Czech University of Life Sciences Prague

Faculty of Economics and Management

Department of Statistics

Master's Thesis

Convergence of Selected Economic Indicators in Regional Development in the European Union

Karel Tomšík

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Economics and Management

DIPLOMA THESIS ASSIGNMENT

Bc. Karel Tomšík

Economics and Management
Economics and Management

Thesis title

Convergence of selected economic indicators in regional development in the European Union

---

Objectives of thesis

The goal of the diploma thesis is to evaluate convergence of selected economic indicators at the regional level in the European Union. Specifically, the thesis aims to compare regional development in time using relevant statistical methods.

Methodology

The diploma thesis is divided into two main parts, theoretical and practical one. The theoretical part will consist of the description of the EU regional policy and the system of data collection from the EU regions. Furthermore, relevant statistical methods such as exploratory data analysis and methods measuring convergence will be presented. The practical part will be composed on an analysis of collected regional data, their synthesis and drawing relevant conclusions.

The proposed extent of the thesis

60 – 80 pages

Keywords

Convergence, European Union, regional development, statistical analysis

Recommended information sources

ANDERSON, T W. *An introduction to multivariate statistical analysis.* Hoboken, N.J.: Wiley-Interscience, 2003. ISBN 0471360910.

BALDWIN, R E. – WYPLOSZ, C. *The economics of European integration.* London: McGraw-Hill Higher Education, 2012. ISBN 978-0-07-713172-2.

CAPELLO, R. Regional Economics. Abingdon: Routledge, 2016. ISBN 978-1-138-85588-5

HINDLS, R. – HRONOVÁ, S. – SEGER, J. *Statistika pro ekonomy.* Praha: Professional publishing, 2006. ISBN 80-86419-99-1.

KÁBA, B. – SVATOŠOVÁ, L. *Statistické nástroje ekonomického výzkumu.* Plzeň: Vydavatelství a nakladatelství Aleš Čeněk, 2012. ISBN 978-80-7380-359-9.

LARSON, R. – FARBER, E. *Elementary statistics : picturing the world.* Boston: Pearson Prentice Hall, 2015. ISBN 9780321693624.

OTT, L. – LONGNECKER, M. *An introduction to statistical methods & data analysis.* Australia: Cengage Learning, 2016. ISBN 9781305269477.

Expected date of thesis defence

2021/22 SS – FEM

The Diploma Thesis Supervisor

Ing. Tomáš Hlavsa, Ph.D.

Supervising department

Department of Statistics

Electronic approval: 8. 9. 2021

prof. Ing. Libuše Svatošová, CSc.

Head of department

Electronic approval: 19. 10. 2021

Ing. Martin Pelikán, Ph.D.

Dean

Prague on 31. 03. 2022

Declaration

I declare that I have worked on my master's thesis titled 'Convergence of Selected Economic Indicators in Regional Development in the European Union' by myself and I have used only the sources mentioned at the end of the thesis. As the author of the master's thesis, I declare that the thesis does not break copyrights of any their person.

In Prague on 31 March 2022   _____

Convergence of Selected Economic Indicators in Regional Development in the European Union

Abstract

The master's thesis deals with the topic of convergence of selected economic indicators in regional development in the European Union. The aim of the thesis is to evaluate the economic convergence among the regions at the NUTS 2 level in the period 2015-2019 and the subsequent analysis of selected indicators influencing this process. The master's thesis contains the theoretical basis of regional development and regional policy with a focus on the regional policy of the European Union. In the practical part, the thesis uses statistical methods to analyse convergence between regions in the European Union. Specifically, it is beta-convergence and sigma-convergence to calculate economic convergence and cluster analysis to identify trends affecting certain groups of regions. The synthesis of the results points to the fact that economic convergence among the regions occurs, but with significant differences among the groups of regions defined in the cluster analysis.

Keywords: convergence, European Union, regional development, statistical analysis

Konvergence vybraných ekonomických indikátorů v regionálním rozvoji v Evropské unii

Abstrakt

Diplomová práce se zabývá tématem konvergence vybraných ekonomických indikátorů v regionálním rozvoji v Evropské unii. Cílem práce je zhodnocení ekonomické konvergence mezi regiony na úrovni NUTS 2 v období 2015-2019 a následná analýza vybraných indikátorů ovlivňujících tento proces. Diplomová práce obsahuje teoretická východiska regionálního rozvoje a regionální politiky se zaměřením na regionální politiku Evropské unie. V praktické části práce využívá statistických metod k analýze konvergence mezi regiony v Evropské unii. Konkrétně se jedná o beta-konvergenci a sigma-konvergenci pro výpočet ekonomické konvergence a shlukovou analýzu pro zjištění trendů, které ovlivňují určité skupiny regionů. Syntéza výsledků poukazuje na skutečnost, že k ekonomické konvergenci mezi regiony dochází, avšak s výraznými rozdíly mezi skupinami regionů definovanými ve shlukové analýze.

Klíčová slova: konvergence, Evropská Unie, regionální rozvoj, statistická analýza

Contents

List of Images

List of Figures

List of Tables

List of Abbreviations

CAP	Common Agricultural Policy
CF	Cohesion Fund
CFP	Common Fisheries Policy
CVCE	Virtual Centre for Knowledge on Europe
CZSO	Czech Statistical Office
EAFRD	European Agricultural Fund for Rural Development
ECU	European Currency Unit
EDA	Exploratory Data Analysis
EMFF	European Maritime and Fisheries Fund
EPRS	European Parliamentary Research Service
ERDF	European Regional Development Fund
ESF	European Social Fund
ESI Funds	European Structural and Investment Funds
EU	European Union
EUR	Euro (currency of the European Union)
GDP	Gross Domestic Product
GERD	Gross Domestic Expenditure on Research and Development
IQR	Interquartile Range
ISCED	International Standard Classification of Education
JTF	Just Transition Fund
LAU	Local Administrative Units
MFF	Multiannual Financial Framework
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organisation for Economic Co-operation and Development
SAE	Single European Act
WWI	First World War
WWII	Second World War

Codes of Territorial Units

BE	Belgium
BE10	Région de Bruxelles-Capitale/Brussels Hoofdstedelijk Gewest
BE21	Prov. Antwerpen
BE22	Prov. Limburg (BE)
BE23	Prov. Oost-Vlaanderen
BE24	Prov. Vlaams-Brabant
BE25	Prov. West-Vlaanderen
BE31	Prov. Brabant wallon
BE32	Prov. Hainaut
BE33	Prov. Liège
BE34	Prov. Luxembourg (BE)
BE35	Prov. Namur
BG	Bulgaria
BG31	Severozapaden
BG32	Severen tsentralen
BG33	Severoiztochen
BG34	Yugoiztochen
BG41	Yugozapaden
BG42	Yuzhen tsentralen
CZ	Czechia
CZ01	Praha
CZ02	Střední Čechy
CZ03	Jihozápad
CZ04	Severozápad
CZ05	Severovýchod
CZ06	Jihovýchod
CZ07	Střední Morava
CZ08	Moravskoslezsko
DK	Denmark
DK01	Hovedstaden
DK02	Sjælland
DK03	Syddanmark
DK04	Midtjylland
DK05	Nordjylland
DE	Germany
DE11	Stuttgart
DE12	Karlsruhe
DE13	Freiburg
DE14	Tübingen
DE21	Oberbayern
DE22	Niederbayern
DE23	Oberpfalz
DE24	Oberfranken
DE25	Mittelfranken
DE26	Unterfranken
DE27	Schwaben
DE30	Berlin

DE40	Brandenburg
DE50	Bremen
DE60	Hamburg
DE71	Darmstadt
DE72	Gießen
DE73	Kassel
DE80	Mecklenburg-Vorpommern
DE91	Braunschweig
DE92	Hannover
DE93	Lüneburg
DE94	Weser-Ems
DEA1	Düsseldorf
DEA2	Köln
DEA3	Münster
DEA4	Detmold
DEA5	Arnsberg
DEB1	Koblenz
DEB2	Trier
DEB3	Rheinhessen-Pfalz
DEC0	Saarland
DED2	Dresden
DED4	Chemnitz
DED5	Leipzig
DEE0	Sachsen-Anhalt
DEF0	Schleswig-Holstein
DEG0	Thüringen
EE	Estonia
EE00	Eesti
IE	Ireland
IE04	Northern and Western
IE05	Southern
IE06	Eastern and Midland
EL	Greece
EL30	Attiki
EL41	Voreio Aigaio
EL42	Notio Aigaio
EL43	Kriti
EL51	Anatoliki Makedonia, Thraki
EL52	Kentriki Makedonia
EL53	Dytiki Makedonia
EL54	Ipeiros
EL61	Thessalia
EL62	Ionia Nisia
EL63	Dytiki Ellada
EL64	Sterea Ellada
EL65	Peloponnisos
ES	Spain
ES11	Galicia
ES12	Principado de Asturias

ES13	Cantabria
ES21	País Vasco
ES22	Comunidad Foral de Navarra
ES23	La Rioja
ES24	Aragón
ES30	Comunidad de Madrid
ES41	Castilla y León
ES42	Castilla-la Mancha
ES43	Extremadura
ES51	Cataluña
ES52	Comunitat Valenciana
ES53	Illes Balears
ES61	Andalucía
ES62	Región de Murcia
ES63	Ciudad de Ceuta
ES64	Ciudad de Melilla
ES70	Canarias
FR	France
FR10	Île de France
FRB0	Centre - Val de Loire
FRC1	Bourgogne
FRC2	Franche-Comté
FRD1	Basse-Normandie
FRD2	Haute-Normandie
FRE1	Nord-Pas-de-Calais
FRE2	Picardie
FRF1	Alsace
FRF2	Champagne-Ardenne
FRF3	Lorraine
FRG0	Pays-de-la-Loire
FRH0	Bretagne
FRI1	Aquitaine
FRI2	Limousin
FRI3	Poitou-Charentes
FRJ1	Languedoc-Roussillon
FRJ2	Midi-Pyrénées
FRK1	Auvergne
FRK2	Rhône-Alpes
FRL0	Provence-Alpes-Côte d'Azur
FRM0	Corse
FRY1	Guadeloupe
FRY2	Martinique
FRY3	Guyane
FRY4	La Réunion
FRY5	Mayotte
HR	Croatia
HR03	Jadranska Hrvatska
HR04	Kontinentalna Hrvatska
IT	Italy

ITC1	Piemonte
ITC2	Valle d'Aosta/Vallée d'Aoste
ITC3	Liguria
ITC4	Lombardia
ITH1	Provincia Autonoma di Bolzano/Bozen
ITH2	Provincia Autonoma di Trento
ITH3	Veneto
ITH4	Friuli-Venezia Giulia
ITH5	Emilia-Romagna
ITI1	Toscana
ITI2	Umbria
ITI3	Marche
ITI4	Lazio
ITF1	Abruzzo
ITF2	Molise
ITF3	Campania
ITF4	Puglia
ITF5	Basilicata
ITF6	Calabria
ITG1	Sicilia
ITG2	Sardegna
CY	Cyprus
CY00	Kypros
LV	Latvia
LV00	Latvija
LT	Lithuania
LT01	Sostines regionas
LT02	Vidurio ir vakaru Lietuvos regionas
LU	Luxembourg
LU00	Luxembourg
HU	Hungary
HU11	Budapest
HU12	Pest
HU21	Közép-Dunántúl
HU22	Nyugat-Dunántúl
HU23	Dél-Dunántúl
HU31	Észak-Magyarország
HU32	Észak-Alföld
HU33	Dél-Alföld
MT	Malta
MT00	Malta
NL	Netherlands
NL11	Groningen
NL12	Friesland (NL)
NL13	Drenthe
NL21	Overijssel
NL22	Gelderland
NL23	Flevoland
NL31	Utrecht

NL32	Noord-Holland
NL33	Zuid-Holland
NL34	Zeeland
NL41	Noord-Brabant
NL42	Limburg (NL)
AT	Austria
AT11	Burgenland (AT)
AT12	Niederösterreich
AT13	Wien
AT21	Kärnten
AT22	Steiermark
AT31	Oberösterreich
AT32	Salzburg
AT33	Tirol
AT34	Vorarlberg
PL	Poland
PL21	Malopolskie
PL22	Slaskie
PL41	Wielkopolskie
PL42	Zachodniopomorskie
PL43	Lubuskie
PL51	Dolnoslaskie
PL52	Opolskie
PL61	Kujawsko-Pomorskie
PL62	Warminsko-Mazurskie
PL63	Pomorskie
PL71	Lódzkie
PL72	Swietokrzyskie
PL81	Lubelskie
PL82	Podkarpackie
PL84	Podlaskie
PL91	Warszawski stoleczny
PL92	Mazowiecki regionalny
PT	Portugal
PT11	Norte
PT15	Algarve
PT16	Centro (PT)
PT17	Área Metropolitana de Lisboa
PT18	Alentejo
PT20	Região Autónoma dos Açores (PT)
PT30	Região Autónoma da Madeira (PT)
RO	Romania
RO11	Nord-Vest
RO12	Centru
RO21	Nord-Est
RO22	Sud-Est
RO31	Sud - Muntenia
RO32	Bucuresti - Ilfov
RO41	Sud-Vest Oltenia

RO42	Vest
SI	Slovenia
SI03	Vzhodna Slovenija
SI04	Zahodna Slovenija
SK	Slovakia
SK01	Bratislavský kraj
SK02	Západné Slovensko
SK03	Stredné Slovensko
SK04	Východné Slovensko
FI	Finland
FI19	Länsi-Suomi
FI1B	Helsinki-Uusimaa
FI1C	Etelä-Suomi
FI1D	Pohjois- ja Itä-Suomi
FI20	Åland
SE	Sweden
SE11	Stockholm
SE12	Östra Mellansverige
SE21	Småland med öarna
SE22	Sydsverige
SE23	Västsverige
SE31	Norra Mellansverige
SE32	Mellersta Norrland
SE33	Övre Norrland

1 Introduction

In a European context, regional development is inextricably linked to the regional policy of the European Union. Except for non-EU countries, the EU regional policy represents a significant contribution and a new element that acts beside the original national policies for every member state.

After the terrifying experience of WWII, it was evident that a stable development of Europe and the prevention of potential war conflicts required the establishment of solid relationships based on cooperation in the economic field. Unfortunately, the start of the so-called 'Cold War' and the division of the World into the Eastern and Western Bloc caused that the 'European Project' was for 40 years possible to be realised only in the western part of Europe.

Negative experiences stemming from history and mutual rivalry motivated the original six countries (Belgium, Germany, France, Italy, Luxembourg, and the Netherlands) to establish the European Coal and Steel Community in 1952 and subsequently also the European Economic Community in 1958. This denoted the first step in building a united Europe. The Treaty Establishing the European Economic Community (Treaty of Rome) specified the conditions for creating a common market, one of the stages of integration. However, this also required the elimination of disparities among states and regions. Although the original states did not face significant regional differences, this changed during the enlargement of the communities. The regional aspect increased in importance as new members with considerable regional disparities joined.

A new challenge came when the European Union accepted the former Eastern Bloc countries as members. Different development lasting 40 years caused enormous disparities. From that point, the EU had to deal not only with disparities in specific regions but also with an economic imbalance between the states in the West and the East. For this reason, the regional policy experienced huge evolution in the last two decades.

Despite the regional policy's huge effort, it is necessary to mention that many of the disparities persist. There still exist regions that report considerable differences compared to the rest of the EU. Simultaneously, it is still challenging for the community to deal with disparities caused by the violent division of Europe after WWII.

Crucial for the future of the EU were the events concerning the withdrawal of the United Kingdom from the European Union (commonly called 'Brexit'). Besides the fact that that is for the first time when a country leaves the EU, it is also necessary to mention that the United Kingdom was a net contributor to the EU budget. Therefore, Brexit has nowadays a strong impact on the formulation of the future not just of the regional policy but also on the development of the EU as a whole.

All of the aspects mentioned before were important for formulating this thesis topic, 'Convergence of selected economic indicators in regional development in the European Union'. The thesis concerns evaluating the most recent period to investigate current trends among the regions and give a framework to support the implementation of the regional policy in the future.

2 Objectives and Methodology

2.1 Objectives

The objective of the master's thesis is to evaluate the convergence of selected economic indicators in Regional Development in the EU within the time period 2015-2019. Specifically, the thesis aims to determine tendencies of the economic convergence of NUTS 2 regions in the EU and find common attributes of certain groups of areas that characterise their development in the selected period.

2.2 Methodology

The basis of the thesis is working with relevant data concerning the economic development of units within the EU. The data concern mainly economic indicators, which are supplemented by socioeconomic indicators. The source of the data is the European Statistical Office (Eurostat). The data are taken from the five-year period 2015-2019. The selection of the time framework was decided based on two factors. Firstly, the year 2019 is the most recent one for which most data are available. Secondly, the period was chosen due to the fact that there were no changes in NUTS 2 classification within the EU. Moreover, the period fits into the programming period 2014-2020 and therefore disruption due to the formulation of new goals and strategies does not occur.

