



**INSTYTUT BADAŃ SYSTEMOWYCH
POLSKIEJ AKADEMII NAUK**

TECHNIKI INFORMACYJNE TEORIA I ZASTOSOWANIA

Wybrane problemy
Tom 5 (17)

poprzednio

**ANALIZA SYSTEMOWA W FINANSACH
I ZARZĄDZANIU**

Pod redakcją
Andrzeja MYŚLIŃSKIEGO

Warszawa 2015



**INSTYTUT BADAŃ SYSTEMOWYCH
POLSKIEJ AKADEMII NAUK**

**TECHNIKI INFORMACYJNE
TEORIA I ZASTOSOWANIA**

Wybrane problemy
Tom 5 (17)

poprzednio

**ANALIZA SYSTEMOWA W FINANSACH
I ZARZĄDZANIU**

Pod redakcją
Andrzeja Myślińskiego

Warszawa 2015

Wykaz opiniodawców artykułów zamieszczonych
w niniejszym tomie:

Dr inż. Tatiana JAWORSKA

Dr inż. Jan OWSIŃSKI

Dr hab. inż. Andrzej MYŚLIŃSKI, prof. PAN

Prof. dr hab. inż. Piotr SIENKIEWICZ

Prof. dr hab. inż. Andrzej STRASZAK

Dr hab. Dominik ŚLĘZAK, prof. UW

Prof. dr hab. inż. Stanisław WALUKIEWICZ

Copyright © by Instytut Badań Systemowych PAN
Warszawa 2015

ISBN 83-894-7558-8

ECONOMIES OF EU COUNTRIES SEEN VIA GRADE DATA ANALYSIS

Stanisław Lenkiewicz

*Systems Research Institute, Polish Academy of Sciences
Ph.D. Studies, Warsaw, Poland
e-mail: stan@lenkiewicz.eu*

Abstract. In this paper we analyze the changes that took place in the European Union in the period 2004-2014, it is in a period of its intense expansion. Based on a set of eight characteristics, we split the EU into clusters of countries with similar profiles; we also split characteristics into clusters of linked features. The study covers 28 countries currently being members of the EU irrespective of when they have joined it. The research method we use is Grade Data Analysis, and our data processing tool is GradeStat program.

Keywords: European Union, economic condition, cluster analysis, Grade Data Analysis, overrepresentation, outlier

1 ECONOMIC CONDITION

The economic condition of a country may be understood as its position in relation to other countries in a specific period of time. It may be measured through use of statistics involving GDP information, unemployment rates, stock market data, and many other metrics.

This definition points to two major problems in assessing the economic condition. Firstly, there is no universal set of measures of economic health. Secondly, we cannot assess the economic condition of a certain country in isolation from other countries; economic evaluations almost always take the form of rankings or comparative analysis.

These factors make every assessment of the economic health more or less subjective. We do not deny that our analysis – an assessment of the economies of European Union countries over the past ten years – also has this disadvantage.

1.1 Macroeconomic Measures

By macroeconomic measures we understand indicators of the current status of particular areas of the economy (industry, labour market, trade, etc.).

They describe real national income, spending, and output. They indicate whether the economy is growing or in recession. They are published regularly by governmental or international agencies, such as Eurostat.

The most important macroeconomic indicators are:

- Gross Domestic Product (GDP). It represents the total market value of all goods and services produced in a country during a given year. There are two most commonly used indicators based on the GDP: GDP per capita and GDP growth rate.
- Inflation rate. It is a sustained increase in the general price level of goods and services in an economy over a period of time. It is the result of an imbalance between the amount of money circulating in the economy and the amount of goods available on the market.
- Unemployment rate. It is the percentage of the total labour force that is unemployed but actively seeking employment and willing to work.
- Financial indicators: balance of payments, public debt, savings ratio, etc.

1.2 Measures of Quality of Life

It is difficult to define the concept as abstract as quality of life. Usually it is understood as the general well-being of individuals and societies. There are many aspects of this "well-being", and some of them are closely related to macroeconomics. Thus, the division of economic measures into macroeconomic indicators and indicators of quality of life is somewhat artificial. Areas covered by quality of life are: political freedom, security, healthcare, education, access to technologies, employment, and many others.

