication of high-intensity cognitive activities*. It is noticeable the spatial autocorrelation and concentration around big cities. It could cover urban and periurban areas. The highest agglomeration is around Bucharest and Constanța. Mainly around Bucharest, Brașov and Constanța are locations included in Cluster 4 (mustard). This cluster is principally described by employees with the following characteristics: secondary level of education, adults and males. Their presence in the neighbourhood of Cluster 5 indicates operationalisation and execution tasks.

Image 4. Multivariate Clustering Chart outputs for H51, N=6 by shares in total for the level of education (1-Low, 2- Medium and 3- high), age (T- youth, A – adults and V – aged) and gender (M- males, F- Females)



Source: RPL 2011 NUTs 5 data, NIS source. Map made by authors in Arc GIS Pro with ESRI Romania shapefiles

4.2. Clusters profiles for I55 Hotels and other accommodation facilities

The I55 sector employees mainly females adults with a secondary level of education. The average share of employees with tertiary level of education accounts 10% means at NUTS 5 level. (Table 4)

Table 4. Statistics for employment in I55 at NUTS5 level, by shares in total for the level of education (1-Low, 2- Medium and 3- high), age (T- youth, A – adults and V – aged) and gender (M- males, F- Females) and R^2

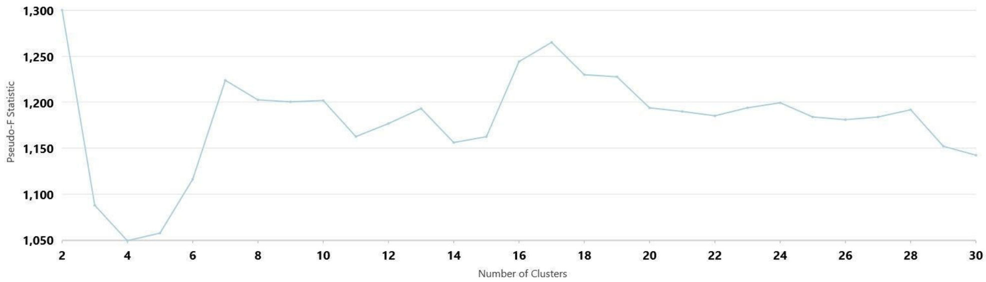
Variable	Mean	Std. Dev.	Min	Max	R^2 N=2
PP_55_A	51.868713	42.981966	0.00	100.00	0.629062
PP_55_F	47.995086	41.5676	0.00	100.00	0.575889
PP_55_2	41.665612	9.864743	0.00	100.00	0.471881
PP_55_M	21.838717	30.640288	0.00	100.00	0.219444
PP_55_1	17.651683	29.363785	0.00	100.00	0.156100
PP_55_T	15.31516	27.263364	0.00	100.00	0.136314
PP_55_3	**10.516511**	**21.576547**	**0.00**	**100.00**	**0.102621**
PP_55_V	2.649915	10.607766	0.00	100.00	0.026957

Note: N number of clusters
Source: data calculated by authors

The optimal number of cluster is 30. (Image 5) For the policies purpose, this number of clusters is too large. The criteria selection is given by the rule "the highest peak is the largest F-statistic, indicating how many clusters will be most effective at distinguishing the features and variables you specified."

We run the MCA a number times, selecting the results for N 20 and 7. While N=7 is the first maximum in the pseudo F statistic share, distinguishes feature similarities and differences in an acceptable level. F level indicates that this variable divides the employees into clusters most effectively.

Image 5. Optimized Pseudo-F Statistic Chart for I55, N Optimal = 30 clusters

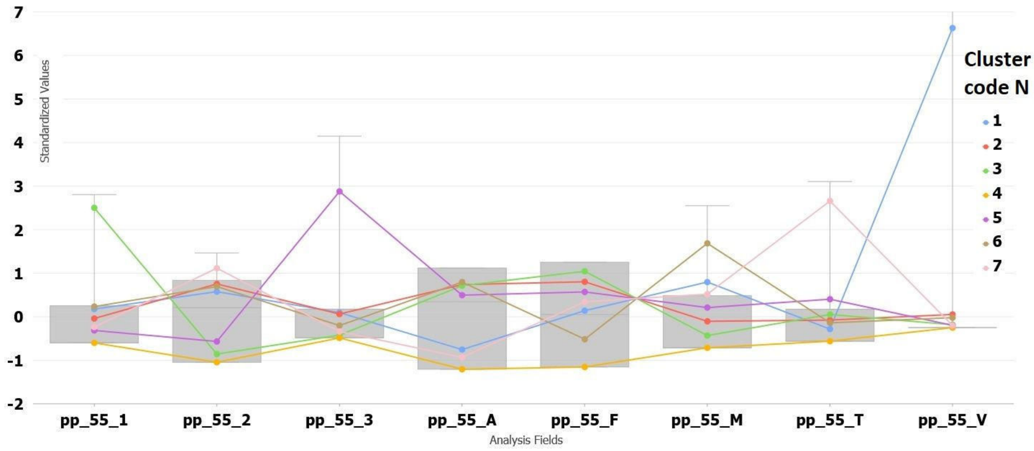


Source: graphic made by authors in Arc Gis Pro

The results of MCA are of 3 types: box plots, features per cluster chart and the MCA cluster map. All these results are linked to each other, each cluster identified is differentiated by a unique colour. In our case, MCA run on census tract to create 7 clusters.

The Box-Plots (Image 5) show information about "both the characteristics of each cluster as well as characteristics of each variable used in the analysis". Our cluster of interest is Cluster number 5 (magenta colour). (Image 7) The cluster with pp_55_3 is the cluster with the highest proportion of employees with tertiary education. In opposition, Cluster 3 reflect the tract of employees with the low level of education. Cluster number 2 includes the locations with the highest share of adult women with a medium level of education from the same sector. Cluster number 1 offer the image for tracts of the place with the highest percentage of aged employees.

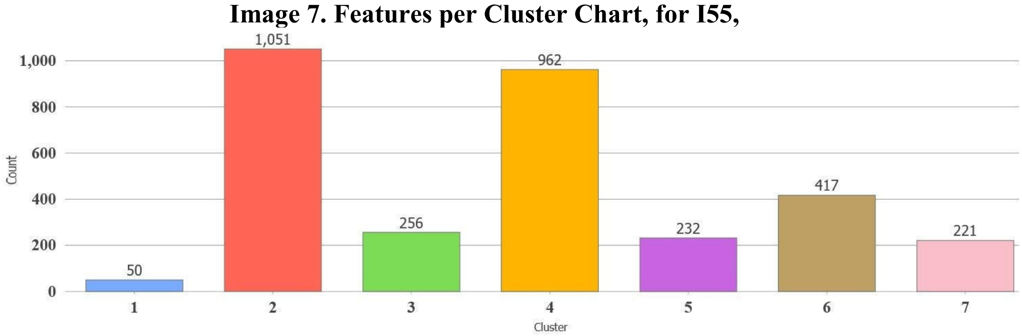
Image 6. Multivariate clustering Box-Plots, for I55, N=7



Source: graphic made by authors in Arc Gis Pro

High concentration of I55 employees with tertiary education are in 232 features (NUTS5). (Image3) and for the rural tourism Cluster 2 there are included 1051 locations.