There are four main methods used in the thesis: exploratory data analysis for basic characterisation of selected data, beta-convergence and sigma-convergence for determining the prevailing trend of economic development of the units, and cluster analysis used to find similarities among NUTS 2 regions in the selected period.

For individual calculations, relevant software is used. Basic calculations and creating graphs of beta-convergence and sigma-convergence are done in Microsoft Excel. The exploratory data analysis and statistical verification of the models are calculated in SAS Studio. For the cluster analysis, the software Statistica 14 is chosen due to its suitable user interface for this type of analysis. The cartograms are created in Inkscape 1.1.2, an open-source vector graphics editor.

2.2.1 Exploratory Data Analysis

Exploratory data analysis (EDA) is a basic statistical approach that allows the researcher to get acquainted with the data set. It does not require any preconceived idea about the data, which hypothesis testing does. The framework of EDA was developed by John W. Tukey in 1977. Through the exploratory data analysis, it is possible to investigate the location and variation of the data set. (Martinez, 2017)

Mean

The arithmetic mean (also called arithmetic average) is a fundamental statistical summary of central tendency in a sample. It is defined as the sum of all individual values divided by the number of observations.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}$$

(Myatt, 2014)

Median

The Median could be described as the middle value of the sample. In other words, half of the values in the data set sorted from low to high are below median and half of the values are below. If the data set consists of an even number of observations, the middle value is found as an arithmetic average of two values in the middle. Median is often denoted by $\tilde{x}$. In contrast with the arithmetic mean, it is not sensitive to extreme values. (Myatt, 2014)

Variance

"The variance describes the spread of the data and measures how much the values of a variable differ from the mean. For variables that represent only a sample of some population and not the population as a whole, the variance formula is

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1} \tag{2}$$

The sample variance is referred to as s^2. The actual value (x_i) minus the mean value ($\bar{x}$) is squared and summed for all values of a variable. This value is divided by the number of observations minus 1 ($n-1$)." (Myatt, 2014, p. 32)

Standard Deviation

"The standard deviation is the square root of the variance. For a sample from a population, the formula is

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{3}$$

where s is the sample standard deviation, x_i is the actual data value, $\bar{x}$ is the mean for the variable, and n is the number of observations." (Myatt, 2014, p. 33)

"The standard deviation is the most widely used measure of the deviation of a variable. The higher the value, the more widely distributed the variable's data values are around the mean. Assuming the frequency distribution is approximately normal (i.e., a bell-shaped curve), about 68% of all observations will fall within one standard deviation of the mean (34% less than and 34% greater than)." (Myatt, 2014, p. 33)

"It is possible to calculate a normalized value, called a z-score, for each data element that represents the number of standard deviations that element's value is from the mean. The following formula is used to calculate the z -score:

$$z = \frac{x_i - \bar{x}}{s} \tag{4}$$

where z is the z-score, x_i is the actual data value, $\bar{x}$ is the mean for the variable, and s is the standard deviation. A z-score of 0 indicates that a data element's value is the same as the mean, data elements with z-scores greater than 0 have values greater than the mean, and elements with z-scores less than 0 have values less than the mean." (Myatt, 2014, p. 34)

Five-Number Summary

$$x_{min} < \tilde{x}_{0.25} < \tilde{x} < \tilde{x}_{0.75} < x_{max} \tag{5}$$

The fundamental element of the five-number summary is the **range** of a sample. The lowest value in the data set is denoted by x_{min}, the highest by x_{max}. All values between these two form the range. The range is further divided into **quartiles**. The first quartile (lower quartile), denoted by $\tilde{x}_{0.25}$, separates the bottom 25% of values from the highest 75%. The second quartile is the same as the median described above (denoted by $\tilde{x}$). The third quartile (upper quartile), denoted by $\tilde{x}_{0.75}$, separates the bottom 75% of values from the highest 25%. The area between the lower and upper quartile is called the **interquartile range (IQR)**. The interquartile range is important for the determination of outliers. The value is considered to be an outlier when it is below $\tilde{x}_{0.25}$ by more than $1.5 \cdot IQR$ or above $\tilde{x}_{0.75}$ by more than $1.5 \cdot IQR$. (Lipi, 2022)

2.2.2 Economic Convergence

The term convergence indicates the tendency of regions to get closer in the term of a selected indicator. The opposite term is divergence. Originally the only indicator used was Gross Domestic Product (GDP). Nowadays, the measuring of convergence could be used in many other areas. However, this thesis focuses only on measuring economic convergence represented by GDP per capita. The absolute convergence or divergence could be only observed when comparing two regions. In the case of more regions, it is understood as a tendency towards one of the possibilities. The two main ways for calculating convergence are Beta-convergence and Sigma-convergence. (Minařík, 2013)

Beta-Convergence

For usage of this method, the values from the beginning and the end of a selected period are employed. It is assumed that units with low initial values reported higher economic growth than units with high initial values. If this assumption occurs, it is possible to speak about convergence.

The first step for calculating beta-convergence is to create a scatter plot with logarithms of data from the beginning on the horizontal axes and logarithm of the average growth coefficients on the horizontal axes. The average growth coefficient is calculated using the following formula:

$$\bar{k} = \sqrt[n]{\frac{y_n}{y_0}} \tag{6}$$

where $\bar{k}$ stands for average growth coefficient, y_n is the value at the end of the selected period, y_0 is the value at the beginning of the selected period, and n denotes the number of periods.[1]

When the scatter plot is created, the second step is to construct the regression line using the ordinary least squares method. The general equation of the regression line is:

$$\log \bar{k}' = \alpha + \beta \log y_0 \tag{7}$$

where $\log \bar{k}'$ is the logarithm of the predicted average growth coefficient, α and β are the parameters of the regression line, and $\log y_0$ stands for the logarithm of values in the beginning of the selected period. The parameter β decides whether convergence or divergence occurs. If the parameter β is lower than 0 and the regression line is downwards sloping, it is possible to speak about the tendency towards convergence. If the parameter β is greater than 0 and the regression line is upward sloping, it is possible to speak about the tendency towards divergence. If $\beta = 0$, then neither convergence nor convergence occurs.

To analyse beta-convergence, it is also necessary to consider the coefficient of determination. The coefficient of determination r^2 ranges between 0 and 1. It denotes the proportion of the variation in the dependent variable that is predictable from the independent variable. In the beta-convergence it is calculated using the following formula:

$$r^2 = \frac{var \log \bar{k}'}{var \log \bar{k}} \tag{8}$$

(Minařík, 2013)

Sigma-Convergence

It is based on the assumption that in the case of convergence, the variability of the selected indicator declines in time. The variability is measured by the standard deviation calculated

[1] The firs value is denoted by '0'. Therefore, the actual number of periods is n+1, and in the calculation the n-root is used (instead of n-1-root).

for each year. Usually, a logarithm of computed values is used. If the variability increases over time, it is possible to speak about divergence tendency. (Minařík, 2013)

2.2.3 Cluster Analysis

The cluster analysis is used in the practical part for the creation of groups of regions in which similarities during the time period could be found. Such groups are used for the description of common characteristics in the convergence process. (Hebák, 2015)

"The task of cluster analysis or grouping is to divide the set of objects into homogeneous groups: two arbitrary objects belonging to the same group are more similar to each other than two arbitrary objects belonging to different groups. If we wish to apply this recipe in practice, we must find the answers to two basic questions: (a) how to define the similarity between the objects, and (b) in what manner should one make use of the thus defined similarity in the process of grouping." (Wierzchoń, 2018, p. 1)

In this thesis, the Agglomerative Hierarchical Clustering Algorithm is used. The procedure is described in the following way: "First, using a particular proximity measure a dissimilarity matrix is constructed and all the data points are visually represented at the bottom of the dendrogram. The closest sets of clusters are merged at each level and then the dissimilarity matrix is updated correspondingly. This process of agglomerative merging is carried on until the final maximal cluster (that contains all the data objects in a single cluster) is obtained." (Aggarwal, 2013, p. 101)

For the procedure, the method for measuring the distance and the method for grouping objects together is necessary.

Distance Measures

The assumption is, that the objects are characterised by quantitative values. These values must be in the same units, or they must be standardized. According to Hebák, the most commonly used method is the **Euclidean distance**. For measuring the distance between two objects represented by vectors x_i and $x_{i'}$ the following formula is used:

$$d_E(x_i, x_{i'}) = \sqrt{\sum_{j=1}^{p}(x_{ij} - x_{i'j})^2} \qquad (9)$$

(Hebák, 2015)

Methods of Hierarchical Clustering

For the purpose of clustering, **Ward's method** is chosen. "Ward's linkage merges two clusters with minimum between-cluster distance, that is, two clusters that lead to the minimum increase in total within-cluster variance after merging. Therefore, the distance function between two clusters in Ward's linkage is defined as within-cluster variance by considering them as one cluster." (Abu-Jamous, 2015, p. 161)

3 Theoretical Part

3.1 Starting Points of Regional Development

3.1.1 Definition of the Term Region

The term region is nowadays broadly used either in everyday life or in media. It is one of the most frequently used terms in many scientific fields. It is used in the economy, sociology, geography, demography, or political science. However, one exact definition does not exist that would cover the meaning of the term region in all scientific disciplines. Different perspectives on understanding regional issues make it impossible to make one solid interpretation of the term.

One of the possible definitions is provided by the Czech Statistical Office (CZSO). It says that a region could be defined "as a tract of land with more or less definitely marked boundaries, which often serves as an administrative unit below the level of the nation state. Regions have an identity which is made up of specific features such as their landscape (mountains, coast, forest), climate (arid, high-rainfall), language (for example in Belgium, Finland, Spain), ethnic origin (for example Wales, northern Sweden and Finland, the Basque country) or shared history." (CZSO, 2022)

According to the definition, there are many factors that determine the nature of what we understand as a region. Because the goal of this thesis is a statistical analysis of the economic convergence indicators, the view of a region is taken mainly from the administrative perspective. It is because, as an administrative unit, the region is able to report relevant economic data which can be analysed by a proper statistical method.

For effective analysis of a region, it is not sufficient to focus only on regional data but also evaluate approaches to regional development and regional policy. Regions are always involved in state or supranational policies, and as objects of it, "regions are subjected to attempted policy-led transformations designed to ameliorate uneven development for reasons of social justice, welfare and economic efficiency." (Gregory, 2009, p. 636) Therefore, for a proper understanding of regional problematics, the following chapters outline the main concepts of regional development and regional policy.

3.1.2 Regional Development

Throughout history, it has been possible to observe human effort to enhance living conditions in the place of living. Humankind has always tried to find a way to develop such changes that could improve life in a specific area. Therefore, the natural human desire to live better is the driving force of regional development. (Minařík, 2013)

One of the broadest definitions of regional development states that the term "can be seen as a general effort to reduce regional disparities by supporting (employment and wealth-generating) economic activities in regions." (OECD, 2022) To better understand regional development, it is suitable to consider different aspects of it. Wokoun states that the understanding of regional development could be generally taken from two perspectives:

- Practical approach
- Academic approach

According to the practical approach, regional development means increasing the potential of a defined area, mainly by encouraging socioeconomic activities and measures leading to better utilisation of natural resources. Such an approach is important mainly for non-academic institutions, e. g. municipalities, regional councils, or private companies, which primarily use data to evaluate the potential of that region. Such data may be gross domestic product per capita, average wage, education attainment level, or infrastructure quality and availability.

The academic approach uses the knowledge of academic disciplines, such as economy, sociology, and geography, to understand the region's processes and regularities. Based on the findings, the theories are developed, and tools for the application of measures are sought. The academic approach typically deals with issues like uneven deployment of inhabitants, economic activities, and the following searching for tools that may influence these processes. Such an approach is sometimes called regional science. (Wokoun, 2008) (Minařík, 2013)

Image 1: Academic and Practical Approach to Regional Development
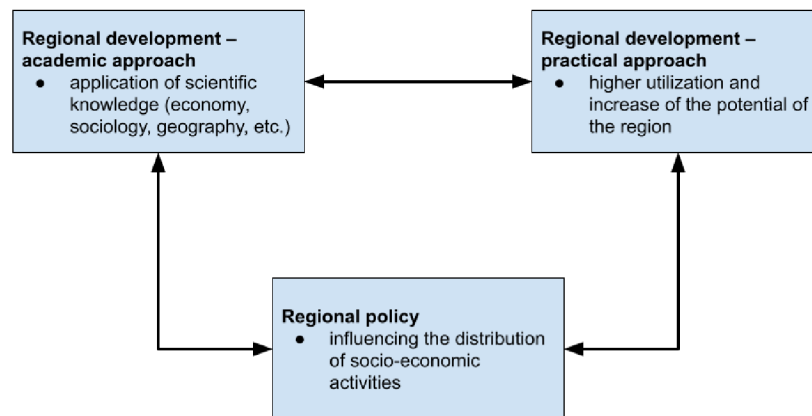
In this diagram, it is possible to see both effects of the regional development understanding. From one point of view, both approaches affect regional policy. At the same time, they are in some way mutually dependent on each other, and it is not possible to isolate them, although, in the short run, they could act independently to some extent. The academic approach brings the intellectual framework and knowledge which may be used in practice. Together with the practical approach, it brings a scheme for the development of regional policy. The regional policy then affects the real regional development, which then affects it back.

The driving forces that direct regional development could be essentially divided into three main categories: economic, social, and environmental. The balance of these three pillars is necessary for the smooth and sustainable development of the region. From the economic perspective, the main aim is the maximal utilisation of production factors, especially the labour force. The ineffective utilisation of the labour force leads to unemployment which is one of the crucial problems of poorer regions. In this case, the effort is to relocate the economic activities into less developed areas. Further economic motives are economic growth and optimal allocation of firms. The second category, social motives, stress the importance of full employment, the welfare of inhabitants, and optimal distribution of income. Finally, there are environmental issues which become a discussed topic mainly since the 1970s. The environmental motives try to lower the increase of harmful effects of pollution. (Wokoun, 2008)

3.1.3 Regional Policy

Like in the case of the region and regional development, there are many definitions of the term regional policy. None of the definitions is generally accepted as universal. Therefore, for defining regional policy, it is necessary to consider more opinions and, based on them, develop own attitude.

Bannock describes the regional policy as "the framework for measures taken in the attempt to reduce disparities between economic development in general and unemployment in particular among different parts of the country. All countries have prosperous and depressed regions, though in some the disparities are greater than the others. In most cases, depressed areas result from the decline of once important industries or other economic activities." (Bannock, 2003, p. 328)

According to Vanhove, regional policy "includes all forms of public intervention intended to ameliorate the geographical distribution of economic activities; in reality regional policy tries to correct certain spatial consequences of the free market economy in order to achieve two interrelated objectives: economic growth and improved social distribution." (Vanhove, 2018, p. 57)

Although the basic understanding of regional policy may be connected only to the redistribution of resources, Gregory states that "regional policy is rarely purely redistributive. It is intended to be transformative. Thus it may, for example, involve bringing work to (unemployed/low-productivity) workers or attempt to address the uneven distribution of cultural facilities (e.g. symphony orchestras, art galleries and theatres) or the regional availability of educational facilities (such as university disciplines, for example). Such policies are usually driven and financed from outside the region – albeit often with regional participation – by national or supra-national state bodies." (Gregory, 2009, p. 636)

The common framework of all definitions motioned above is the idea that regional policy is a purposeful activity of organisations or representatives of state and local administration that provides economic growth to the administrated territory. It follows from the definitions that a general goal of regional policy is to lower disparities among regions and ensure equal chances for fulfilling regional potential.

3.1.4 Theories of Regional Development and Regional Policy

Approaches to grab the topic of regional development have led to the creation of a considerable number of theories. According to Blažek, although the views are conceptually diverse and their principles very often opposite, it is possible to classify them into two main groups:

- Convergence theories
- Divergence theories

Convergence theories assume that market forces will level out disparities among regions. In contrast, divergence theories claim that the natural tendency of regions is to deepen mutual disparities. The division between convergence and divergence theories is, however, considerably problematic due to ambiguous definitions. Moreover, both groups of theories report a significant difference in understanding of the time horizon, although generally said, the convergence theories operate in a longer time horizon than the divergence theories. Nevertheless, the division is necessary because it is one of the basic possible divisions of rural development theories.

Because the regional policy is a state policy, naturally, it deals with one of the most fundamental questions of economics how big state interventions in the economy should be. The main directions are represented by noninterventionism (neoclassical and partly also neoliberal approaches) and interventionism (Keynesian and neo-Marxist approaches).