Image 7. Features per Cluster Chart, for I55,



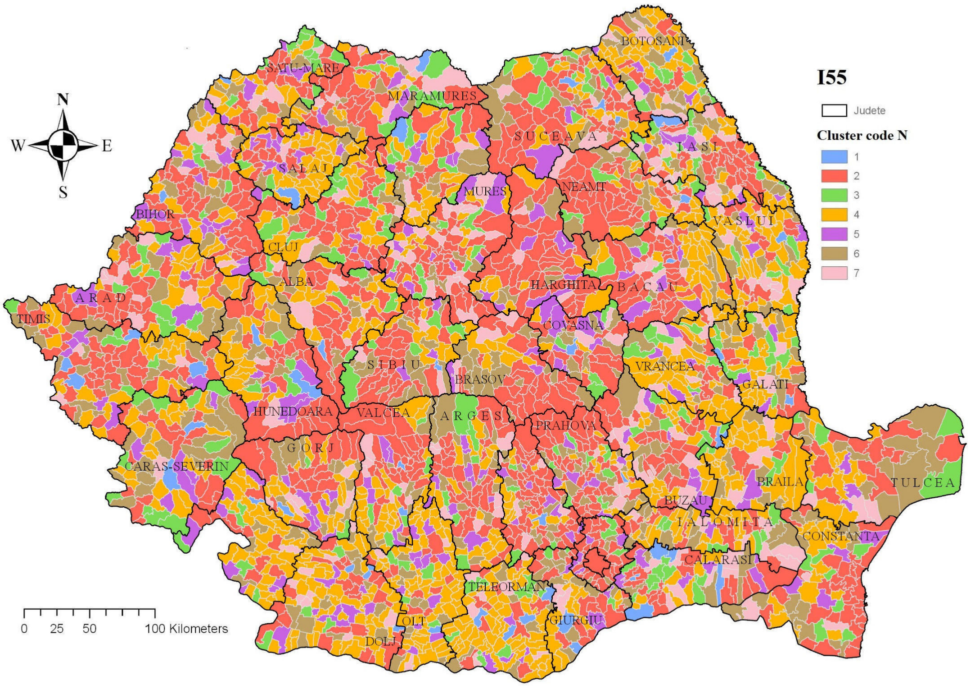
N=7

Source: graphic made by authors in Arc Gis Pro

Features with high tertiary employment in H51 are visible in Image 8 in Cluster 5 (magenta), the *indication of high-intensity cognitive activities*. It is evident the spatial autocorrelation and concentration around some touristic balnear areas: Covasna, Tușnad, Sovata, Vatra Dornei, Herculane. This person present high cognitive activities in balneary treatment – sophisticated treatments, indicating **Marshallian sources externalities**. The

Cluster 2 (red) is more spatially dispersed, especially in main touristic destinations mountain stations, seaside locations and Prahova Valley, mainly in rural areas. Bran, Valea Prahovei, etc. are agrotourism areas; employees are mostly adult women with a medium level of education. Also, cluster 2 indicates **Marshallian source externalities**. For this sector, locations from Cluster 4 do not provide activities, regardless of the eight characteristics, there is no employment visible.

Image 8. Multivariate Clustering Chart outputs for I55, N=7 by shares in total for the level of education (1-Low, 2- Medium and 3- high), age (T- youth, A – adults and V – aged) and gender (M - males, F- Females)



Source: RPL 2011 NUTs 5 data, NIS source. Map made by authors in Arc GIS Pro with ESRI Romania shapefiles,

4.3. Clusters profiles for N79 Activities of tourist agencies and tour operators; other reservation services and tourist assistance

The N79 sector employees mainly females adults with a secondary level of education. The average share of employees with tertiary level of education accounts 13% mean at NUTS 5 level. (Table 5) This sector (see 1.2.1.) is a KIA and an LKIMS sector, have a high-intensity information use, **based on ICT and digitisation adoption is increasing its efficiency and quality of services, services provided in highly competitive markets.**

Table 5. Statistics for employment in N79 at NUTS5 level, by shares in total for the level of education (1-Low, 2- Medium and 3- high), age (T- youth, A – adults and V – aged) and gender (M- males, F- Females) and R^2

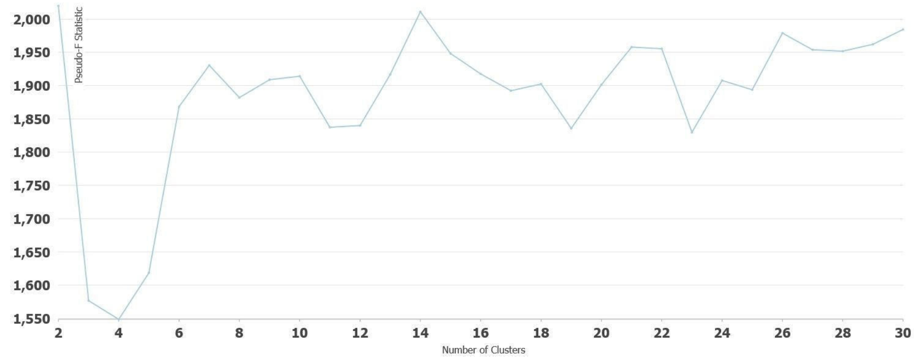
Variable	Mean	Std. Dev.	Min	Max	R^2 N=2	R^2 N=7
PP_79_A	26.150705	41.50857	0.00	100.00	0.796063	0.941548
PP_79_F	18.442486	34.195265	0.00	100.00	0.583396	0.746154
PP_79_2	16.747211	32.913021	0.00	100.00	0.519285	0.80747
PP_79_M	14.828132	30.465268	0.00	100.00	1.475137	0.68019
PP_79_3	**13.078031**	**29.374194**	**0.00**	**100.00**	**0.397566**	**0.797129**
PP_79_T	4.989581	18.253100	0.00	100.00	0.149869	0.773412
PP_79_1	3.445376	14.808696	0.00	100.00	0.108567	0.714129
PP_79_V	2.130332	11.129073	0.00	100.00	0.073491	0.797145

Note: N number of clusters
Source: data calculated by authors

The optimal number of cluster is 14. (Image 9) For the policies purpose, this number of clusters is too large. The criteria selection is given by the rule "the highest peak is the largest F-statistic, indicating how many clusters will be most effective at distinguishing the features and variables you specified."

We run the MCA a number times, selecting the results for N 2 and 7. While N=7 is the first maximum in the pseudo F statistic share, distinguishes feature similarities and differences in an acceptable level. F level indicates that this variable divides the employees into clusters most effectively.