Approaches to regional development are:

- **Neoclassical theories** – fall mainly into the period between WWI and WWII. They stress the importance of a free market and low state intervention in the economy. From the regional development perspective, the neoclassical theories assume convergence which might be achieved by tools supporting an increase in labour mobility.
- **Keynesian theories** – prevailing theories after WWII. They profile themselves as a contrast to neoclassical theories. Therefore, the main assumption is divergence. Keynesian theories justify state interventions for balancing regional disparities. They support tools for the inflow of investments for both public and private sectors in the problematic region.

- **(Neo-)Marxian theories** – do not bring any specific approach to regional development. The main aspect of these theories may be characterised by extreme interventionism from the side of the state. In some countries (for instance in the former communist Czechoslovak Socialist Republic), economic disparities among regions were eliminated but at the expense of the efficiency of the state economy.

- **Neoliberal theories** – partly follow up on neoclassical theories. They expanded in the 1970s during a crisis of Keynesian theories, which were predominant until that time. Neoliberal approaches are characteristic of the encouragement of local initiatives. This includes support of small and medium enterprises, decentralisation, and effort to lower regulations.

- **Institutional theories** – from the 1980s. Important is the creation of institutions and communication which may support communication and cooperation among regional actors. Just like neoliberal theories, institutional theories support small and medium enterprises.

Nowadays, the prevailing theories are neoliberal and institutional. Keynesian theories still have their importance in general macroeconomics; however, from the perspective of regional development, they are more or less disqualified. The current view on theories of regional development and regional policy is called 'eclectic'. It means that the concept does not try to find only one ideal theory but to combine tools, mainly from neoliberal and institutional approaches.

Another division of concepts of regional development is into endogenous and exogenous. The exogenous model enables external interventions of central authorities into the functioning of regions. This model is based on the assumption that regional development is determined by the lack of some production factors and assumes that it is possible to move needed production factors in the region which suffers from their scarcity. The endogenous model stresses the efficient utilization of regional resources. This model prefers a decentralised way of management. The current conception uses the benefits of both approaches – utilisation of local potential and support from the centre. This trend is visible in the development of the regional policy of the EU. (Blažek, 2011)

3.2 Regional Policy of the European Union

The regional policy belongs to the main priorities of the European Union. Together with the Common Agriculture Policy, it represents the primary area for allocation from the EU budget. The main principle is that the EU regional policy is an investment policy. It is assumed that investments into regions may improve the socioeconomic situation of disadvantaged areas. The main aim of the community is to support job creation, competitiveness, economic growth, sustainable development, and improvement of the quality of life within the regions. The official name of the EU regional policy is Cohesion Policy. There are two fundamental values that shape the form of the policy:

- **Solidarity** – "economically more developed states finance less developed states through contributions to the common budget." (Stejskal, 2009, p. 40)
- **Cohesion** – "reducing disparities between the various regions and the backwardness of the least-favoured regions." (EUR-Lex, 2022c)

It is necessary to mention that the Cohesion Policy aims not to replace the national regional policies but rather to serve as a support for regional development in general.

3.2.1 EU Institutions and Cohesion Policy

The institutional base for the Cohesion Policy is provided mainly by the following EU institutions:

- **European Commission** – Cohesion Policy falls under the Directorate-General for Regional and Urban Policy (REGIO). The agenda also has a relevant commissioner, nowadays called 'European Commissioner for Cohesion and Reforms'.
- **Council of the European Union** – Cohesion Policy does not have its own configuration; however, the agenda is one part of the General Affairs Council configuration (GAC).
- **European Parliament** – the portfolio belongs to the European Parliament Committee on Regional Development (REGI).
- **Committee of the Regions** – consists of representatives from the sub-national level (e. g., members of regional governments, mayors). It does not have decision making power; it is only a consultative body for the other and executive bodies of the EU. (European Commission, 2022c)

3.2.2 Historical Development of the Cohesion Policy

Although nowadays, the Cohesion Policy is perceived as one of the main priorities of the EU, this was not the case at the origin of the European idea. The six founding member states of the European Economic Community were not in the position in which they had to solve significant regional disparities. The exceptions were minority only, e. g. Mezzogiorno in Italy. At the same time, it was believed that stronger economic cooperation and the removal of trade barriers would eliminate local disparities to the minimum. (Baldwin, 2012)

Despite a low need for common regional policy, the first mention of this could be found in the Treaty establishing the European Economic Community (Treaty of Rome). In the preamble, the founding member states determine an objective "to strengthen the unity of their economies and to ensure their harmonious development by reducing the differences existing between the various regions and by mitigating the backwardness of the less favoured" (CVCE, 2022). However, it is evident that such a statement was rather proclamatory and did not commit signatories to any particular action.

The situation changed in 1973 with the first enlargement of the communities. Ireland, which joined the communities together with the United Kingdom and Denmark, represented the first 'poor' member state where significant regional disparities in contrast with the rest of the communities occurred. As a consequence of this, the European Regional Development Fund (ERDF) was established. The important initiative for creating the fund came from the side of the United Kingdom. It was due to the fact that the United Kingdom was expected to be a significant contributor to the community budget, especially to the Common Agriculture Policy. By establishing the fund, the UK expected to balance an unfavourable position. However, the origin of the fund experienced a hard start, and therefore the fund was finally introduced in 1975.

The crucial reason for a change of spending priorities was the admission of Greece in 1981 and the subsequent admission of Spain and Portugal in 1986 into the communities. These three states were significantly poorer. Moreover, the communities experienced an increase in interregional disparities since 1975 and further deepening of disparities was expected with the accession of new members. (Baldwin, 2012) (Senior Nello, 2012)

3.2.3 Single European Act and Origin of the Cohesion Policy

As it was already said, the first mention of regional policy could be found already in the Treaty of Rome. However, it was the Single European Act that gave the treaty base for cohesion for the first time. The Single European Act (SAE) was the first substantial review of the Treaties. Its main aim was the completion of the internal market and deepening cooperation within the communities. However, this could only be achieved if there was an improvement in regional cohesion, particularly in Greece, Spain, and Portugal. Simultaneously, the question arose of where to take the funding for the newly defined policies. Therefore, the need for reforming the regional policy and the way of financing was indisputable. (Senior Nello, 2012)

The origin of Cohesion Policy as such can be traced back to 1988, when significant reform of regional policy took place. Insufficient coordination between the agricultural policy and the social policy proved to be a problem for seamless reform. Therefore, the policies were integrated into the so-called 'Structural Policy', which also included better cooperation among funds. The reform also included the formulation of basic principles of Cohesion Policy (see chapter 3.2.10). (Stejskal, 2009)

The reform also brought up a new classification of territorial units below the national level. The new methodology has been named NUTS (Nomenclature of Territorial Units for Statistics, originally from French: *Nomenclature des unités territoriales statistiques*). Below the national level denoted by NUTS 0, there have been five levels set up: NUTS 1 for sections of a country, NUTS 2 as main regions for implementation of the Cohesion Policy, NUTS 3 for small regions, and lastly, NUTS 4 and NUTS 5 for the smallest units[2]. (Bache, 2007)

All changes made have also been reflected in the European budget. As already mentioned, the question concerned the problem of sources of funding. The result was a new agreement between the Commission, the Council, and the European Parliament on long-term financial perspectives (later renamed Multiannual Financial Frameworks). This agreement brought up a financial outlook for the period 1988-1992 (known as the 'First Delors Package' after the president of the European Commission, Jacques Delors). (Senior Nello, 2012)

---

[2] NUTS 4 and NUTS 5 were later replaced by Local Administrative Units LAU 1 and LAU 2

3.2.4 Cohesion Policy 1989-1993

The programming period 1989-1993 was the first case when the community, according to the reform in 1988, developed so-called priority objectives for which the allocations of the Funds should have been concentrated. Specifically, these were the five objectives:

1. "promoting the development and structural adjustment of the regions whose development is lagging behind;
2. converting the regions, frontier regions or parts of regions (including employment areas and urban communities) seriously affected by industrial decline;
3. combating long-term unemployment;
4. facilitating the occupational integration of young people;
5. with a view to reform of the common agricultural policy:
   a. speeding up the adjustment of agricultural structures, and
   b. promoting the development of rural areas" (EUR-Lex, 2022a)

Regarding the funding, the aim was to double the volume of Structural Funds allocations by 1993. Overall, the expenditure on cohesion policy was about 22% of the whole budget. At this point, it is necessary to mention that until 2000 the Multiannual Financial Frameworks (MFF) did not fully fit into the programming periods. Therefore, the time frameworks are slightly different. (European Parliament, 2022)

Table 1: MFF 1988-1992

1988-1992 (1988 prices)	ECU[3] million	Share in %
Multiannual Policies	9,310	4%
Cohesion	53,140	22%
Agriculture	142,200	59%
Other Policies	12,488	5%
Administration	22,700	9%
Total	239,838	100%

Source: own elaboration, based on Benedetto, 2019

3.2.5 Cohesion Policy 1994-1999

For the programming period 1994-1999, two main novelties were prepared: the establishment of the Cohesion Fund and the origin of a new consultative body – the Committee of the Regions. The Cohesion Fund (CF) was agreed at Maastricht in 1991. Its

---

[3] European Currency Unit, the unit of account of the Communities, before it was replaced by the euro

aim was to support members states whose GDP per capita was less than 90% of the Community average. Therefore, it was an instrument that aimed to help poorer countries, not regions. The Committee of the Regions was created as an advisory capacity that aimed to assist the legislative process as a consultative body. (Teasdale, 2012)

The priority objectives remained the same; however, due to the accession of Finland and Sweden in 1995 (together with Austria), the 6th goal was formed:

- "Objective 6: Regions with a low density of population in the extreme north of Finland and Sweden" (Senior Nello, 2012, p. 347)

The budget for the period 1993-1999 (again still not fully corresponding with the programming period) was designed according to the 'Second Delors Package'. Its main aim was to increase spending on Structural Funds. The proportion of budget spending on cohesion increased from 22% in the previous programming period to 34%. Regarding the percentage amount, this number has remained approximately similar until the present, when this thesis is elaborated (2022).

Table 2: MFF 1993-1999

1993-1999 (1992 prices)	ECU million	Share in %
Internal Policies	31,587	6%
Cohesion	176,398	34%
Agriculture	255,570	49%
External Action	32,400	6%
Administration	25,480	5%
Total	521,435	100%

Source: own elaboration, based on Benedetto, 2019

3.2.6 Cohesion Policy 2000-2006

In March 1999, at the Berlin European Council, the future development of the EU was discussed. The main topic was the impact of the enlargement of new countries from the former Eastern bloc. These countries were significantly poorer, Bache states, that the average GDP of new theses states was around one-third of the EU-15 average. (Bache, 2007)

The output of the Berlin European Council was an agreement on Agenda 2000. The deal, among other things, reformulated the original six objectives into three new points:

1. "The less well developed areas of the EU, which are defined as those whose GDP per capita is less than 75% of the EU average.
2. The economic and social conversion of regions that were facing natural difficulties including declining rural areas and those dependent of fishing.
3. Improvement of human capital by promoting employment, education and professional training." (Senior Nello, 2012, p. 346)

Besides the priority objectives, for the programming period 2000-2006, the EU also defined the so-called Community Initiatives. The Initiatives were launched already in the programming period 1989-1993. There was a considerable amount of them focused on specific targets. However, for the programming period 2000-2006, they were clearly defined and reduced to the following four:

- "LEADER+ (rural development)
- INTERREG II (cross-border, transnational and interregional co-operation)
- URBAN (economic and social regeneration of cities and urban neighbourhood)
- EQUAL (transnational co-operation to combat all kinds of discrimination and inequalities in the labour market)" (Senior Nello, 2012, p. 345)

The main aim of the whole European Union in the programming period 2000-2006 was the preparation for the enlargement which took place in 2004. For this purpose, pre-accession instruments were implemented. There were three instruments which took approximately 3% of the 2000-20006 budget:

- **Phare** – to set up necessary administrative and institutional capacity to fulfil accession conditions. It also aimed to help in strengthening economic and social cohesion.
- **ISPA** – investment into infrastructure in the environmental field and into the trans-European network.
- **SAPARD** – investment into agriculture and rural development. (euroskop.cz, 2022)

After the accession of particular countries into the EU, the pre-accession instruments were transformed into European Structural and Investment Funds (ESI Funds).

Table 3: MFF 2000-2006

2000-2006 (1999 prices)	EUR million	Share in %
Internal Policies	43,830	7%
Cohesion	213,010	33%
Agricultural and Fish	267,370	42%
Environment and Rural Development	30,370	5%
Pre-Accession Aid	21,840	3%
External Action	32,060	5%
Administration	33,660	5%
Total	642,140	100%

Source: own elaboration, based on Benedetto, 2019

3.2.7 Cohesion Policy 2007-2013

It was the period of the last enlargement of the EU so far. In 2007, Bulgaria and Romania joined the EU. In 2013, at the very end of this programming period, the last enlargement so far occurred when Croatia joined the EU. Priority Objectives and Community Initiatives from the previous programming period were integrated into three new objectives:

- Convergence
- Regional competitiveness and employment
- European territorial cooperation

The funding of cohesion policy was simplified, and the responsibility of funding was given to only three of the ESI Funds – ERDF, ESF (European Social Fund), and CF.

Convergence aimed to support the least developed regions. The help was concerned mainly with stimulating economic growth and employment. The support came from all three main ESI Funds – ERDF, ESF and CF. Whereas the support from ERDF and ESF was concerned with NUTS 2 regions, the eligibility for taking contributions from the CF was dedicated to the whole state.

The second objective, Regional competitiveness and employment, covered regions not eligible for the convergence objective. Its main goal was to support competitiveness, the attractiveness of a particular region, and employment. The financing came from ERDF and ESF.

The last objective, European territorial cooperation, had two subcategories. The first one was cross-border cooperation. Unlike the convergence, it aimed to NUTS 3 level regions to

support connection among bordering areas. The second category was transnational cooperation. The funding came from the ERDF. (European Commission, 2007)

Due to the enlargements in 2004 and 2007, the average GDP per capita in the EU decreased. Therefore, the system of phasing-out regions was adopted. This applies to the regions which were eligible for the support in EU-15 but not in EU-25, respectively EU-27. As the pre-accession aid concerned only two states (Romania and Bulgaria), there was no need to create an extra item in the EU budget. Therefore, the funding of cohesion increased from 33% to 36%.

Table 4: MFF 2007-2013

2007-2013 (2004 prices)	EUR million	Share in %
Competitiveness	74,098	9%
Cohesion	308,041	36%
Agricultural and Fish	293,105	34%
Environment and Rural Development	78,239	9%
Freedom, Citizenship	10,770	1%
Global Europe	49,463	6%
Administration	49,800	6%
Total	863,516	100%

Source: own elaboration, based on Benedetto, 2019

3.2.8 Cohesion Policy 2014-2020

For the programming period 2014-2020, the Cohesion Policy set 11 thematic objectives:

1. "Strengthening research, technological development and innovation
2. Enhancing access to, and use and quality of, information and communication technologies
3. Enhancing the competitiveness of SMEs
4. Supporting the shift towards a low-carbon economy
5. Promoting climate change adaptation, risk prevention and management
6. Preserving and protecting the environment and promoting resource efficiency
7. Promoting sustainable transport and improving network infrastructures
8. Promoting sustainable and quality employment and supporting labour mobility
9. Promoting social inclusion, combating poverty and any discrimination
10. Investing in education, training and lifelong learning
11. Improving the efficiency of public administration" (European Commission, 2014)

All objectives were supported by funds from the ERDF. However, objectives 1-4 were considered to be the main priorities for the ERDF. The ESF's main priorities were objectives 8-11. The CF supported objectives 4-7 and 11. (European Commission, 2014)

Table 5: MFF 2014-2020

2014-2020 (2011 prices)	EUR million	Share in %
Competitiveness	125,614	13%
Cohesion	325,149	34%
Agricultural and Fish	277,851	29%
Environment and Rural Development	95,328	10%
Freedom, Citizenship	15,686	2%
Global Europe	58,704	6%
Administration	61,629	6%
Total	959,961	100%

Source: own elaboration, based on Benedetto, 2019

It is necessary to mention that this period was also affected by the withdrawal of the United Kingdom from the EU. Based on the referendum from 2016, the UK withdrew from the EU on the 31st of January 2020. On the 31st of December, after the end of the transitional period, the United Kingdom left the EU and became the first state which ended its membership in the community. Since then, the EU again consists of 27 states (EU-27).

3.2.9   Cohesion Policy 2021-2027

The new period, which started in 2021, belongs to one of the most difficult in the history of the community. Firstly, the EU must cope with the consequences of Brexit and the resulting loss of one of the net contributors to the EU budget. Secondly, the EU and the rest of the world have to face the Covid-19 pandemic.

For the programming period 2021-2027, the community has set the following five policy objectives:

1. "a more competitive and smarter Europe
2. a greener, low-carbon transitioning towards a net zero carbon economy
3. a more connected Europe by enhancing mobility
4. a more social and inclusive Europe
5. Europe closer to citizens by fostering the sustainable and integrated development of all types of territories" (European Commission, 2022d)

The following table summarises the Multiannual Financial Framework allocations for the programming period 2021-2027. However, it is necessary to mention that the EU budget also consists of the so-called 'NextGenerationEU', which contributes to the budget by an additional approximately EUR 750 billion (in 2018 prices).

Table 6: MFF 2021-2027[4]

2021-2027 (2018 prices)	EUR million	Share in %
Market, Innovation, Digital	166,303	15%
Cohesion and Values	391,974	35%
Agriculture and Fish	254,247	22%
Environment and Rural Development	82,379	7%
Migration and Border Management	30,829	3%
Security and Defence	24,323	2%
Neighbourhood and the World	108,929	10%
Administration	75,602	7%
Total	1,134,586	100%

Source: own elaboration, based on Benedetto, 2019

3.2.10 Principles of Cohesion Policy

As it was mentioned, the main principles of the Cohesion Policy are solidarity and cohesion. Since the reform of Structural Funds in 1988, there exist four principles:

- **Concentration** – the community defines its priority objectives to ensure the concentration of support from funds into areas of the greatest need. The effort is to direct resources and measures to the poorest regions and states. The applied policies should be closely coordinated.

- **Partnership** – includes collaboration between the European Commission and entities on all levels (national, regional, or local). It is necessary to develop horizontal cooperation among organisations at the same level as well as vertical cooperation among different levels. The development should be a result of a collective process.

- **Programming** – the help from the fund is not designed for individual projects but according to multiannual programmes. These programming periods last seven years and nowadays also correspond with the financial perspectives. The main aim of the programming principle is to encourage actors to develop long-term strategies regarding the development of a particular region.

---

[4] NextGenerationEU not included

- **Additionality** – this principle may ensure that the allocations from the Funds do not aim to substitute structural financing of Member States, in other words, to replace national measures. The objective is that the Structural and Investment Funds should serve as an additional contribution to national public spending. (Senior Nello, 2012)

Among other important principles also regarding the Cohesion Policy belong:

- **Subsidiarity** – "to ensure that decisions are taken as closely as possible to the citizen. Except in cases where the EU has exclusive competence, action at European level should not be taken unless it is more effective than action taken at national, regional or local level. Subsidiarity is closely bound up with the principles of proportionality and necessity, meaning that any action by the Union should not go beyond what is necessary to achieve the objectives of the Treaty." (European Commission, 2022c)
- **Co-financing** – "Grants shall involve co-financing. As a result, the resources necessary to carry out the action or the work programme shall not be provided entirely by the grant. Co-financing may be provided in the form of the beneficiary's own resources, income generated by the action or work programme or financial or in-kind contributions from third parties." (EUR-Lex, 2022b)
- **Monitoring and evaluation** – "it is about continuous monitoring and evaluation of the implemented measures and the overall efficiency of the funds spent. The expected impact of the project before its approval, project implementation, and its real benefits are evaluated." (Stejskal, 2009, pp. 40-41)

3.2.11 European Structural and Investment Funds

The European Structural and Investment Funds (ESI) serve as the primary tool of the Cohesion Policy. They aim to fulfil defined goals and priorities within multiannual cycles. Through them, financial means are distributed. The main goal is to reduce disparities among regions in particular areas. There are 5 Structural and Investment Funds:

- European Regional Development Fund
- European Social Fund
- Cohesion Fund
- European Agricultural Fund for Rural Development
- European Maritime and Fisheries Fund (dotaceeu.cz, 2022)

Primary funds for implementing the EU Cohesion Policy are the European Regional Development Fund, European Social Fund, and Cohesion Fund.

European Regional Development Fund

Regarding the value of monetary means, the European Regional Development Fund (ERDF) is the largest of the Structural Funds. The fund was established in 1975. Its main goal is to eliminate disparities among regions and support regional development, mainly in the case of poor regions. The support is addressed to measures aiming to create jobs, improve infrastructure, invest in education, or increase awareness of environmental issues. (European Commission, 2022c)

European Social Fund

Established in 1958, the European Social Fund (ESF) is the oldest of the Structural Funds. It serves as the main tool of social policy and employment in the EU. The primary activity is the investment into human capital. It should support the adaptability of workers, access to employment, or help people from disadvantaged groups. Currently, the European Social Fund is denoted as ESF+. (European Commission, 2022c)

Cohesion Fund

The Cohesion Fund (CF) is designed to help in the field of environment and infrastructure. In contrast with the other ESI funds, the eligibility for support is not determined according to the performance of a specific region but on the performance of the whole state. The funding from the CF can access those member states whose GNI per capita is lower than 90% EU-27 average. (European Commission, 2022c)

European Agricultural Fund for Rural Development

Even though it does not belong to the main instruments for Cohesion Policy, the European Agricultural Fund for Rural Development (EAFRD) has an important role. Because disadvantaged regions are often rural, they might use funding aimed to support agriculture. The EAFRD belongs to the instruments of Common Agriculture Policy (CAP). Its main goal is to support the competitiveness of agriculture, improve the environment, and promote quality of life in rural areas. (European Commission, 2022c)

European Maritime and Fisheries Fund

Out of the 5 Structural Funds, the last one is the European Maritime and Fisheries Fund (EMFF). Its main goal is the fulfilment of the Common Fisheries Policy (CFP). It aims to achieve economically feasible, socially and environmentally responsible fisheries and aquaculture. Because the EMFF is relatively small, it usually goes together with support from other ESI funds. (European Commission, 2022c)

Just Transition Fund

The Just Transition Fund (JTF) is a newly introduced instrument for the programming period 2021-2027. It was established within the new challenges regarding the European Green Deal and the aim of climate neutrality by 2050. Because the transition towards climate neutrality may deepen current regional disparities and harm poorer regions, the fund focuses on helping such regions in adapting to new conditions. (European Commission, 2022b)

3.2.12 EU Budget

The areas which belong to competencies of the EU must be financed by the common European budget. Since 1970 the budget has been based on the system of own resources. The resources were affected by the development of the EU integration. At present, the system consists of the following resources:

- Own resource based on customs duties
- Gross National Income-based own resource (approximately 70 % of the revenue)
- Contributions based on the value-added tax collected by the Member States
- Contribution based on non-recycled plastic packaging waste (newly introduced resource in the programming period 2021-2027)
- Other resources (e. g. surpluses, penalties)

For the current programming period 2021-2027 also exists an additional tool that aims to eliminate the negative impacts of the Covid-19 pandemic. It is called 'NextGenerationEU' and, unlike the Multiannual Financial Frameworks, is not based on the own resources, but the financial sources will come from borrowings and grants. (European Commission, 2021)

3.2.13 Relationship of Cohesion Policy to the Single Market

Since 1993 the Sigle Market exists within the EU. The objective of the Single Market was to encourage economic development by eliminating all existing barriers. The Cecchini report from 1988 stated that the potential gain of removing such barriers would have been ECU 200 billion (EPRS, 2014). To ensure smooth running of the Single Market, it is necessary to reduce heterogeneity among countries and their regions. For this reason, the Cohesion Policy contributes to reducing disparities and thus to the functioning of the Single Market.

3.2.14 Relationship of Cohesion Policy to the Rural Development and CAP

The main goal of the Cohesion Policy is to help less developed regions in their development. Less developed regions are often rural areas and, therefore, also, agricultural areas. For this reason, the objectives of the Cohesion Policy are also interconnected with Common Agriculture Policy (CAP). As an example could be noted that the European Rural Development Policy is an integral part of CAP. (Tomšík, 2009)

3.2.15 Classification of Regions

The necessity of Cohesion Policy also required a system of administrative division of the regions. Therefore, the EU developed a system called NUTS (Nomenclature of territorial units for statistics, *Nomenclature des unités territoriales statistiques*). This system has its roots in the 1970s when the European Communities started applying this division based on bilateral agreements among states. The NUTS system became necessary in 1988 during the reform of EU Structural Funds. Nowadays, it is the primary tool for the statistical office of the European Union (Eurostat) when providing regional data.

There are three main NUTS levels defined:

- NUTS 1: major socio-economic regions
- NUTS 2: basic regions for the application of regional policies
- NUTS 3: small regions for specific diagnoses

There are also two levels below NUTS 3, called LAU (Local Administrative Unit).

- LAU 1 (formerly NUTS 4)
- LAU 2 (formerly NUTS 5)

For the NUTS division, there are recommended population sizes:

Table 7: Recommended Population for NUTS Classification

Level	Recommended min. population	Recommended max. population
NUTS 1	3,000,000	7,000,000
NUTS 2	800,000	3,000,000
NUTS 3	150,000	800,000

Source: own elaboration, based on Eurostat, 2018

Because the size of the countries is very different (e. g., Germany includes 16 NUTS 1 regions whereas the whole Luxembourg itself is NUTS 3), there is also the NUTS 0 level which denotes EU member states regardless of their size.

For the application of the Cohesion Policy, the NUTS 2 regions represent the most important level because the amount of funding from the EU budget is dependent on the economic performance of each NUTS 2 region. More specifically, the GDP per capita of each NUTS 2 region is expressed as the percentage of the EU average, and according to it, the funds are allocated. The highest amount of support goes to regions whose GDP per capita is lower than 75% of the EU average. (European Commission, 2022a)

4 Practical Part

4.1 Economic Convergence

The aim of the regional policy of the European Union (and previously the European Communities) is to reduce disparities among regions that underwent different economic development in the past. The main criterion for evaluating the trend and the subsequent decision making, to which areas should the support from the EU go, is the GDP per capita. For this reason, the first step is the analysis of economic convergence in the EU.

For the purpose of testing the economic convergence, the GDP per capita at current market prices is used. The data are analysed on two levels. Firstly, the national (or NUTS 0) values are tested to investigate the tendency among EU countries. Secondly, the data are processed on the NUTS 2 level to test whether there is a convergence among NUTS 2 regions which are the primary units for implementing the Cohesion Policy.

4.1.1 Beta Convergence – EU Countries

The computation of beta-convergence includes 27 EU member states. The United Kingdom is not considered because it is no longer a member of the EU at the time of writing this thesis. For beta-convergence, it is necessary to collect data from the beginning and the end of the period. The following tables represent the exploratory data analysis from the years 2015 and 2019:

Table 8: Summary Statistics – GDP per Capita, 2015, 27 Countries

Analysis Variable: y_0						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
27	27192.59	18840.12	6400.00	12900.00	21100.00	38600.00	91400.00

Source: own elaboration in SAS Studio, data from Eurostat, 2022

Table 9: Summary Statistics – GDP per Capita, 2019, 27 Countries

Analysis Variable: y_4						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
27	31585.19	20870.09	8800.00	17100.00	25300.00	43600.00	102200.00

Source: own elaboration in SAS Studio, data from Eurostat, 2022

From the initial and ending values, the average growth coefficient is computed for each member state. The scatter plot is created with the logarithm of initial values on the horizontal

axis and the logarithm of average growth coefficients on the vertical axis. Then the regression line is calculated using the ordinary least squares method. This line represents the tendency of EU member states to either convergence or divergence.

Figure 1: Beta-Convergence for GDP per Capita, 27 Countries
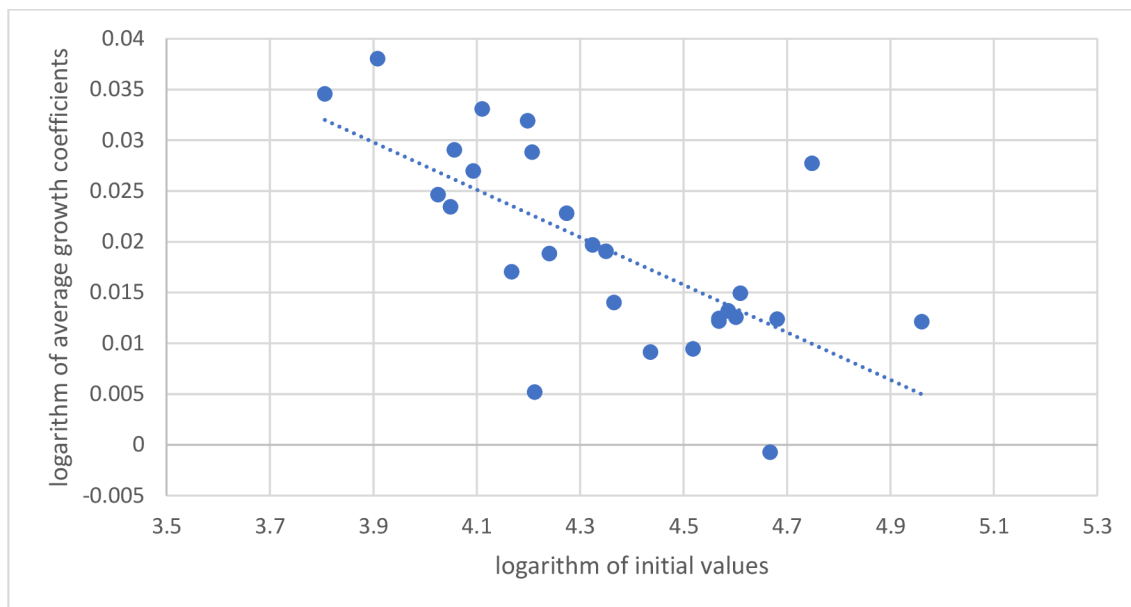
Source: own elaboration in MS Excel, data from Eurostat, 2022

As can be seen from the graph, the regression line is downwards sloping. This indicates a tendency towards economic convergence. To make sure that the tendency is statistically significant, a statistical analysis has to be done.

Table 10: Statistical Analysis of Beta-Convergence for GDP per Capita, 27 Countries

Number of Observations	27

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.00113	0.00113	21.70	<.0001
Error	25	0.00131	0.00005225		
Corrected Total	26	0.00244			

Root MSE	0.00723	R-Square	0.4647
Dependent Mean	0.01937	Adj R-Sq	0.4433
Coeff Var	37.31968		

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	0.12110	0.02188	5.53	<.0001
$\log y_0$	logarithm of initial values	1	-0.02341	0.00502	-4.66	<.0001

Source: own elaboration in SAS Studio, data from Eurostat, 2022

50

The statistical analysis of the regression reports that both parameters are statistically significant as their p-values are lower than 0.01%. The slope of the function is negative, which indicates a tendency towards convergence. However, the coefficient of determination of 46.47% indicates a low quality of the model. If the regression model is not conclusive, there is the possibility to analyse the data by dividing the scatter plot into four quadrants. These quadrants are separated by the mean of horizontal and vertical axis values.

The division gives four interpretations of the units included. Units in the first quadrant report high initial value and above-average growth. Therefore, they tend to diverge. The second quadrant includes units with low initial value and above-average growth. Together with the fourth quadrant, which represents units with high initial value and below-average growth, it creates a group of units that tend to converge. Lastly, the third quadrant represents units that, despite the low value of the initial position, report below-average growth, and therefore they tend to diverge.

Figure 2: Correlation Diagram for GDP per Capita, 27 Countries
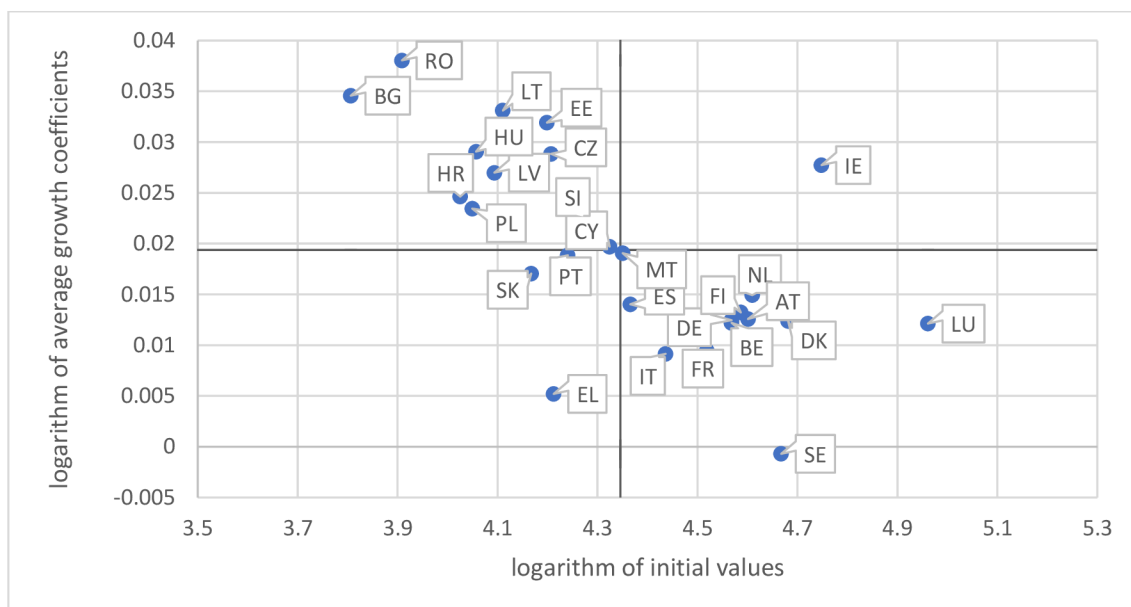
Source: own elaboration in MS Excel, data from Eurostat, 2022

According to the diagram, the only country which belongs to the first group is Ireland which reports the second highest GDP per capita at the beginning and above-average growth in the observed time period. The second quadrant contains all countries which were part of the former Eastern Bloc except Slovakia. This demonstrates that those countries with lower GDP per capita at the beginning of the period report a tendency towards convergence to the EU

average. Simultaneously it is necessary to mention that the Czech Republic, Estonia, and Slovenia are close to the first quadrant, which indicates that they might move to the first quadrant in the future. On the other side in the fourth quadrant are countries of the so-called EU-15 (states that entered the EU before 2004) except Greece and Portugal. This result is expectable because the convergence assumes that units with high initial value report lower growth. Again, it is possible to observe a possible tendency for Italy and Spain to move to the third quadrant. The third quadrant contains three countries, specifically Greece, Portugal, and Slovakia, which means countries that report lower GDP per capita growth than expected according to their initial position. Cyprus and Malta lie close to the intersection of the mean of initial positions and the mean of average growths. Therefore, they do not tend to belong to any of the four groups.