Image 9. Optimized Pseudo-F Statistic Chart for N79, N Optimal = 14 clusters

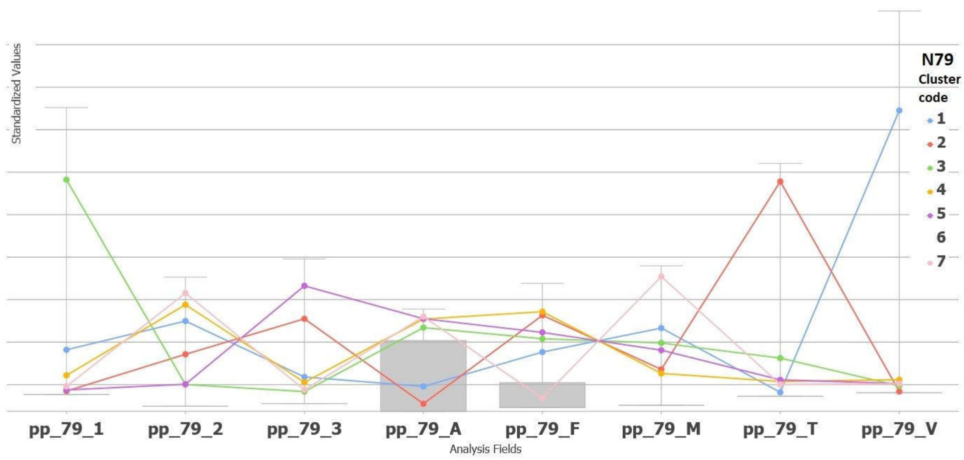


Source: graphic made by authors in Arc Gis Pro

The results of MCA are of 3 types: box plots, features per cluster chart and the MCA cluster map. All these results are linked to each other, each cluster identified is differentiated by a unique colour. In our case, MCA run on census tract to create 7 clusters.

The Box-Plots (Image 10) show information about "both the characteristics of each cluster as well as characteristics of each variable used in the analysis". Our cluster of interest is Cluster number 5 (magenta colour). (Image 10) The cluster with pp_79_3, is the cluster with the highest proportion of employees with tertiary education. In opposition, Cluster 3 reflect the tract of employees with the low level of education, adults and youth, in equal measure males and females. Cluster number 2 includes the locations with the highest share of youth, mainly women with a high level of education from the same sector**.** Cluster number 1 offer the image for tracts of the location with the highest share of aged employees, mainly males with a secondary level of education.

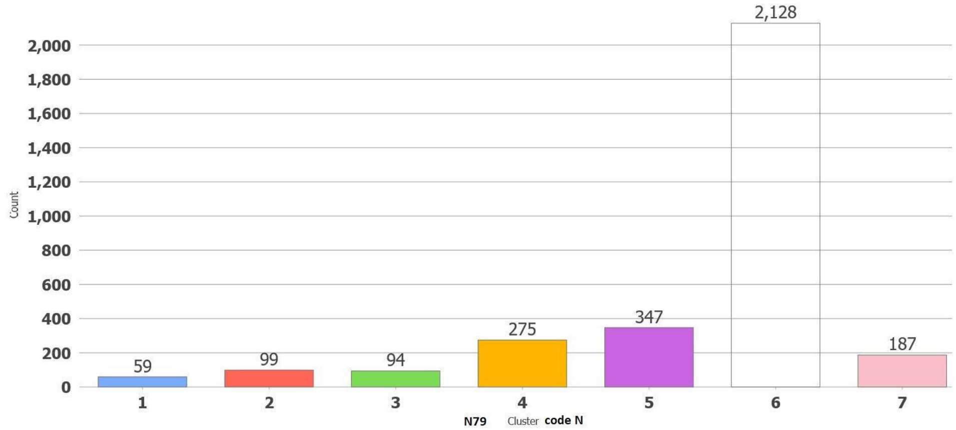
Image 10. Multivariate clustering Box-Plots, for N79, N=6



Source: graphic made by authors in Arc Gis Pro

The high concentration of N79 employees with tertiary education are in Cluster 5 (magenta) 347 features (NUTS5). (Image11) and for the youth tourism Cluster 2 (red) there are included 99 locations. Next to agglomerations of cognitive activities are the Cluster 4 (mustard), which group locations with employees with a medium level of education, mainly adult females, covering 275 features. Cluster 4 is in the neighbourhood of Cluster 5 and 2, reflecting the role of operationalising.

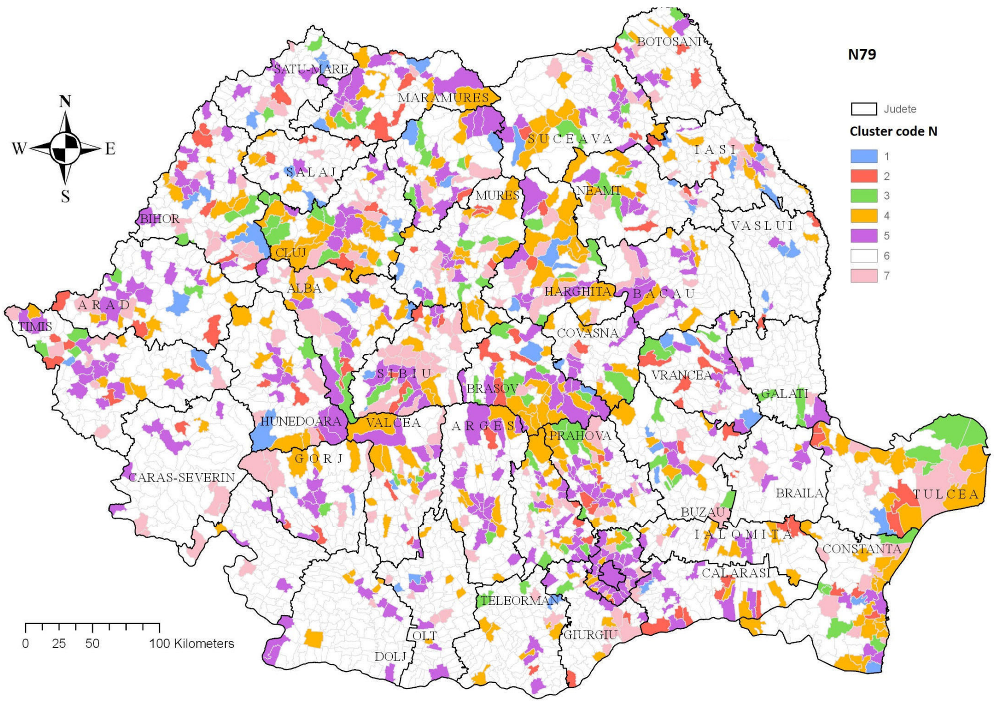
Image 11. Features per Cluster Chart, for N79, N=7



Source: graphic made by authors in Arc Gis Pro

Features with high tertiary employment in N79 are visible in Image 12 in the Cluster 5 (magenta), an *indication of high-intensity cognitive activities*.

Image 12. Multivariate Clustering Chart outputs for N79, N=7 by shares in total for the level of education (1-Low, 2- Medium and 3- high), age (T- youth, A – adults and V – aged) and gender (M- males, F- Females)



Source: RPL 2011 NUTs 5 data, NIS source. Map made by authors in Arc GIS Pro with ESRI Romania shapefiles

It is visible the spatial autocorrelation and location in the large urban agglomerations. Bucharest, Brașov are locations indicating **Urban sources externalities**. Maramures and Hunedoara locations from Cluster 5 indicating Marshallian sources externalities. Cluster 2 (red) but visible on the map, showing activities with the highest degree of digitisation. For this sector, locations from Cluster 6 do not provide activities, regardless of the eight characteristics, there is no employment visible. We emphasise the case of Cluster 3 (green), cluster with employees with a low level of education. It is clear the case of location from the North Tulcea, area of Delta Danube, a highly isolated area on the water. Here a lot o people leave from catching fishing and even rowing.