Because the first model is found to be inconclusive, there is the option to exclude some variables according to specific characteristics and improve the model. In this case, it is decided to make two adjustments. Firstly, the two strongest economies, Luxembourg and Ireland, are removed from the model. Secondly, Greece is excluded because, in the selected period, it was still facing the debt crisis, which influenced its GDP growth.

The new model, therefore, includes 24 states with values presented in Table 11:

Table 11: Summary Statistics – GDP per Capita, 2015, 24 Countries

Analysis Variable: y_0						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
24	23770.83	12982.58	6400.00	12650.00	19950.00	37000.00	48000.00

Source: own elaboration in SAS Studio, data from Eurostat, 2022

Table 12: Summary Statistics – GDP per Capita, 2019, 24 Countries

Analysis Variable: y_4						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
24	27550.00	13318.54	8800.00	16550.00	24250.00	41450.00	53800.00

Source: own elaboration in SAS Studio, data from Eurostat, 2022

Subsequently, a new regression model is created.

Figure 3: Beta-Convergence for GDP per Capita, 24 Countries
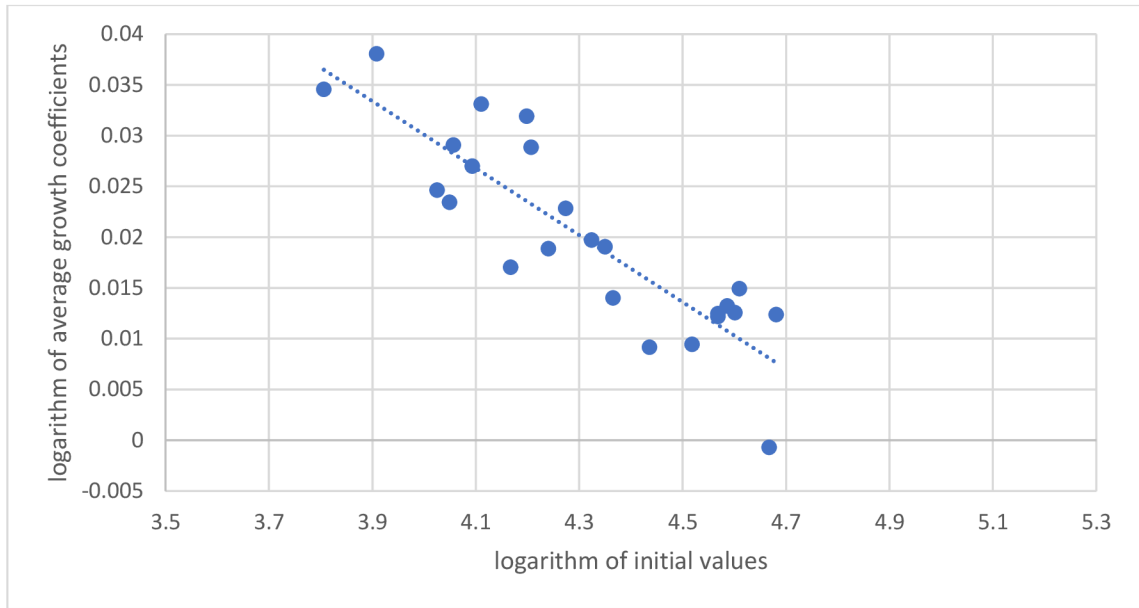
Source: own elaboration in MS Excel, data from Eurostat, 2022

The tendency to convergence does not change as the regression line is downwards sloping. Following statistical analysis tests the appropriateness of the model.

Table 13: Statistical Analysis of Beta-Convergence for GDP per Capita, 24 Countries

Number of Observations	24

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.00161	0.00161	70.56	<.0001
Error	22	0.00050144	0.00002279		
Corrected Total	23	0.00211			

Root MSE	0.00477	R-Square	0.7623
Dependent Mean	0.01991	Adj R-Sq	0.7515
Coeff Var	23.97698		

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	0.16203	0.01695	9.56	<.0001
$\log y_0$	logarithm of initial values	1	-0.03298	0.00393	-8.40	<.0001

Source: own elaboration in SAS Studio, data from Eurostat, 2022

The regression parameters remain statistically significant, again with p-values lower than 0.01%. Compared to the previous model, the most important change is the increase of the coefficient of determination to the value of 76.23%. It means that the quality of the model is sufficient as the usual threshold is 70% or 75%. Using this model, it is possible to interpret

the results of beta-convergence that convergence occurs in GDP per capita among the EU member states.

4.1.2 Sigma Convergence – EU Countries

The second option for investigating economic convergence is the sigma-convergence calculation. For this purpose, from each year of the selected period, the data are taken from all 27 countries, and their standard deviation is calculated. Then the logarithms of standard deviations are taken, and the graph is created.

Figure 4: Sigma-Convergence for GDP per Capita, 27 Countries
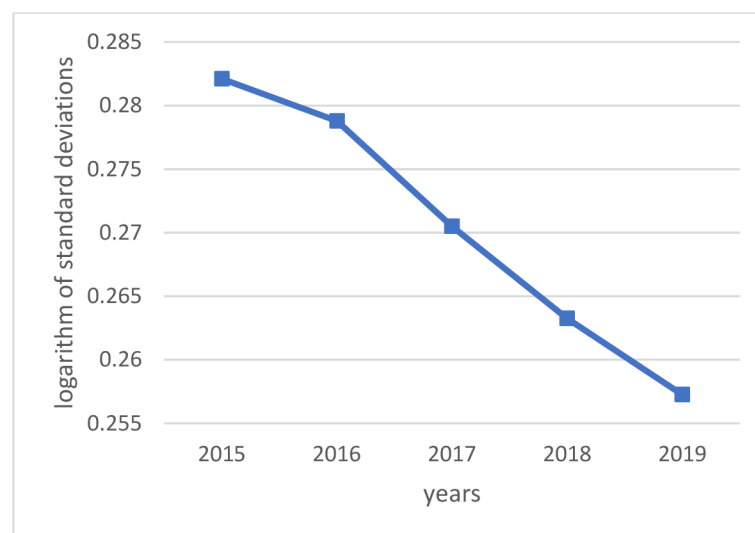
Source: own elaboration in MS Excel, data from Eurostat, 2022

The result of sigma-convergence shows the decline of variability among units in time. Therefore, from this point of view, it is possible to state that the convergence tendency occurs.

4.1.3 Beta-Convergence – EU Regions

After calculating beta-convergence for states, the beta-convergence of NUTS 2 regions is examined. The data set includes 240 NUTS 2 regions of the EU, which are present in the selected period. Again, the regions of the United Kingdom are not considered.

Table 14: Summary Statistics – GDP per Capita, 2015, 240 Regions

Analysis Variable: y_0						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
240	26272.92	14014.24	3900.00	14600.00	25800.00	34900.00	91400.00

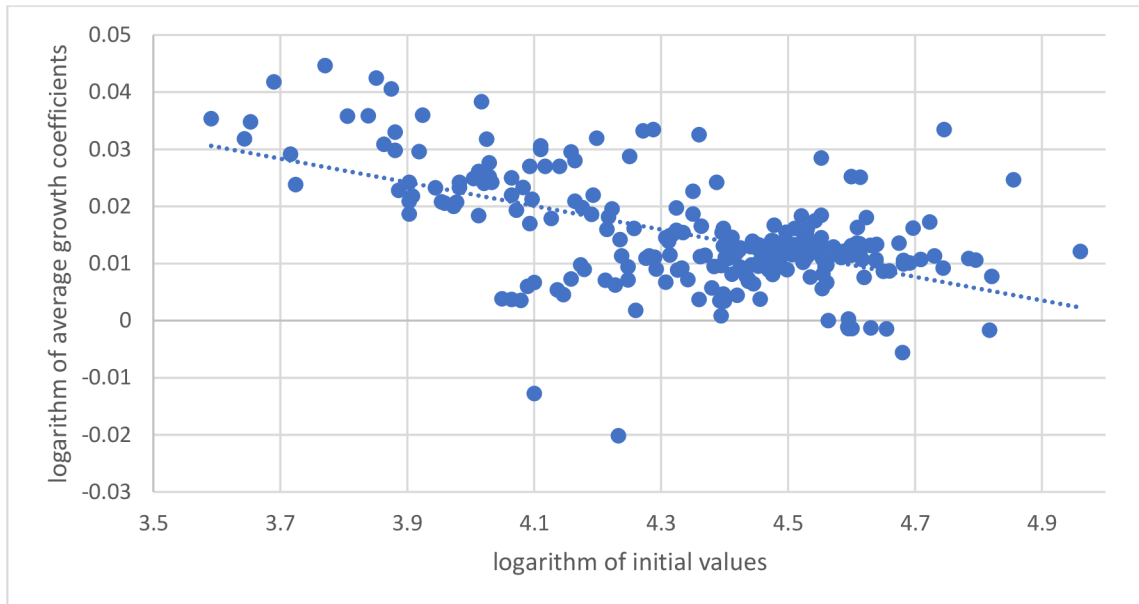Source: own elaboration in SAS Studio, data from Eurostat, 2022

54

Table 15: Summary Statistics – GDP per Capita, 2019, 240 Regions

Analysis Variable: y_4							
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
240	29604.17	15380.07	5400.00	17250.00	29000.00	38900.00	102200.00

Source: own elaboration in SAS Studio, data from Eurostat, 2022

After computation of the average growth coefficients for each NUTS 2 region, the scatter plot with a regression line is created. Similarly to the EU member states, the regression line for beta-convergence of NUTS 2 regions is downwards sloping, indicating the tendency towards convergence of the regions.

Figure 5: Beta-Convergence for GDP per Capita, 240 Regions

Source: own elaboration in MS Excel, data from Eurostat, 2022

Afterwards, the statistical analysis of the model is performed.

Table 16: Statistical Analysis of Beta-Convergence for GDP per Capita, 240 Regions

Number of Observations	240

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.00709	0.00709	112.85	<.0001
Error	238	0.01495	0.00006282		
Corrected Total	239	0.02204			

Root MSE	0.00793	R-Square	0.3216
Dependent Mean	0.01491	Adj R-Sq	0.3188
Coeff Var	53.14109		

| Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	0.10490	0.00849	12.36	<.0001
$\log y_0$	logarithm of initial values	1	-0.02069	0.00195	-10.62	<.0001

Source: own elaboration in SAS Studio, data from Eurostat, 2022

The p-value of both parameters is lower than 0.01%, and therefore both parameters are statistically significant. However, the coefficient of determination with a value of 32.16% again indicates low model quality, and the result is consequently unsure. The approach used in the case of states to divide the diagram into quadrants is also used for regions, but it is not possible to characterise all units because of the limited size of the diagram. Therefore, only specific cases are listed:

- Among the regions from the first quadrant, there are two regions, IE05 (Southern) and IE06 (Eastern and Midland), that report considerable divergence as their GDP growth is high as well as their initial position. Another region with a similar tendency is, among others, also CZ01 (Prague).

- Out of the group of regions with low initial value and above-average growth, regions LT01 (Sostines regionas) and RO32 (Bucuresti - Ilfov) tend to move to the first quadrant in the future.

- Regions EL41 (Voreio Aigaio) and EL53 (Dytiki Makedonia) report extreme divergence as their average growth does not correspond with their initial position.

- There is the possibility that regions FRI2 (Limousin), FRY2 (Martinique), and ITF1 (Abruzzo) may move to the third quadrant in the future due to relatively low average growth. However, in the case of Martinique, the result cannot be thoroughly compared as it is the overseas department of France with specific parameters compared to the regions in Europe.

Figure 6: Correlation Diagram for GDP per Capita, 240 Regions
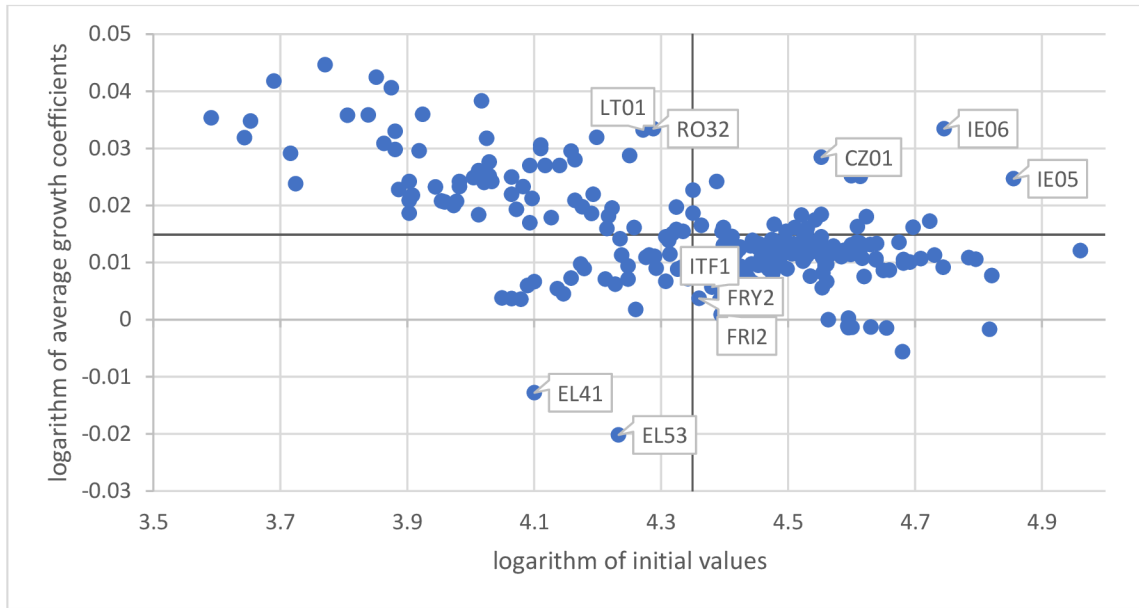
Source: own elaboration in MS Excel, data from Eurostat, 2022

A complete overview of the classification of the NUTS 2 regions into quadrants is available in Appendix C.

To achieve a more conclusive result in terms of beta-convergence, it is necessary to modify the original model. In this case, the same approach as in the case of states is applied. The second model for beta convergence excludes regions of Ireland and Luxembourg as well as all regions of Greece.

Table 17: Summary Statistics – GDP per Capita, 2015, 223 Regions

Analysis Variable: y_0							
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
223	26335.87	13049.56	3900.00	15600.00	26300.00	35100.00	66200.00

Source: own elaboration in SAS Studio, data from Eurostat, 2022

Table 18: Summary Statistics – GDP per Capita, 2019, 223 Regions

Analysis Variable: y_4							
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
223	29673.99	13826.72	5400.00	18900.00	29400.00	38900.00	71100.00

Source: own elaboration in SAS Studio, data from Eurostat, 2022

Figure 7: Beta-Convergence for GDP per Capita, 223 Regions
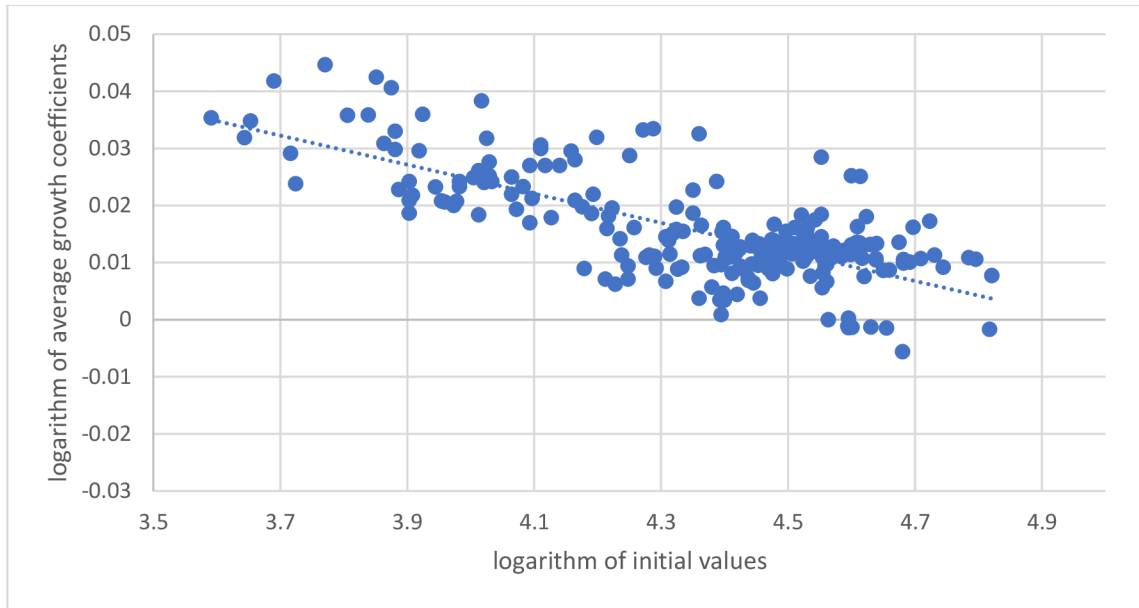
Source: own elaboration in MS Excel, data from Eurostat, 2022

As can be seen from the graph, removing regions from Luxembourg, Ireland, and Greece results in a decrease in the variability of the model. The following table summarises the statistical model of the regression line:

Table 19: Statistical Analysis of Beta-Convergence for GDP per Capita, 223 Regions

Number of Observations	223

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.00983	0.00983	253.61	<.0001
Error	221	0.00857	0.00003877		
Corrected Total	222	0.01840			

Root MSE	0.00623	R-Square	0.5344
Dependent Mean	0.01558	Adj R-Sq	0.5322
Coeff Var	39.96089		

| Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	0.12650	0.00698	18.13	<.0001
log y_0	logarithm of initial values	1	-0.02547	0.00160	-15.93	<.0001

Source: own elaboration in SAS Studio, data from Eurostat, 2022

The most notable change is the change in the coefficient of determination from the original value of 32% to 53% in the new model. This indicates a medium tendency to convergence. However, according to Minařík, such a result may be sufficient to prove the tendency

58

towards convergence among the regions. Nevertheless, higher variability in comparison with the model with states refers to the presence of disparities in the economic development of the regions.

4.1.4   Sigma Convergence – EU Regions

Figure 8: Sigma-Convergence for GDP per Capita, 240 Regions
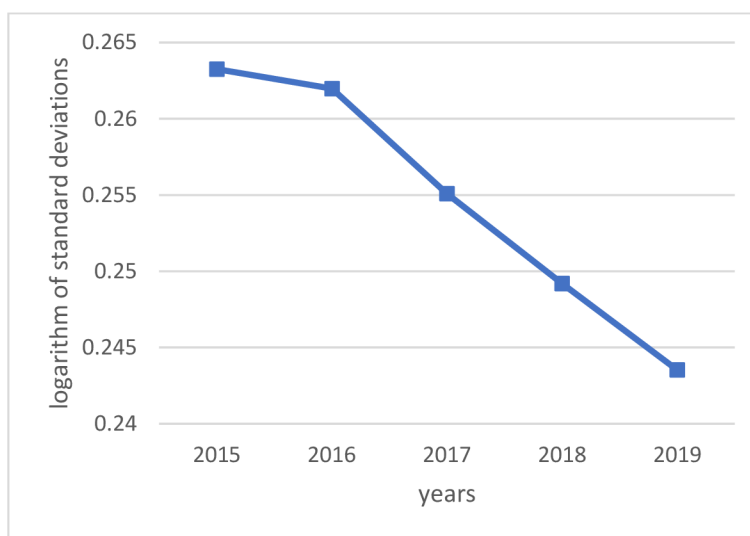
Source: own elaboration in MS Excel, data from Eurostat, 2022

In contrast with beta-convergence, from the perspective of sigma-convergence, it is evident that the variability represented by the standard deviation of values from each year declines in time. Therefore, the sigma-convergence results support the beta-convergence conclusion that there is a tendency towards convergence among the EU NUTS 2 regions.

4.2 Cluster Analysis

In previous examples, beta-convergence and sigma-convergence detect a general tendency towards convergence among EU NUTS 2 regions. However, the variability in the economic development of the areas indicates a different nature of their development. For this reason, it is decided to find similarities among the regions using cluster analysis.

For the cluster analysis, relevant data were selected with the aim to capture appropriate indicators for the economic situation of the regions. Based on Minařík (2013) and Vystrčil (2008), seven indicators are selected. These are the leading indicators that determine the economic situation of the regions as well as their socio-economic condition and ability to

innovate. The following table lists the indicators with a code used in the following calculations:

Table 20: Indicators for Cluster Analysis

Indicator	Description
X1	GDP per capita in Euros at current market prices
X2	unemployment (in %)
X3	net disposable income of households per capita in Euros
X4	population aged 25-64 with tertiary education (ISCED 5-8) (in %)
X5	gross domestic expenditure on research and development (GERD) (in %)
X6	employment in technology and knowledge-intensive sectors (in %)
X7	share of population with tertiary education on employment (in %)

Source: own elaboration

To make the cluster analysis applicable for the whole period, the average values for the period 2015-2019 are calculated for each region in order to overcome breaks in time series which occur in some cases.

It is also necessary to deal with cases when some countries do not provide specific data for the NUTS 2 level but only for the NUTS 0 level. Regarding GERD, this occurs in the case of France, Ireland, and the Netherlands. Therefore, the average GERD value for NUTS 2 regions in these three countries are approximated according to the national value. In other words, the GERD value is in these countries the same for all NUTS 2 regions. It is assumed that the possible differences among these regions will be determined by the other indicators. A similar situation occurred in the case of Greece and employment in technology and knowledge-intensive sectors.

Simultaneously, the regions which are not located in continental Europe or the Mediterranean Sea are excluded. This adjustment is made because these regions lie in distant areas with different conditions, and their integration into the model would not contribute to the aim of finding similarities among the regions. Specifically, this adjustment was applied to following NUTS 2 regions: ES63 (Ciudad de Ceuta), ES64 (Ciudad de Melilla), ES70 (Canarias), FRY1 (Guadeloupe), FRY2 (Martinique), FRY3 (Guyane), FRY4 (La Réunion), FRY5 (Mayotte), PT20 (Região Autónoma dos Açores), PT30 (Região Autónoma da Madeira). Moreover, it is necessary to exclude four NUTS 2 from continental Europe where much of the data is missing: DEB2 (Trier), ITC2 (Valle d'Aosta/Vallée d'Aoste), PT15 (Algarve), FI20 (Åland).

4.2.1 Exploratory Data Analysis

The first indicator used in the Cluster Analysis is the GDP per capita in Euros at current market prices. In the time period 2015-2019, the average GDP per capita calculated for all NUTS 2 regions used in the analysis was EUR 28,176. The lowest value, EUR 4,640, was reported by the region BG31 (Severozapaden), whereas the highest GDP per capita was performed by region LU00, which indicates the whole state of Luxembourg because it does not have any division to NUTS 2 regions due to its size. Luxembourg's average GDP per capita in the selected period was EUR 96,260.

Table 21: Summary Statistics – GDP per Capita, Average 2015-2019, 226 Regions

Analysis Variable: X1						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
226	28175.93	14871.18	4640.00	15280.00	27460.00	37200.00	96260.00

Source: own elaboration in SAS Studio, data from Eurostat, 2022

The second selected indicator is the unemployment rate. Regional unemployment belongs among the most fundamental problems of disadvantaged regions, especially the rural ones. The highest unemployment occurred in the region EL53 (Dytiki Makedonia), where it reached almost 29%. In contrast, the lowest value was reported in region CZ01 (Praha), where the average unemployment for the period 2015-2019 did not exceed 2%.

Table 22: Summary Statistics – Unemployment (%), Average 2015-2019, 226 Regions

Analysis Variable: X2						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
226	8.020443	5.338052	1.860000	4.380000	6.470000	9.680000	28.560000

Source: own elaboration in SAS Studio, data from Eurostat, 2022

Besides the GDP per capita, it is also necessary to consider the net disposable income of households, in other words, the amount of money that a household spends on goods and services. Taking the average values from the period 2015-2019, it is possible to state that the average household in the EU regions could dispose of EUR 15,658 per capita on a yearly basis. The income was the lowest in the region BG31 (Severozapaden) and the highest in LU00 (Luxembourg).

Table 23: Summary Statistics – Net Disposable Income of Households per Capita,

Average 2015-2019, 226 Regions

Analysis Variable: X3						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
226	15658.23	6351.54	3200.00	9340.00	18540.00	20800.00	32900.00

Source: own elaboration in SAS Studio, data from Eurostat, 2022

Considering data that could be characterised as socio-economic, one of the important indicators is the proportion of the population with tertiary education. The contribution of higher education institutions, especially to innovations, may result in a transition of the particular region's economy. The average percentage of people with tertiary education in 2015-2019 was 29.7%. The lowest value (11.9%) reported the region RO21 (Nord-Est), the highest value (55.5%) reported the region BE31 (Prov. Brabant wallon).

Table 24: Summary Statistics – Population With Tertiary Education (%),

Average 2015-2019, 226 Regions

Analysis Variable: X4						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
226	29.704071	9.066551	11.920000	23.060000	28.770000	35.400000	55.500000

Source: own elaboration in SAS Studio, data from Eurostat, 2022

Another important indicator is the expenditure on Research and Development. Eurostat expresses this as a percentage of the regional GDP, called GERD (Gross Domestic Expenditure on R&D). The lowest value, approximately 0.08%, was reported by RO22 (Sud-Est), the highest value (8.34%) was reported by DE91 (Braunschweig).

Table 25: Summary Statistics – GERD (%), Average 2015-2019, 226 Regions

Analysis Variable: X5						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
226	1.709841	1.210002	0.078000	0.832000	1.487917	2.208000	8.340000

Source: own elaboration in SAS Studio, data from Eurostat, 2022

The importance of R&D is also reflected in the percentage of people employed in this branch. In the Eurostat database, the value of employment in technology and knowledge-intensive Sectors is used. In the period 2015-2019, the values ranged between 0.96% in RO22 (Sud-Est) and 8.34% in IE06 (Eastern and Midland).

Table 26: Summary Statistics – Employment in Technology and Knowledge-Intensive Sectors (%),
Average 2015-2019, 226 Regions

Analysis Variable: X6						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
226	3.603142	1.878062	0.960000	2.360000	3.120000	4.440000	9.680000

Source: own elaboration in SAS Studio, data from Eurostat, 2022

The last indicator is the share of population with tertiary education on employment. The value is calculated as the number of employed people with tertiary education divided by the total amount of employed people. The result shows by which percentage the population with tertiary education participates on total employment. The lowest value was reported by RO21 (Nord-Est, 11.7%), the highest value was reported by BE31 (Prov. Brabant wallon, 59%)

Table 27: Summary Statistics – Share of Population With Tertiary Education on Employment (%),
Average 2015-2019, 226 Regions

Analysis Variable: X7						
N	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
226	32.235457	9.300710	11.708114	25.000417	31.140768	38.283583	59.023003

Source: own elaboration in SAS Studio, data from Eurostat, 2022

The summary statistics can also be demonstrated on boxplots. For all indicators, the boxplots are available in Appendix E.

4.2.2 Computation of the Cluster Analysis

Before the actual computation of the cluster analysis, it is necessary to test potentially high correlation among the indicators. A high correlation could undermine the credibility of the model as the highly correlated variables are not independent. If such a situation occurs, it is necessary to make adjustments to the model. Because the data are in different units, the first step is to standardise data using a z-score.

Table 28: Correlation Matrix for Indicators

	X1	X2	X3	X4	X5	X6	X7
X1	1						
X2	-0.2895	1					
X3	0.8755	-0.2610	1				
X4	0.5547	-0.0742	0.4064	1			
X5	0.6014	-0.3017	0.6154	0.4398	1		
X6	0.4878	-0.2788	0.2719	0.6166	0.4431	1	
X7	0.4299	0.0686	0.2997	0.9662	0.3424	0.5459	1

Source: own elaboration in MS Excel

For the selected indicators, the high correlation occurs in two cases. The first correlation appears between X1 (GDP per capita) and X3 (net disposable income of households per capita). There is a logical reason for this because the economic performance expressed as GDP per capita has a significant impact on the income of households. For this reason, the solution of the high correlation is to omit the X3 and use only X1. A similar case is between X4 (population with tertiary education) and X7 (share of population with tertiary education on employment). An explanation is offered here that a higher proportion of people with tertiary education also increases their participation on employment. Even in this case, the solution is to use only one variable, namely X4. To sum up, five indicators are used for the cluster analysis: X1, X2, X4, X5, X6.

The next step is the calculation of the cluster analysis itself. As was already mentioned, there are 226 NUTS 2 regions included in the cluster analysis characterised by five indicators. Firstly, the distance matrix is calculated using the Euclidean distance. The matrix is a square matrix of the size 226×226. The next step is to create clusters using Ward's method. The procedure is described using the tree diagram where the horizontal axis represents the distances where the clusters were created, and the vertical axis stands for the regions. Because the amount of included regions is high and the graphical possibilities are limited, only a simplified version of the tree diagram is used. The full version can be found in Appendix F.

Image 2: Simplified Tree Diagram

Tree Diagram for 226 Cases
Ward's method
Euclidean distances
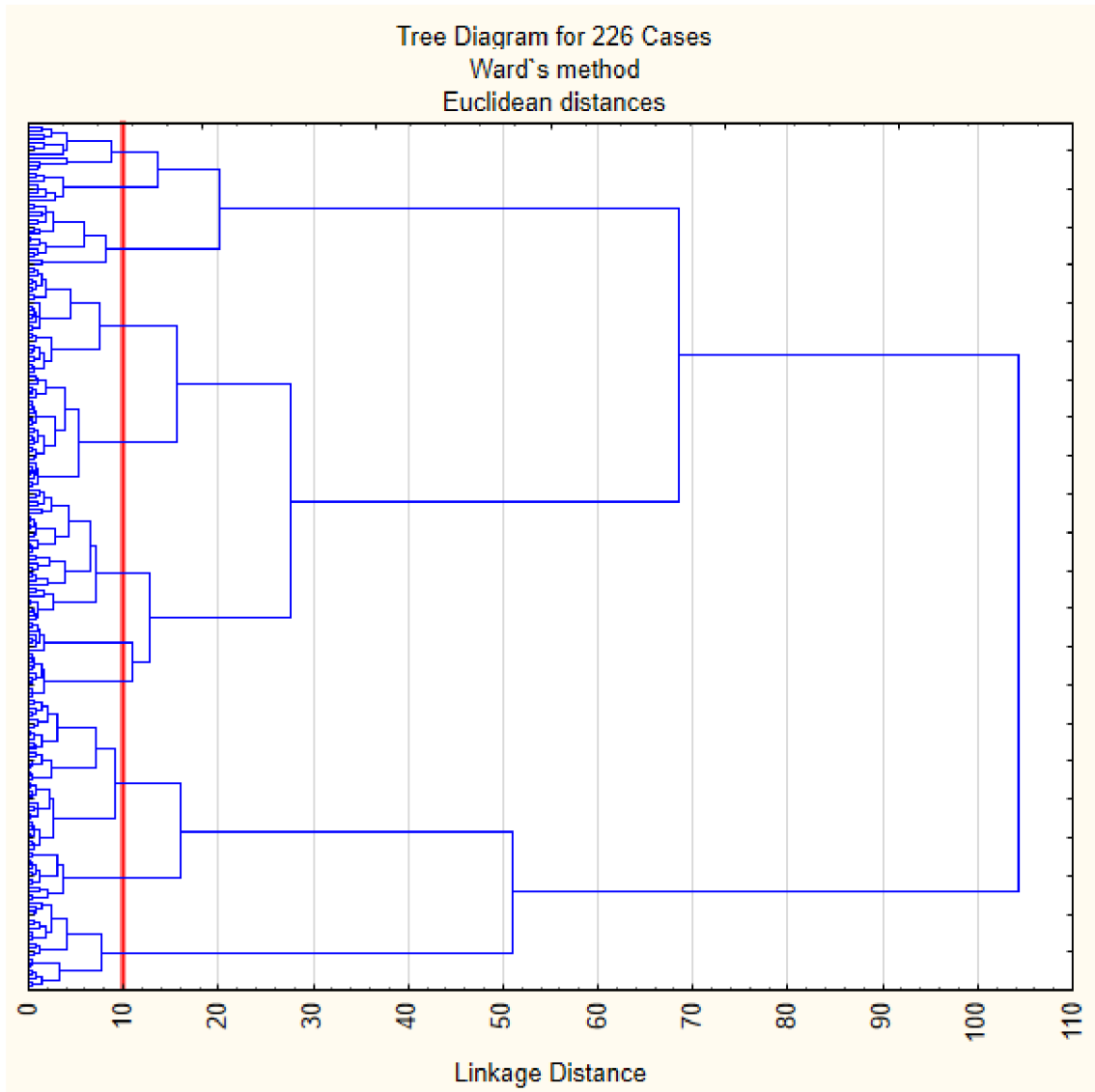
Source: own elaboration in Statistica 14