5. Discussion

Pierre (2014) cite ESRI regarding the fact that MCA is a grouping analysis tool that "perform a classification procedure that tries to **find natural clusters in your data**". This visual detailed and spatial integrated map offer a new perspective to understand the data, i.e. hidden patterns better.

Locations with talent presence could become innovation capabilities investment priorities. (Wei, Feng, and Zhang 2017) examine the impact of innovation capability on the distribution of innovation talents. Authors conclude that for China during 2001-2015: "(a) for areas with low levels of talent, the innovation environment is the most crucial factor; (b) for areas with a medium level of talent, the effects of innovation input and efficiency are moderate; and (c) for areas with a high level of talent, the positive effects of innovation input and efficiency are quite significant".

(Turner, 2018) points out that "Travel & Tourism creates jobs, drives exports, and generates prosperity across the globalised world". Regions are involved in the globalization process to a different extent depending on their structure and specialization. (Capello and Fratesi, 2011, p.2)

(Grigorescu et al. 2019) make a regional mapping of knowledge-intensive job growth from tourism industry using shapeshift analysis. (Lincaru, Ciucă, and Atanasiu 2019) presents the first draft of this article in the recent 2019 EcoSmart.

*

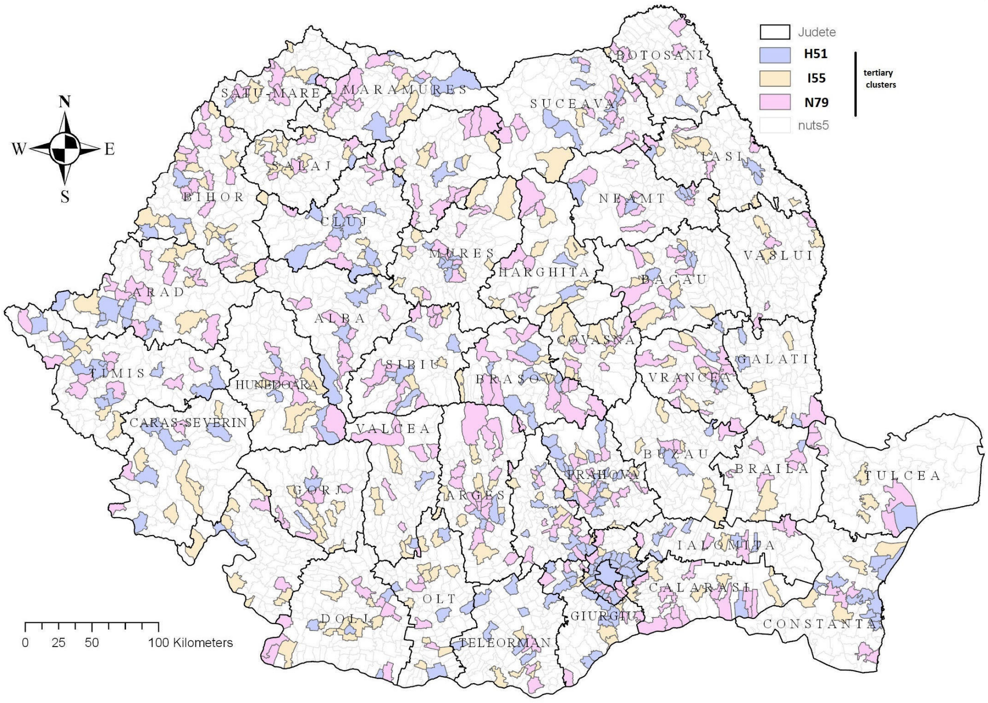
* *

It is remarkable the Bucharest, Brașov and Constanta the neighbouring presence of all analysed sector H51, I55 and N79. (Image 13) These are big cities respond to (Beaudry and Schiffauerova 2009) criterias. These places are sources of both **urbanization and Jacobian externalities. Here exists diversity based mechanisms that support the economic growth of the region.** For these locations are viable tourism ecosystem development. The highest variety allows radical innovation adoption at a rapid pace.

Another case is the Marshallian externalities case for all analysed sectors in Valea Prahovei area and Hunedoara. In Marshallian specialised location is probable the productivity-increasing coupled with incremental innovation adoption.

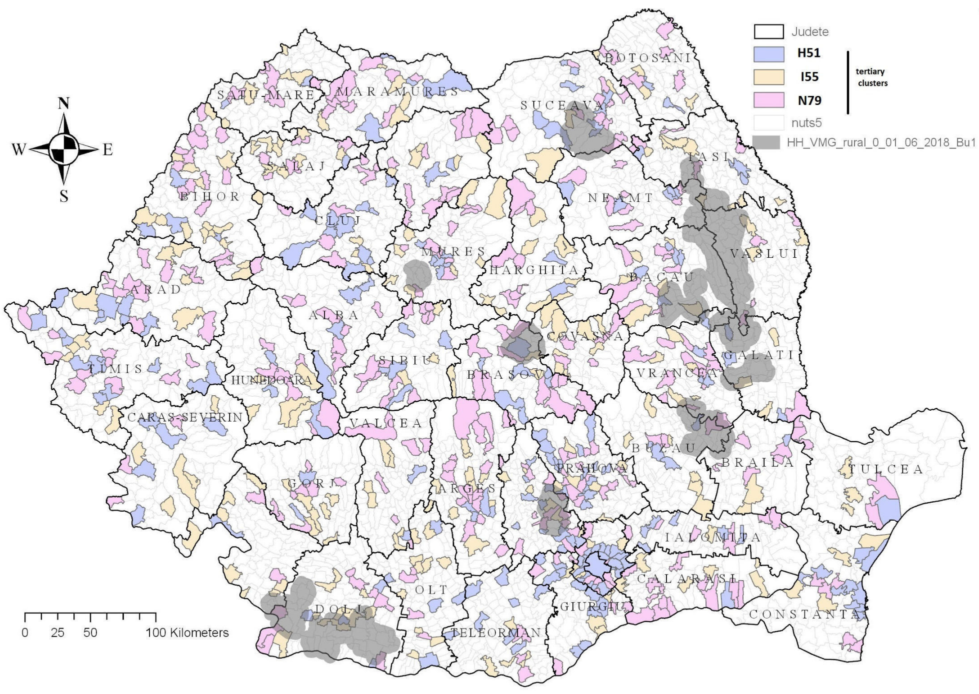
The map overlay initiated by McHarg (1971) "is a procedure for combining the attributes of intersecting features that are represented in two or more geo-registered data layers" DiBiase and Dutton (2009). In Romania, some of the most vulnerable are the Minumum Guarantee Beneficiaries (MGB). If we overlay the Image 13 on the MGB Marginalized Communities from Fl ood Risk Areas published by Lincaru et al. (2020) we obtain the Image 14. We have to emphasise that MGB persons are among the most vulnerable person in Romania, for which there are no sufficiently inclusive jobs! Perles-Ribes et al. (2017) put in discussion the question "Is the tourism-led growth hypothesis valid?", we extrapolate for the case of low educated from underdeveloped areas, as it is the extreme case fo MGB's from historical of 500 years flood risk areas. The response, issued indirectly, is NO, in Romania's poored regions tourism is not inclusive! - the message of Image 14. Result the sad conclusion that location with cumulative disasters as the impact of repetitive crises are not able to attract talents. If tourism, as the most inclusive sector for low educated fails to create inclusive jobs, then, remains still the question: what economic sector could employ persons from historically disadvantaged and vulnerable locations?

Image 13. Human Resources Capabilities inputs for Regions Smart Specialisation Strategies mainly tourism HR with high level of education (3- High) , age(T- youth, A- adults) and gender (mainly F – Females)



Source: made by authors, shape file ESRI RO, RPL 2011 NUTS 5 data, NIS source

Image 14. Human Resources Capabilities inputs for Regions Smart Specialisation Strategies mainly tourism HR with high level of education (3- High), age(T- youth, A- adults) and gender (mainly F – Females) and the VMG agglomeration beneficiaries (June, 2018. ANPIS)



Source: made by authors, shape file ESRI RO, RPL 2011 NUTS 5 data, NIS source; VMG agglomeration beneficiaries (June, 2018. ANPIS)

These results provide insights for spatial based policy decision (employment, education, innovation, investments, infrastructure, etc.) further development. Employment in tourism industries, therefore, should fulfil their potential to "create jobs for economically less advantaged socio-demographic groups or regions" (Eurostat, 2016) as an inclusive growth engine, working for all in a global knowledge economy. Another trend that increases the tourism inclusiveness dimension is the cultural heritage sites, a requisite for a sustainable tourism in a cultural context. (Huete-Alcocer, López-Ruiz, and Grigorescu 2019) Not at least, tourism is a possible link between public and private affairs. (Grigorescu 2006).

6. Conclusions

The first contribution of this work is the creation of spatial patterns of clusters for employment in the main tourism sectors, individually and superposed, at local level granularity – the lowest level of granularity. These maps exploit the recent geo-referenced information related to tourism and other areas. The hidden patters for tourism jobs are sad. Only wealthy locations can develop further sustainable tourism industry, while poor areas, i.e. the poorest ones (MGB Marginalized Communities from Flood Risk Areas) have no chance! The main conclusion of this article is that only the developed regions near highly urban agglomerations have innovation capacity to adopt radical innovation, including to exploit the digital transformation according to global megatrends. Even if the landscape is an attraction, locations (i.e. the MGM location case) these communities fail to develop tourism activities and to create inclusive jobs. This result is in line with the conclusion of (Romão, Guerreiro, and Rodrigues 2017). The results represent a useful input for building a Smart Strategy of the gradual recovery of the sectors following corona Pandemic Impact. To accelerate the digital transformation of the industry according to the demands identified by the tourism Megatrends iterated by (Bloomberg Media Group 2019) is a top priority. This Smart Strategy has to exploit sustainably all the specific territorial resources, especially its talents and human capital, regardless its level of education. In contrast "innovation talents make significant contributions to the improvement of innovation capability, but not vice versa" (Wei, Feng, and Zhang 2017). Authors insist that "attracting and nourishing talents it is the first priority for any region which want to develop smart and sustainable. An inclusive and cohesive territorial development, if exploits the digital transformation opportunities should create the Morettian externalities of high human capital, and finally should close the "diverging gap between ürban and rural"(Lopez-Ruiz et al. 2014).

7. Acknowledgments

This work was supported by a grant from the Romanian Ministry of Research and Innovation- Ministerul Cercetarii si Inovarii din Romania, The Government of Romania, the National Research and Development Plan – Programme NUCLEU, 2019-2022, InovSoc program of the INCSMPS – the National Labor Research Institute of Romania, project no: Perspective funcționale a piețelor locale ale muncii în România, în contextul economiei inteligente și innovative / [Functional perspectives of local labor markets in Romania, in the context of smart and innovative economy] PN 19130101, coordinator Dr. Speranța Pîrciog. An for the Project PN 19130201, from the same Programe: "Analysis of the social risks associated with poverty, with an emphasis on poverty at work /[Analiza riscurilor sociale asociate saraciei, cu accent asupra saraciei in munca]", coordinator Cristina Stroe.

References

Araújo-Vila, Noelia, Jose Antonio Fraiz-Brea, and Arthur Filipe de Araújo. 2020. 'Spanish Economic-Financial Crisis: Social and Academic Interest'. Journal of Business Cycle Research, April. https://doi.org/10.1007/s41549-020-00045-z.

Aydin, Orhun. 2019. 'GIS & Spatial Machine Learning: Transforming Our Planet's Pulse to Action'. ESRI, 73. https://www.lib.uidaho.edu/media/gisday/Aydin_GISDay_Keynote2019.pdf.

Bănilă, Nicoleta. 2018. 'Tourism's Direct Contribution to Romania's GDP to Rise 4.6% in 2018 - WTTC'. See News - Business Intelligence for Southeast Europe. 2018. http://seenews.com/news/tourisms-direct-contribution-to-romanias-gdp-to-rise-46-in-2018-wttc-607733.