As seen in the graph, the distance of 10 is decided to be the threshold for the creation of clusters. To determine what number of clusters should be applied, the graph for the amalgamation schedule is used. From the graph, it is possible to deduce that the first sharp rise occurs at step 216. The step corresponds with the linkage distance of 10, which is decided to be the threshold.
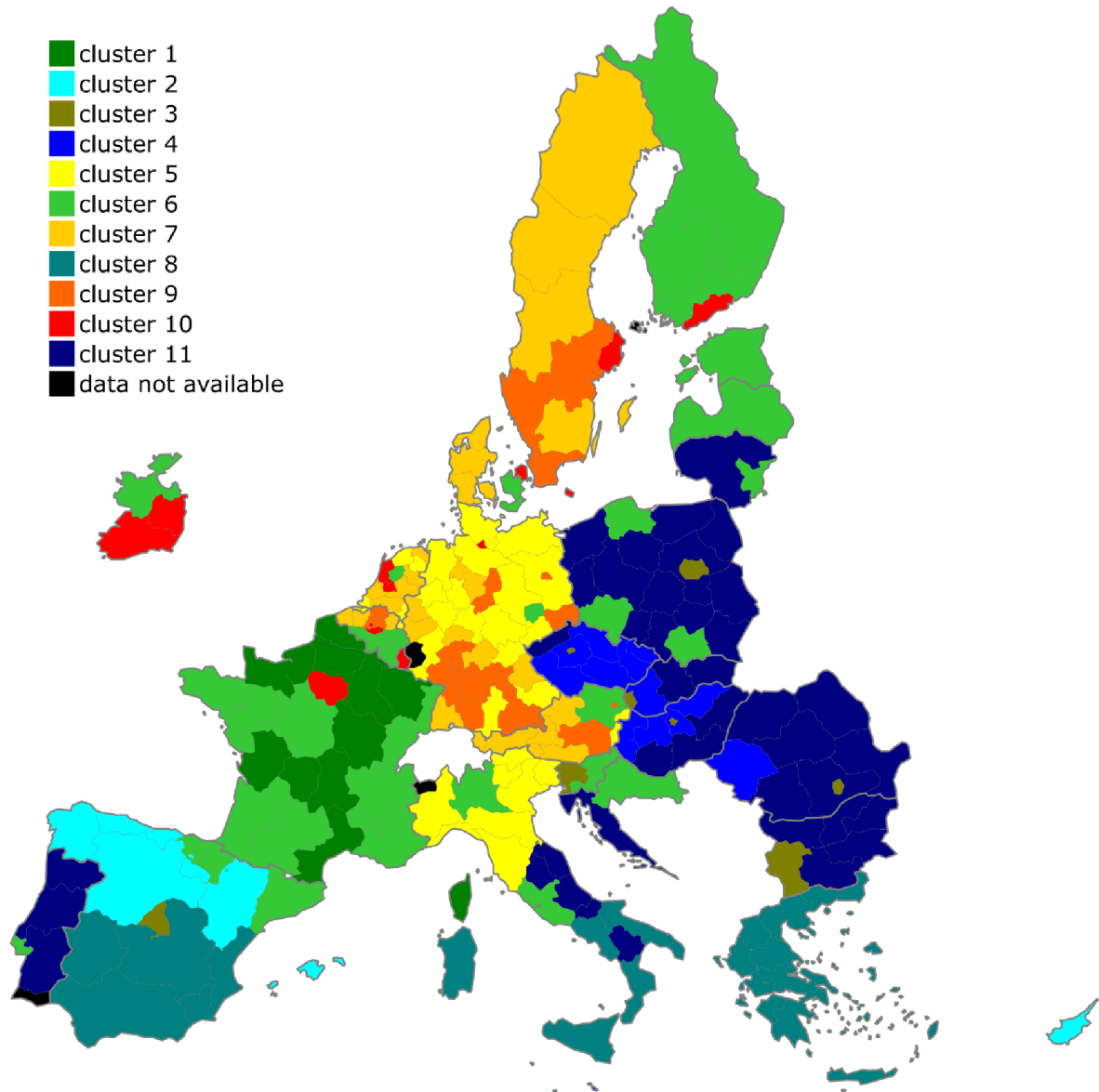
Image 3: Amalgamation Schedule Graph
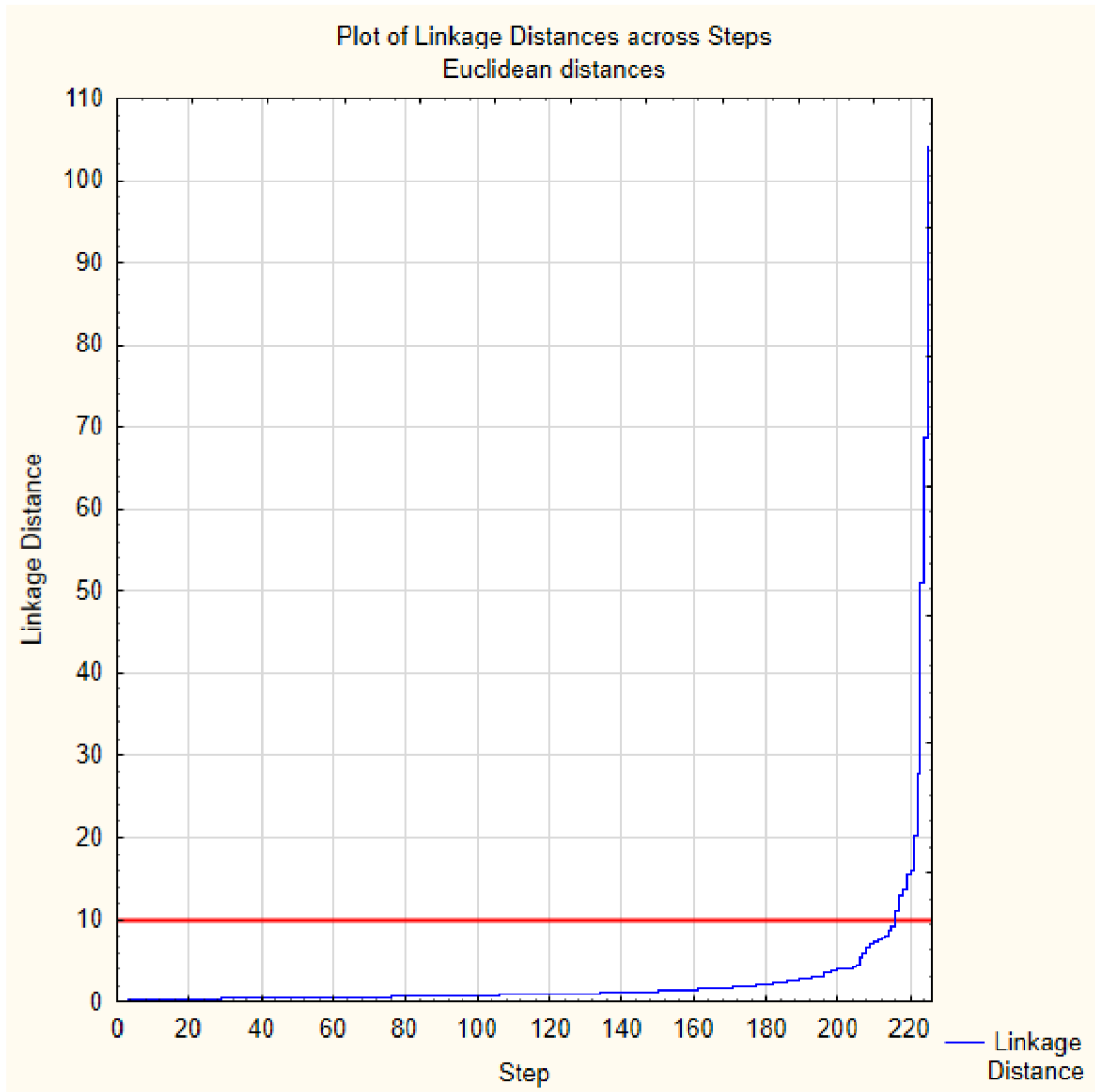
Source: own elaboration in Statistica 14

As it was already mentioned, the high number of NUTS 2 regions involved does not allow to analyse clusters according to the tree diagram. For better clarity, the cartogram is created to enable easy orientation in the composition of clusters.

Image 4: Cartogram of Clusters

Source: own elaboration in Inkscape 1.1.2

The colour scale is divided according to the GDP per capita. Red regions have the highest average GDP per capita, whereas the dark blue regions have the lowest average GDP per capita. However, the colour scale does not necessarily describe which regions are developing better in terms of total convergence. This is the subject for the following interpretation of the cluster analysis (chapter 4.2.3). The overall model shows the general tendencies among regions in the EU. It is possible to observe differences among regions of the former Western and Eastern Block as well as common characteristics of regions of southern Europe which were considerably poorer in the past.

4.2.3 Interpretation of the Cluster Analysis

For the purpose of interpreting the Cluster analysis, the following table is created. The table takes indicator values from each cluster and presents the arithmetic mean of these values:

Table 29: Average Values for Clusters

cluster	X1	X2	X4	X5	X6
1	27,010	9.6	29.0	2.21	2.2
2	24,250	13.5	37.8	0.80	2.5
3	28,175	5.1	43.2	1.88	8.7
4	14,414	3.8	21.0	1.14	4.7
5	32,947	4.9	23.8	1.62	2.7
6	27,205	8.1	35.0	1.72	4.2
7	40,778	4.5	32.0	2.18	3.5
8	16,388	20.8	24.3	0.84	2.3
9	43,772	4.9	35.6	4.45	5.3
10	64,253	7.0	47.4	2.78	7.2
11	11,561	7.4	22.0	0.63	2.0

Source: own elaboration in MS Excel

For better orientation, also a table containing standardized values is created. The values are standardized using a z-score.

Table 30: Standardized Average Values for Clusters

cluster	X1	X2	X4	X5	X6
1	-0.08	0.30	-0.08	0.41	-0.75
2	-0.26	1.03	0.89	-0.75	-0.61
3	0.00	-0.55	1.49	0.14	2.71
4	-0.93	-0.79	-0.96	-0.47	0.56
5	0.32	-0.59	-0.65	-0.07	-0.47
6	-0.07	0.01	0.58	0.01	0.34
7	0.85	-0.65	0.25	0.39	-0.05
8	-0.79	2.40	-0.60	-0.72	-0.71
9	1.05	-0.59	0.65	2.26	0.88
10	2.43	-0.19	1.95	0.88	1.92
11	-1.12	-0.11	-0.85	-0.89	-0.83

Source: own elaboration in MS Excel

From this table, it is also possible to create a graphical interpretation. However, it is necessary to mention one specific issue. The majority of the indicator values could be evaluated as positive if they reach higher values. This applies to GDP per capita, population with tertiary education, GERD, and employment in technology and knowledge-intensive sectors. However, the unemployment rate represents the opposite position. In this case, the

68

aim is to achieve the lowest possible values. To avoid confusion, the X2 values from Table 30 are multiplied by −1 for the purposes of Figure 9. It means that high unemployment (negative phenomenon) causes the bar chart location in negative values, whereas low unemployment (positive phenomenon) places the bar chart into positive values.

Figure 9: Graphical Interpretation of Standardized Average Values for Clusters
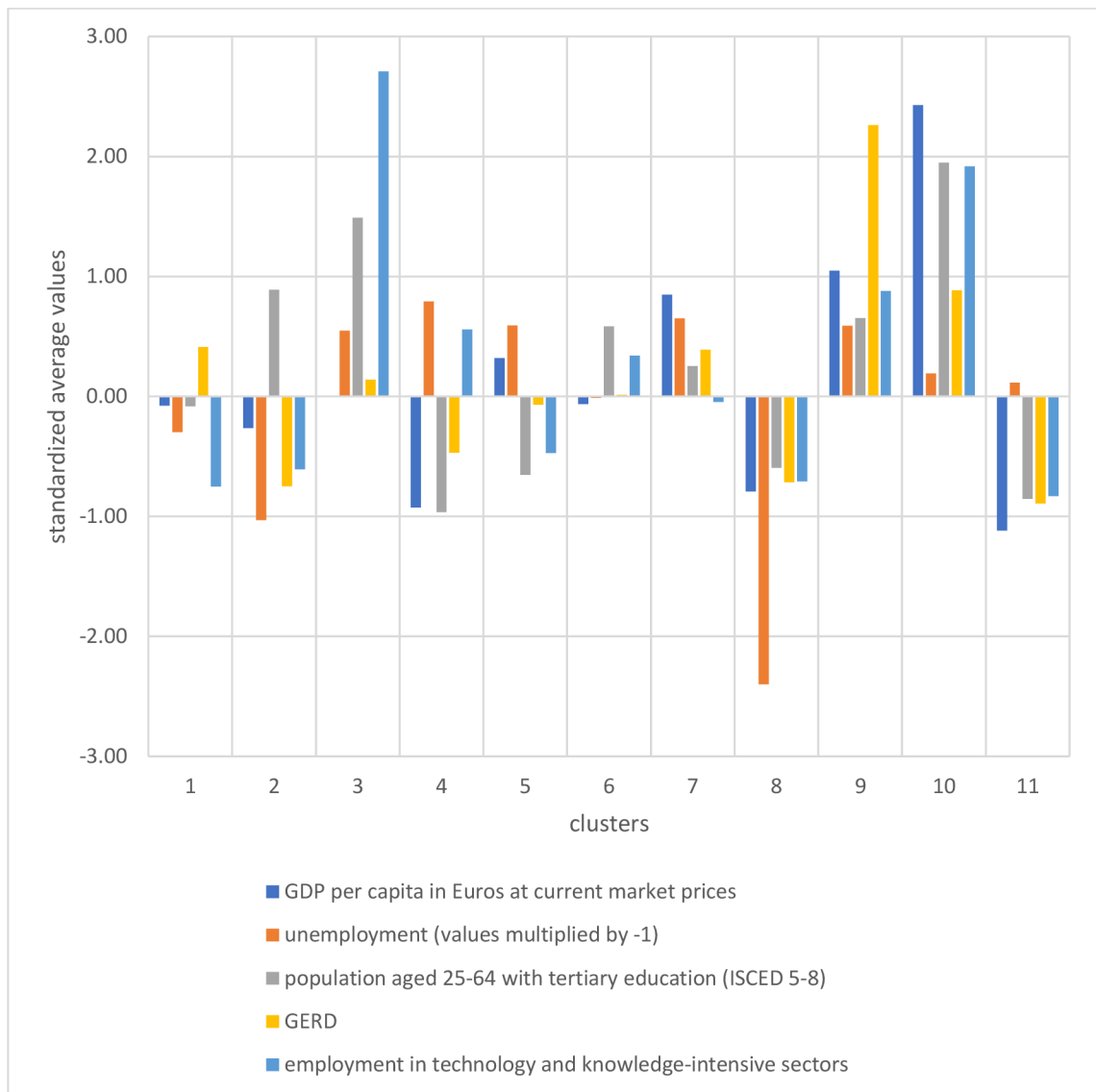
Source: own elaboration in MS Excel

Using the values from Table 29 (respectively Table 30) and Figure 9, it is possible to summarise the characteristics of each cluster.

4.2.4 Interpretation of Individual Clusters

Cluster 1

- Average GDP per capita
- Higher unemployment
- Average population with tertiary education
- Slightly above-average GERD
- Below-average employment in technology and knowledge-intensive sectors

Cluster 1 consists purely of French NUTS 2 regions. These are twelve regions that do not report significantly higher or lower values in any of the indicators compared to the overall average. The only exception is the employment in technology and knowledge-intensive sectors, which is the second lowest.

Cluster 2

- Slightly below-average GDP per capita
- Very high unemployment
- Above-average population with tertiary education
- Below-average GERD
- Below-average employment in technology and knowledge-intensive sectors

In cluster 2, there are regions of northern Spain and Cyprus. This cluster includes areas in the relatively poorer south of Europe. Although the GDP per capita is not significantly lower than the EU average, high unemployment is evident.