Beaudry, Catherine, and Andrea Schiffauerova. 2009. 'Who's Right, Marshall or Jacobs? The Localization versus Urbanization Debate'. Research Policy 38 (2): 318–37. https://doi.org/10.1016/j.respol.2008.11.010.
Bena, Antonella, Massimiliano Giraudo, Roberto Leombruni, and Giuseppe Costa. 2013. 'Job Tenure and Work Injuries: A Multivariate Analysis of the Relation with Previous Experience and Differences by Age'. BMC Public Health 13 (1): 869. https://doi.org/10.1186/1471-2458-13-869.
Bennett, Lauren, Flora Vale, and Jenora d'Acosta. 2015. 'Spatial Statistics: Simple Ways to Do More with Your Data'. Esri User Conference 2015 - Technical Workshop. https://www.esri.com/training/catalog/5765fb3551de57f10993106e/spatial-statistics%3A-simple-ways-to-do-more-with-your-data/.
Bloomberg Media Group. 2019. 'WORLD, TRANSFORMED MEGATRENDS AND THEIR IMPLICATIONS FOR TRAVEL & TOURISM'. World Travel Tourism Council. file:///C:/Users/Cristina/Downloads/Megatrends%202019%20-%20World%20Transformed.pd.
Bogart, William Thomas. 1997. Economics of Cities and Suburbs, The - William Thomas Bogart - 9780135699713 - Economics - Urban and Regional Economics (119). 1st ed. Prentice Hall. Pearson. https://www.pearson.ch/HigherEducation/Pearson/EAN/9780135699713/Economics-of-Cities-and-Suburbs-The.
Bradley, Kacey. 2019. 'Top 5 Trends of IoT in Hospitality : 4Hoteliers'. 4Hoteliers.Com. 2019. https://www.4hoteliers.com/features/article/13069.
Bryson, John, and Peter Daniels. 2007. 'The Handbook of Service Industries'. 2007. https://www.e-elgar.com/shop/the-handbook-of-service-industries.
Cameron, Sperance. 2017. 'Trump Immigration Crackdown Could Take Much-Needed Labor From Booming Hotels'. Bisnow. 2017. https://www.bisnow.com/national/news/hotel/trump-fiscal-plans-give-hotels-business-his-immigration-policies-take-their-workers-82979.
DiBiase, David, and John A Dutton. 2009. '5. Map Overlay Concept | The Nature of Geographic Information'. PennState College of Earth and Mineral Sciences, Department of Geography. 2009. https://www.e-education.psu.edu/natureofgeoinfo/c9_p6.html.
Eurostat. 2014. 'Methodological Manual for Tourism Statistics v3 1 20141219', 250. https://ec.europa.eu/eurostat/documents/3859598/6454997/KS-GQ-14-013-EN-N.pdf/166605aa-c990-40c4-b9f7-59c297154277.
———. 2020. 'Tourism Industries -Economic Analysis. Statistics Explained', April. https://ec.europa.eu/eurostat/statistics-explained/pdfscache/39768.pdf.
Feser, Edward J. 2002. 'Tracing the Sources of Local External Economies'. Urban Studies 39 (13): 2485–2506. https://doi.org/10.1080/0042098022000027077.
Frenken, Koen, Frank Van Oort, and Thijs Verburg. 2007. 'Related Variety, Unrelated Variety and Regional Economic Growth'. Regional Studies 41 (5): 685–97. https://doi.org/10.1080/00343400601120296.
Frocrain, Philippe, and Pierre-Noël Giraud. 2019. 'The Evolution of Tradable and Non Tradable Employment: Evidence from France'. Economie et Statistique / Economics and Statistics, no. 503d (January): 87–107. https://doi.org/10.24187/ecostat.2018.503d.1959.
Fujita, Masahisa, Paul R. Krugman, and Anthony Venables. 1999. The Spatial Economy: Cities, Regions and International Trade. Cambridge, Mass: MIT Press.
Fujita, Masahisa, and Jacques-François Thisse. 1996. 'Economics of Agglomeration'. Journal of the Japanese and International Economies 10 (4): 339–78. https://doi.org/10.1006/jjie.1996.0021.
———. 2013. Economics Agglomeration Cities Industrial Location and Globalization. 2nd ed. Cambridge University Press. https://www.cambridge.org/ro/academic/subjects/economics/microeconomics/economics-agglomeration-cities-industrial-location-and-globalization-2nd-edition, https://www.cambridge.org/ro/academic/subjects/economics/microeconomics.
Grigorescu, Adriana. 2006. 'Marketing of Public and Private Affairs - a Link'. In , KYBERNETES Volume:1179–89. Maribor, Slovenia.
Harrison, Bennett, Maryellen R. Kelley, and Jon Gant. 1996. 'Innovative Firm Behavior and Local Milieu: Exploring the Intersection of Agglomeration, Firm Effects, and Technological Change'. Economic Geography 72 (3): 233. https://doi.org/10.2307/144400.
Henderson, J.Vernon. 2003. 'Marshall's Scale Economies'. Journal of Urban Economics 53 (1): 1–28. https://doi.org/10.1016/S0094-1190(02)00505-3.
Huete-Alcocer, Nuria, Víctor Raúl López-Ruiz, and Adriana Grigorescu. 2019. 'Measurement of Satisfaction in Sustainable Tourism: A Cultural Heritage Site in Spain'. Sustainability 11 (23): 6774. https://doi.org/10.3390/su11236774.
IRTS. 2008. 'IRTS 2008 (International Recommendations, IRTS, Tourism Satellite Account, Tourism Statistics, UNWTO)'. 2008. https://unstats.un.org/unsd/tradekb/Knowledgebase/50551/IRTS-2008.
Jacobs, Jane. 1961. The Death and Life of Great American Cities. New York: Vintage Books A Division of Random House.

———. 1969. 'The City. The Economy of the Cities.' National Civic Review 58 (9): 447–48. https://doi.org/10.1002/ncr.4100580916.

Jain, Anil K. 2010. 'Data Clustering: 50 Years beyond K-Means'. Pattern Recognition Letters 31 (8): 651–66. https://doi.org/10.1016/j.patrec.2009.09.011.

Krugman, Paul. 1991. 'Increasing Returns and Economic Geography'. Journal of Political Economy 99 (3): 483–99. https://doi.org/10.1086/261763.

Langford, Guy, Adam Weissenberg, and Marcello Gasdia. 2019. '2019 US Travel and Hospitality Outlook'. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consumer-business/us-consumer-2019-us-travel-and-hospitality-outlook.pdf.

Lim, Jaewon, and DooHwan Won. 2020. 'How Las Vegas' Tourism Could Survive an Economic Crisis?' Cities 100 (May): 102643. https://doi.org/10.1016/j.cities.2020.102643.

Lincaru, Cristina, Speranţa Pîrciog, Adriana Grigorescu, and Cristina Stroe. 2020. 'Chapter 4. The Flood Risk Information Used in Spatial Policy-Based Design for Socio-Economic Integration of the Most Vulnerable Persons from Rural Agglomerations. Study Case Romania'. In Public Administration: Perspectives, Management and Challenges, Frederik S. Møller (Editor). Political Science and History POL028000. Niva Science Publisher. https://novapublishers.com/shop/public-administration-perspectives-management-and-challenges/.

Llano, Carlos. 2000. 'ECONOMIC GEOGRAPHY: THE CORE-PERIPHERY MODEL'. References for the slides: Fujita, Krugman y Venables: presented at the Lecture 6. https://slideplayer.com/slide/4346625/.

Lopez-Ruiz, Victor Raoul, Domingo Nevado Pena, Jose Luis Alfaro Navarro, and Adriana Grigorescu. 2014. 'Human Development European City Index: Methodology and Results | Request PDF'. Romanian Journal of Economic Forecasting XVII (3). http://www.ipe.ro/rjef/rjef3_14/rjef3_2014p72-87.pdf.

MacQueen, J.B. 1967. 'Some Methods for Classification and Analysis of Multivariate Observations.' Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, , Berkeley, Calif., 1, Styatistics: 281–97.

McHarg, I.L. 1971. Design with Nature. https://books.google.ro/books?id=CovnoQEACAAJ.

Michaelides, Panayotis, George Economakis, and Dimitrios Lagos. 2006. 'Clustering Analysis Methodology for Employment and Regional Planning in Greece', 9. https://pdfs.semanticscholar.org/c596/db07c25ea4d66eee9a5b77194892d8ea3e98.pdf.

Mitchell, Tom M. 1997. Machine Learning. McGraw-Hill Science/Engineering/Math. http://profsite.um.ac.ir/~monsefi/machine-learning/pdf/Machine-Learning-Tom-Mitchell.pdf.

Moretti, Enrico. 2003. 'Human Capital Externalities in Cities'. w9641. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w9641.

———. 2004. 'Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions'. American Economic Review 94 (3): 656–90. https://doi.org/10.1257/0002828041464623.

———. 2012. The New Geography of Jobs. Houghton Mifflin Harcourt.

Muller, Emmanuel, and David Doloreux. 2007. 'The Key Dimensions of Knowledge-Intensive Business Services (KIBS) Analysis: A Decade of Evolution'. U1/2007. Working Papers 'Firms and Region'. Fraunhofer Institute for Systems and Innovation Research (ISI). https://ideas.repec.org/p/zbw/fisifr/u12007.html.

OECD. 2007. 'OECD Science, Technology and Industry Scoreboard 2007 - OECD'. 2007. http://www.oecd.org/science/inno/oecdsciencetechnologyandindustryscoreboard2007.htm.

———. 2017. OECD Science, Technology and Industry Scoreboard 2017: The Digital Transformation. OECD Science, Technology and Industry Scoreboard. OECD. https://doi.org/10.1787/9789264268821-en.

Parsons, Meg, Cilla Brown, Johanna Nalau, and Karen Fisher. 2018. 'Assessing Adaptive Capacity and Adaptation: Insights from Samoan Tourism Operators'. Climate and Development 10 (7): 644–63. https://doi.org/10.1080/17565529.2017.1410082.

Perles-Ribes, José Francisco, Ana Belén Ramón-Rodríguez, Antonio Rubia, and Luis Moreno-Izquierdo. 2017. 'Is the Tourism-Led Growth Hypothesis Valid after the Global Economic and Financial Crisis? The Case of Spain 1957–2014'. Tourism Management 61 (August): 96–109. https://doi.org/10.1016/j.tourman.2017.01.003.

Petrova, Aleksia. 2020. 'Romania, Serbia to Bear Brunt of COVID-19 Induced Value Chain Disruptions - OECD'. See News - Business Intelligence for Southeast Europe. 4 April 2020. http://seenews.com/news/romania-serbia-to-bear-brunt-of-covid-19-induced-value-chain-disruptions-oecd-696326.

Pierre, Daniel. 2014. 'Developing a Tutorial for Grouping Analysis in ArcGIS'. Centre of Geographic Science. https://www.slideshare.net/COGSPresentations/developing-a-tutorial-for-grouping-analysis-in-arcgis.

Porter, Michael E. 1979. 'How Competitive Forces Shape Strategy'. Harvard Business Review, 1 March 1979. https://hbr.org/1979/03/how-competitive-forces-shape-strategy.
Quigley, John M. 1998. 'Urban Diversity and Economic Growth'. Journal of Economic Perspectives 12 (2): 127–38.
Raspe, Otto, and Frank Van Oort. 2006. 'The Knowledge Economy and Urban Economic Growth'. European Planning Studies 14 (9): 1209–34. https://doi.org/10.1080/09654310600933322.
Reichheld, Ashley, Mark Allen, Emiliy Ballbach, and Morgan Seyle. 2020. 'Human Experience: At the Heart of Learning, Living, Connecting, and Business Loyalty The Transportation, Hospitality & Services Perspective'. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/process-and-operations/us-human-experience.pdf.
Romania. OECD Tourism Trends and Policies 2020. 2020. OECD Tourism Trends and Policies. OECD. https://doi.org/10.1787/721999bd-en.
Romão, João, João Guerreiro, and Paulo M. M. Rodrigues. 2017. 'Territory and Sustainable Tourism Development: A Space-Time Analysis on European Regions'. REGION 4 (3): 1. https://doi.org/10.18335/region.v4i3.142.
Ruthartr. 2018. 'New Clustering Tools in ArcGIS Pro 2.1: More Machine Learning at Your Fingertips'. ArcGIS Blog (blog). 22 January 2018. https://www.esri.com/arcgis-blog/products/analytics/analytics/new-clustering-tools-in-arcgis-pro-2-1-more-machine-learning-at-your-fingertips/.
Schwab, Klaus. 2019. 'The Global Competitiveness Report 2019' World Economic Forum: 666. http://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf.
SiteMinder. 2020. 'The Hotel Industry of 2020: Predicting the Future of Hotel Tech Is Difficult : 4Hoteliers'. 4Hoteliers.Com. 2020. https://www.4hoteliers.com/features/article/13096.
Starkov, Max. 2020. 'Do You Have a Coronavirus Strategy? : 4Hoteliers'. 4Hoteliers.Com. 2020. https://www.4hoteliers.com/features/article/13234.
Steinhaus, Hugo. 1957. 'Sur La Division Des Corps Materiels En Parties.' Bull Acad Polon Sci 4 4 (12): 801–4. https://mathscinet.ams.org/mathscinet-getitem?mr=0090073.
Tatarczak, Anna, and Oleksandra Boichuk. 2017. 'The Use of Multivariate Techniques for Youth Unemployment Analysis in Poland'. Working Papers, May. https://ideas.repec.org/p/pes/wpaper/2017no130.html.
———. 2018. 'The Multivariate Techniques in Evaluation of Unemployment Analysis of Polish Regions'. Oeconomia Copernicana 9 (3): 361–80. https://doi.org/10.24136/oc.2018.018.
TSA, ed. 2010. Tourism Satellite Account: Recommended Methodological Framework 2008. Rev. 1. Studies in Methods Series F 80, Rev.1. Luxembourg: Commission of the European Communities, Eurostat [u.a.]. https://unstats.un.org/unsd/publication/Seriesf/SeriesF_80rev1e.pdf.
UNWTO. 2018. European Union Tourism Trends. World Tourism Organization (UNWTO). https://doi.org/10.18111/9789284419470.
Vasilescu, Maria Denisa, Larisa Aparaschivei, and Mihai Roman. 2012. 'EMPLOYMENT IN ROMANIA: EVIDENCE FROM A PANEL DATA ANALYSIS', 15. https://mpra.ub.uni-muenchen.de/38388/1/MPRA_paper_38388.pdf.
Wei, Fenfen, Nanping Feng, and Kevin Zhang. 2017. 'Innovation Capability and Innovation Talents: Evidence from China Based on a Quantile Regression Approach'. Sustainability 9 (7): 1218. https://doi.org/10.3390/su9071218.
Wikipedia. 2020. 'K-Means Clustering'. 2020. https://en.wikipedia.org/wiki/K-means_clustering.
World Tourism Organization (UNWTO). 2018. European Union Tourism Trends. World Tourism Organization (UNWTO). https://doi.org/10.18111/9789284419470.
WTTC. 2020a. '2020 Economic Impact Report on the Economic and Employment Impact of Travel & Tourism'. https://wttc.org/en-us/Research/Economic-Impact.
———. 2020b. 'The Importance of Travel & Tourism in 2019'. World Travel & Tourism Council. https://wttc.org/en-us/Research/Economic-Impact.
———. 2020c. 'TRAVEL & TOURISM ECONOMIC IMPACT FROM COVID-19'. https://wttc.org/en-us/Research/Economic-Impact.