Cluster 3

- Average GDP per capita
- Low unemployment
- Significantly above-average population with tertiary education
- Average GERD
- The highest employment in technology and knowledge-intensive sectors

Cluster 3 is specific because it only involves capital cities or regions in which the capital city is located. It includes the capitals of the Czech Republic, Bulgaria, Hungary, Poland, Romania, Slovakia, and Slovenia – capital cities of countries from the former Eastern Bloc. Additionally, the region, which includes the capital of Spain, is also part of the cluster. In general, the cluster can be characterised as a set of capital cities with very good economic conditions, which are, however, located in countries with lower economic performance or – in the case of Spain – a country with specific problems (high unemployment).

Cluster 4

- Below-average GDP per capita
- The lowest unemployment
- The lowest population with tertiary education
- Slightly below-average GERD
- Slightly above-average employment in technology and knowledge-intensive sectors

In this cluster, there is the majority of Czech regions, together with regions of northern Hungary and one region from Slovakia and Romania. These are, therefore, purely regions of the former Eastern Bloc. The regions have relatively low GDP per capita and the lowest population with tertiary education. Despite all of this, they report the lowest unemployment rate.

Cluster 5

- Slightly above-average GDP per capita
- Low unemployment
- Below-average population with tertiary education
- Average GERD
- Slightly below-average employment in technology and knowledge-intensive sectors

The majority of regions of cluster 5 come from Germany and the northern part of Italy. To them, one region from Austria and four regions from the Netherlands are added. The regions report low unemployment and a low proportion of the population with tertiary education. The rest of the indicators are not considerably far from the average.

Cluster 6

- Average GDP per capita
- Average unemployment
- Above-average population with tertiary education
- Average GERD
- Slightly above-average employment in technology and knowledge-intensive sectors

Cluster 6 is the most heterogeneous one in terms of geographical location. There are regions both from the West and East included. Therefore, it is not possible to conclusively characterise the cluster. The included regions are characterised only by a higher proportion of the population with tertiary education. The other indicators are not significantly different from the average.

Cluster 7

- Above-average GDP per capita
- Low unemployment
- Slightly above-average population with tertiary education
- Slightly above-average GERD
- Average employment in technology and knowledge-intensive sectors

Regions involved in this cluster come from the following countries: Austria, Belgium, Germany, Denmark, Netherlands, and Sweden. The regions are characterised by the combination of higher average GDP per capita and lower average unemployment. The rest of the indicators report values close to the average.

Cluster 8

- Below-average GDP per capita
- The lowest unemployment
- Below-average population with tertiary education
- Below-average GERD
- Below-average employment in technology and knowledge-intensive sectors

Cluster 8 is the only one whose average indicator values are all below the average of the EU. Moreover, the unemployment in these regions is by far the highest. In the cluster, there are regions of southern Spain and Italy, and the whole of Greece. It points to the problem of states, which are on one side former countries of the Western Bloc, but on the other side, part of the poorer south of Europe. This problem is characteristic of the fact that it lasts for several decades.

Cluster 9

- Above-average GDP per capita
- Low unemployment
- Above-average population with tertiary education
- The highest GERD
- Above-average employment in technology and knowledge-intensive sectors

Considering all indicators, cluster 9 comprises highly developed regions that report the second highest average GDP per capita. A notable characteristic of this cluster is by far the highest GERD. Geographically it is similar to cluster 7. The only difference is that there is no region from Denmark and the Netherlands.

Cluster 10

- Highest GDP per capita
- Average unemployment
- The highest population with tertiary education
- Above-average GERD
- Significantly above-average employment in technology and knowledge-intensive sectors

Regions of cluster 10 are the richest ones in terms of GDP. These are mainly the capitals of countries from Western and Northern Europe or their surroundings. The exception is in Germany, where Hamburg is involved instead of Berlin. Together with cluster 9, it is the only cluster with all indicator values above the EU average. It is, however, necessary to mention that unemployment is not significantly low.

Cluster 11

- Lowest GDP per capita

- Average unemployment

- Below-average population with tertiary education

- The lowest GERD

- The lowest employment in technology and knowledge-intensive sectors

Cluster 11 is predominant in countries of the former Eastern Bloc to which most Portuguese regions are added. Indicator values in this cluster are generally below the average of the whole EU. Moreover, regions involved in this cluster have the lowest GERD and the lowest employment in technology and knowledge-intensive sectors.

5 Results and Recommendations

The result of the convergence analysis is that the general tendency of the regions is to converge. This could basically serve as proof of the theories mentioned in the theoretical part. As it was said, the predominant regional development theories assume a general tendency to convergence. However, it is also necessary to mention that the tendencies in the EU are considerably affected by the Cohesion Policy. With a high degree of probability, the tendency to convergence would be substantially lower without the EU and its Cohesion Policy.

Although the computations of beta-convergence and sigma-convergence find the tendencies towards convergence, the convergence among member states is more verifiable than among the NUTS 2 regions. From one point of view, it is natural because the heterogeneity among 240 units should be higher than among 27 units. On the other hand, this fact points out that there are still persisting barriers towards definite regional convergence.

One of the reasons for this imbalance between the convergence among states and the convergence among regions is the concentration of economic activities in richer regions determined by the agglomeration forces. This could be shown in an example of cluster 3. For instance, one of the regions from this cluster – Prague (code CZ01) – is the capital city of the Czech Republic. Its membership in cluster 3 confirms its difference from other Czech regions. At the same time, Figure 6 indicates that Prague may diverge from the other regions as it reported high initial GDP per capita and high average growth of it. It causes a situation when a rich region affects the economic condition of the whole state. This could also contribute to the development of the rest of the region; however, poorer regions' development is limited.

Another example of this imbalance are the so-called Baltic states, where the cluster analysis encounters the limits of the NUTS methodology. Whereas Latvia and Estonia are put in cluster 6 as a whole, only the surrounding of the Lithuanian capital is involved in this cluster. The rest of the country is put into cluster 11. This is, however, caused due to the fact that Lithuania is the only Baltic state with more than one NUTS 2 region. The intrastate disparities in Latvia and Estonia cannot be found on the NUTS 2 level; the only solution would be to go to the NUTS 3 level. Nevertheless, this is not possible as the methodology

of the EU needs a unified approach, involving the decision of which level of regions should be taken as the main one for the implementation of the Cohesion Policy.

Another problem identified is the persistence of long-term disparities concerning geographical location. The cluster analysis detects a worse economic situation in the south of Europe. This is, however, a problem that the Cohesion Policy has dealt with since the 1980s when Greece, Portugal, and Spain joined the community.

The differences are also evident between the western and eastern parts of the EU. Although the Cohesion Policy should support the connection among regions no matter in which country they are located, the national dimension still prevails. According to the cluster analysis, there are rather diverse clusters on different sides of the east-west border. This might also be caused by the cross-border cooperation problems of connection. There are several barriers, such as legislation, administration, or language. The recommendation concerning this problem should be that the EU should accentuate more the regional dimension.

6 Conclusion

The cooperation of European countries after WWII helped to break out of the consequences of the wort war conflict in the history of humankind and enabled the connection of the states in Europe. This also continued after the end of the Cold War, and the process was unique because it was able to integrate states of different cultures and different development. Nevertheless, the different development caused considerable disparities, mainly among the regions. During the enlargements, the European Communities and subsequently the EU implemented policies based on a regional approach aimed to reduce these disparities.

The result of this thesis confirms that the general tendency of the regions in the period 2015-2019 was to converge in terms of GDP per capita. However, there are still persisting disparities caused by different development of the regions and by specific problems, e.g., high unemployment in Southern Europe. As the cluster analysis shows, the position of the regions in clusters is highly affected by the state in which they are located. For the future development of the Cohesion Policy, it might be beneficial to focus even more on the regional dimension and interregional cooperation.

7 References

ABU-JAMOUS, B., FA, R., NANDI, A. K. 2015. *Integrative Cluster Analysis in Bioinformatics*. Chichester: John Wiley & Sons. 454 p. ISBN 978-1-118-90653-8.

AGGARWAL, C. C., REDDY, C. K. 2013. *Data Clustering: Algorithms and Applications*. Boca Raton: CRC Press. 653 p. ISBN 978-1-4665-5821-2.

BACHE, I. 2007. *Europeanization and Multilevel Governance: Cohesion Policy in the European Union and Britain*. Lanham: Rowman & Littlefield Publishers. 180 p. ISBN 978-0-7425-7911-8.

BALDWIN R., WYPLOSZ C. 2012. *The Economics of European Integration*. 4th ed. Maidenhead: McGraw-Hill Education. 584 p. ISBN 9780077131722.

BANNOCK, G., BAXTER, R. E. DAVIS, E. 2003. *The Penguin Dictionary of Economics*. 7th ed. London: Penguin Books. 416 p. ISBN 0-14-101075-4.

BENEDETTO, G. *The History of the EU Budget: IN-DEPTH ANALYSIS* [online]. 2019. (PDF). https://www.europarl.europa.eu/RegData/etudes/IDAN/2019/636475/IPOL_IDA(2019)636 475_EN.pdf. Accessed 27 February 2022.

BLAŽEK, J., UHLÍŘ, J. 2011. *Teorie regionálního rozvoje: nástin, kritika, implikace*. 2nd ed. Praha: Karolinum. 344 p. ISBN 978-80-246-1974-3.

CVCE. *Treaty establishing the European Economic Community* [online]. https://www.cvce.eu/en/obj/treaty_establishing_the_european_economic_community_rom e_25_march_1957-en-cca6ba28-0bf3-4ce6-8a76-6b0b3252696e.html. Accessed 26 January 2022.

CZSO. *Regional breakdown* [online]. https://www.czso.cz/csu/czso/regional_breakdown. Accessed 26 January 2022.

DOTACEEU.CZ. *Evropské strukturální a investiční fondy* [online]. https://www.dotaceeu.cz/cs/informace-a-dokumenty/slovnik-pojmu/e/evropske-strukturalni-a-investicni-fondy-(esi-fond. Accessed 27 February 2022.

EPRS. *The Cost of Non-Europe in the Single Market: I - Free Movement of Goods* [online]. 2014. (PDF). https://www.europarl.europa.eu/EPRS/EPRS_STUDY_536353_CoNE_Single_Market_I.p df. Accessed 27 February 2022.

EUR-LEX. *Council Regulation (EEC) No 2052/88 of 24 June 1988 on the tasks of the Structural Funds and their effectiveness and on coordination of their activities between themselves and with the operations of the European Investment Bank and the other existing financial instruments* [online]. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A31988R2052&qid=1648584015002. Accessed 27 February 2022a.

EUR-LEX. *Regulation (EU, Euratom) 2018/1046 of the European Parliament and of the Council of 18 July 2018 on the financial rules applicable to the general budget of the Union, amending Regulations (EU) No 1296/2013, (EU) No 1301/2013, (EU) No 1303/2013, (EU) No 1304/2013, (EU) No 1309/2013, (EU) No 1316/2013, (EU) No 223/2014, (EU) No 283/2014, and Decision No 541/2014/EU and repealing Regulation (EU, Euratom) No 966/2012* [online]. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018R1046&qid=1648584328269. Accessed 27 February 2022b.

EUR-LEX. *Single European Act* [online]. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A11986U%2FTXT. Accessed 27 February 2022c.

EUROPEAN COMMISSION. *An introduction to EU Cohesion Policy 2014-2020* [online]. 2014. (PDF). https://ec.europa.eu/regional_policy/sources/docgener/informat/basic/basic_2014_en.pdf. Accessed 27 February 2022.

EUROPEAN COMMISSION. *Cohesion policy 2007-13: Commentaries and official texts* [online]. 2007. (PDF). https://ec.europa.eu/regional_policy/sources/docoffic/official/regulation/pdf/2007/publicati ons/guide2007_en.pdf. Accessed 27 February 2022.

EUROPEAN COMMISSION. *Cohesion Policy 2021-2027* [online]. https://ec.europa.eu/regional_policy/en/2021_2027/#4. Accessed 27 February 2022a.

EUROPEAN COMMISSION. *Just Transition Fund* [online]. https://ec.europa.eu/regional_policy/en/funding/jtf/. Accessed 27 February 2022b.

EUROPEAN COMMISSION. *Regional Policy: Glossary* [online]. https://ec.europa.eu/regional_policy/en/policy/what/glossary. Accessed 27 February 2022c.

EUROPEAN COMMISSION. *Regional Policy: Priorities for 2021-2027* [online]. https://ec.europa.eu/regional_policy/en/policy/how/priorities. Accessed 27 February 2022d.

EUROPEAN COMMISSION. *The EU's 2021-2027 long-term budget and NextGenerationEU: facts and figures* [online]. 2021. (PDF). https://data.europa.eu/doi/10.2761/91357. Accessed 27 February 2022.

EUROPEAN PARLIAMENT. *Multiannual financial framework* [online]. https://www.europarl.europa.eu/factsheets/en/sheet/29/multiannual-financial-framework. Accessed 27 February 2022.

EUROSKOP.CZ. *Předvstupní pomoc* [online]. https://euroskop.cz/evropska-unie/cr-a-eu/vstup-cr-do-eu/predvstupni-pomoc/. Accessed 27 February 2022.

EUROSTAT. *Regional statistics by NUTS classification* [online]. https://ec.europa.eu/eurostat/web/regions/data/database. Accessed 20 March 2022.

EUROSTAT. *Regions in the European Union* [online]. 2018. (PDF). https://ec.europa.eu/eurostat/documents/3859598/9397402/KS-GQ-18-007-EN-N.pdf. Accessed 27 February 2022.

GREGORY, D. et al. 2009. *The Dictionary of Human Geography*. 5th ed. Chichester: John Wiley & Sons. 1072 p. ISBN 978-1-4051-3288-6.

HEBÁK, P. et al. 2015. *Statistické myšlení a nástroje analýzy dat*. 2nd ed. Praha: Informatorium. 880 p. ISBN 978-80-7333-118-4.

LIPI, A., KUMAR, K., CHAKRABORTY, S. 2022. *Outliers: Detection and Analysis*. New York: Nova Science Publishers. 154 p. ISBN 978-1-68507-554-5.

MARTINEZ, W. L., MARTINEZ, A. R., SOLKA, J. L. 2017. *Exploratory Data Analysis with MATLAB*. 3rd ed. Boca Raton: CRC Press. 625 p. ISBN 978-1-4987-7606-6.

MINAŘÍK, B., BORŮVKOVÁ, J., VYSTRČIL, M. 2013. *Analýzy v regionálním rozvoji*. Průhonice: Professional Publishing. 244 p. ISBN 978-80-7431-129-1.

MYATT, G. J., JOHNSON, W. P. 2014. *Making Sense of Data I: A Practical Guide to Exploratory Data Analysis and Data Mining*. 2nd ed. Hoboken: John Wiley & Sons. 250 p. ISBN 978-1-118-40741-7.

OECD. *Regional Development Policy* [online]. https://www.oecd.org/regional/regional-policy/regionaldevelopment.htm. Accessed 27 February 2022.

SENIOR NELLO, S. 2012. *The European Union: Economics, Policies and History*. 3rd ed. Maidenhead: McGraw-Hill Education. 502 p. ISBN 9780077129668.

STEJSKAL, J., KOVÁRNÍK, J. 2009. *Regionální politika a její nástroje*. 1st ed. Praha: Portál. 216 p. ISBN 978-80-7367-588-2.

TEASDALE, A., BAINBRIDGE, T. 2012. *The Penguin Companion to European Union*. 4th ed. London: Penguin Books. 866 p. ISBN 978-0-141-02118-8.

TOMŠÍK, K. 2009. *Vývoj a perspektivy evropského venkova*. 1st ed. Praha: Wolters Kluwer. 208 p. ISBN 978-80-7357-495-6.

VANHOVE, N. 2018. *Regional Policy: A European Approach*. 3rd ed. Abingdon: Routledge. 656 p. ISBN 978-1-138-33033-7.

VYSTRČIL, M., ŠULC, M. 2008. *Jak se žije na Vysočině*. Jihlava: Petros Martakidis. (42 p.). ISBN 978-80-254-3805-3.

WIERZCHOŃ, S. T., KŁOPOTEK M. A. 2018. *Modern Algorithms of Cluster Analysis*. Cham: Springer International Publishing. 433 p. ISBN 978-3-319-69307-1.

WOKOUN, R. et al. 2008. *Regionální rozvoj: východiska regionálního rozvoje, regionální politika, teorie, strategie a programování*. Praha: Linde. 475 p. ISBN 978-80-7201-699-0.

8 Appendices

Appendix A: GDP per Capita in Euros at Current Market Prices, EU Member States

Appendix B: GDP per Capita in Euros at Current Market Prices, EU NUTS 2 Regions

Appendix C: Division of NUTS 2 Regions Into Quadrants of the Correlation Diagram

Appendix D: Values of Individual Indicators for NUTS 2 Regions, Average 2015-2019

Appendix E: Boxplots of Summary Statistics

Appendix F: Tree Diagram

Appendix G: Division of NUTS 2 Regions Into Clusters

Appendix H: Cartograms of Individual Clusters