Appendices

Table 6. Romanian tourism industries detailed profile in the context of economy (absolute values)

		Enterprises - number	**Turnover or gross premiums written - million euro**	**Value added at factor cost - million euro**	**Persons employed - number**
INDIC_SB	*var*	*V11110*	*V12110*	*V12150*	*V16110*
TIME/NACE_R2	*Code*	*2017*	*2017*	*2017*	*2017*
Total business economy; repair of computers, personal and household goods; except financial and insurance activities	**(1)**	**485,215**	**292,990.2**	**66,908.8**	**4,020,121**
Total services (2)	**(2)**	**204909**	**52077**	**21465.653**	**1328892**
Total tourism industries (3)	**(3)**	**40,654**	**7,140**	**2,227**	**241,812**
Tourism industries (mainly tourism) (4)	**(4)**	**7,011**	**2,658**	**716**	**58,009**
Tourism industries (mainly tourism) (H51+I55+N79) RIS data		**8,960**	**2,913**	**800**	**65,941**
Tourism industries (partially tourism) (5)	(5)	33,643	4,482	1,512	183,803
Transport related (total)	**(6)**	**43,698**	**12,993**	**3,621**	**273,947**
Land transport and transport via pipelines	H49	43,409	12,044.4	3,620.9	267,183
Passenger rail transport, interurban	*H491*	*22*	*207.6*		*13,782*
Taxi operation	*H4932*	*9,490*	*227.7*	*110.4*	*20,704*
Other passenger land transport n.e.c.	*H4939*	*3,430*	*612.9*	*227.4*	*22,420*
Water transport	H50	211	187.5		2,397
Sea and coastal passenger water transport	*H501*	*25*	*0.9*		*50*
Inland passenger water transport	*H503*	*72*	*8.3*	*6.1*	*369*
Air transport	H51	78	761.4		4,367
Passenger air transport	*H511*	*48*	*719.2*		*4,079*
Accommodation	**I55**	**6,074**	**1,303.7**	**631.6**	**50,610**
Hotels and similar accommodation	*I551*	*2,403*	*1,079.2*	*541.4*	*39,136*
Holiday and other short-stay accommodation	*I552*	*1,995*	*95.8*	*36.3*	*5,338*
Camping grounds, recreational vehicle parks and trailer parks	*I553*	*81*	*6.2*	*2.3*	*331*
Food and beverage service activities	**I56**	**20,339**	**3,082.5**	**944.4**	**131,745**
Restaurants and mobile food service activities	*I561*	*9,903*	*2,208.3*	*652.4*	*92,271*
Beverage serving activities	*I563*	*9,158*	*585.7*	*176.7*	*28,474*
Rental and leasing activities	**N77**	**2,465**	**865.7**	**473.3**	**8,500**
Renting and leasing of motor vehicles	*N771*	*990*	*530.7*	*299.8*	*3,476*
Renting and leasing of recreational and sports goods	*N7721*	*229*	*9.4*	*5.9*	*418*
Travel agency, tour operator and other reservation service and related activities	**N79**	**2,808**	**847.8**	**168.4**	**10,964**
Travel agency and tour operator activities	*N791*	*2,484*	*757.2*	*135.5*	*9,125*
Other reservation service and related activities	*N799*	*324*	*90.6*	*32.9*	*1,839*

Note: Due to unreliable data at country level and rounding, deviations can occur between total and subtotal
(1) B-N_S95_X_K; (2) NACE sections H, I, J, L, M, N, S95; (3) NACE classes H491, H4932, H4939, H501, H503, H511, I551, I552, I553, I561, I563, N771, N7721 and division N79; (4) NACE classes H511, I551, I552, I553 and N791; (5) NACE classes H491, H4932, H4939, H501, H503, I561, I563, N771, N7721 and N799; (6) H49+H50+H51; Source: Eurostat (online data code: sbs_na_sca_r2, sbs_na_1a_se_r2)

Table 7. Romanian tourism industries detailed profile in the context of economy (relative values)

		Enterprises - number	Turnover or gross premiums written - million euro	Value added at factor cost - million euro	Persons employed - number
INDIC_SB	Code/	*V11110*	*V12110*	*V12150*	*V16110*
TIME / NACE_R2	*Year*	*2017*	*2017*	*2017*	*2017*
Total tourism industries as share of total non-financial business economy(1)	**(2)**	**8.4%**	**2.4%**	**3.3%**	**6.0%**
Total tourism industries as share of total Services(2)	**(3)**	**19.8%**	**13.7%**	**10.4%**	**18.2%**
Total tourism industries(3), of which:		*100%*	*100%*	*100%*	*100%*
Tourism industries (mainly tourism) (4)	**(4)**	17%	37%	32%	24%
Tourism industries (partially tourism) (5)	(5)	83%	63%	68%	76%
Total tourism industries, of which:		100%	100%	100%	100%
Transport related (total)	**(6)**	107.5%	182.0%	162.6%	113.3%
Land transport and transport via pipelines	**H49**	**106.8%**	**168.7%**	**162.6%**	**110.5%**
Passenger rail transport, interurban	*H491*	0.1%	2.9%	0.0%	5.7%
Taxi operation	*H4932*	23.3%	3.2%	5.0%	8.6%
Other passenger land transport n.e.c.	*H4939*	8.4%	8.6%	10.2%	9.3%
Water transport	**H50**	**0.5%**	**2.6%**	**0.0%**	**1.0%**
Sea and coastal passenger water transport	*H501*	0.1%	0.0%	0.0%	0.0%
Inland passenger water transport	*H503*	0.2%	0.1%	0.3%	0.2%
Air transport	**H51**	**0.2%**	**10.7%**	**0.0%**	**1.8%**
Passenger air transport	*H511*	0.1%	10.1%	0.0%	1.7%
Accommodation	**I55**	**14.9%**	**18.3%**	**28.4%**	**20.9%**
Hotels and similar accommodation	*I551*	5.9%	15.1%	24.3%	16.2%
Holiday and other short-stay accommodation	*I552*	4.9%	1.3%	1.6%	2.2%
Camping grounds, recreational vehicle parks and trailer parks	*I553*	0.2%	0.1%	0.1%	0.1%
Food and beverage service activities	**I56**	**50.0%**	**43.2%**	**42.4%**	**54.5%**
Restaurants and mobile food service activities	*I561*	24.4%	30.9%	29.3%	38.2%
Beverage serving activities	*I563*	22.5%	8.2%	7.9%	11.8%
Rental and leasing activities	**N77**	**6.1%**	**12.1%**	**21.3%**	**3.5%**
Renting and leasing of motor vehicles	*N771*	2.4%	7.4%	13.5%	1.4%
Renting and leasing of recreational and sports goods	*N7721*	0.6%	0.1%	0.3%	0.2%
Travel agency, tour operator and other reservation service and related activities	**N79**	**6.9%**	**11.9%**	**7.6%**	**4.5%**
Travel agency and tour operator activities	*N791*	6.1%	10.6%	6.1%	3.8%
Other reservation service and related activities	*N799*	0.8%	1.3%	1.5%	0.8%

Note: Due to unreliable data at country level and rounding, deviations can occur between total and subtotal
(1) B-N_S95_X_K; (2) NACE sections H, I, J, L, M, N, S95; (3) NACE classes H491, H4932, H4939, H501, H503, H511, I551, I552, I553, I561, I563, N771, N7721 and division N79; (4) NACE classes H511, I551, I552, I553 and N791; (5) NACE classes H491, H4932, H4939, H501, H503, I561, I563, N771, N7721 and N799; (6) H49+H50+H51; Source: Eurostat (online data code: sbs_na_sca_r2, sbs_na_1a_se_r2)