Quality of life should not be confused with standard of living which is based on income. Standard of living refers to the level of wealth available to a certain society and may be considered as one of measures of quality of life (but not its synonym).

2 EUROPEAN UNION

The European Union (EU) was established on November 1st, 1993 by the Treaty of Maastricht. Its creation was the next step towards full political and economic integration of its members. EU is a successor of the European Economic Community founded on January 1st, 1958 by the Treaty of Rome.

The purpose of starting closer cooperation of European countries was to ensure their balanced development, to enable close contacts of European

nations, and as a consequence: to prevent wars in Europe. A "triggering factor" of European integration were tragic experiences of World War II.

For half-century members of the EU enjoy peace and a relatively high standard of living. For this reason, the EU is often perceived as a "land of prosperity", a club of rich countries with no problems. Indeed, EU members are relatively affluent, but the EU is not monolithic; it includes both: rich countries and those who have just started to build their wealth. The year 2004 began the process of EU enlargement, and this process is still ongoing. Accession to the EU has changed not only the situation of new Member States, but also that of "old" ones, and of the Union as a whole.

Since 2007 the EU is facing one of the most difficult periods in its history. The crisis in the real estate market, whose first signals came from the US, began to spread in other countries, soon affecting the majority of EU members. Problems of banks forced to intervene both governments and central banks of many countries, as well as EU central institutions. In 2009 the new Greek government announced that the economic statistics published by the previous governments were falsified, and the country has long been in a recession. Soon serious economic problems hit Spain and Ireland. The crisis – by way of a domino effect – spilled across the EU. It turned out that the "club of the rich" may also have problems.

In this paper we are going to try to assess the evolution of the EU in the period 2004-2014. There are many studies on this issue; however, very few of them use data mining tools, including cluster analysis. We want to fill this gap – we will look at the European Union through the prism of Grade Data Analysis.

3 GRADE DATA ANALYSIS REVISITED

This paper is focused on Grade Data Analysis (GDA) applications, not on the method itself; an in-depth description of GDA may be found in [5] and [12]. However, for clarity of subsequent sections, it is necessary to provide a brief summary of the methodology we are going to use. We have already presented it in details in [15] and [16]; here we will make a short *résumé*.

Grade Data Analysis (GDA) is an advanced data analysis technique for ordering both objects and their features.¹ Its starting point is a table having m rows and n columns; it presents m objects (rows) having n characteristics (columns). All values in such a table must be non-negative; if there are any negative values, they must be eliminated. Data sets for all objects

¹A comprehensive presentation of cluster analysis may be found in [1], [2], and [17].

must be complete; missing values must be supplemented. The sum of each column of the table must be greater than zero.

Table 1 on page 47 contains 8 characteristics of 28 members of the European Union. Values marked in bold were missing, and we had to supplement them (the procedure will be described in Section 4.4). This table will be used for an analysis of 28 countries (present and future EU members) in 2004. However, before we perform it, we must normalise the data.

Data normalisation in GDA is performed in two steps. In the first step each value is divided by the sum of its group. A group may be composed of one or more columns. After this operation all values in the table belong to the interval $[0; 1]$. This prevents the analysis from being dominated by one (or some) of characteristics of objects (in our case these would be: GDP and Debt).²

Each group of characteristics is assigned a weight which reflects its "importance" in the analysis. Then each value in the table is multiplied by the weight of its group and divided by the sum of the whole table – this is the second step of the normalisation.

We decided to put Life F and Life M³ in one group, and to treat all other columns as individual groups. We assigned a weight equal to 1.5 to Debt and to all other characteristics: weights equal to 1. Table 2 on page 48 presents normalised values.

Such a pre-processed table is an input of Grade Correspondence Analysis (GCA) algorithm. In the first step GCA counts for each value its overrepresentation index c_{ij}

$$c_{ij} = \frac{p_{ij}}{p_{i+} \times p_{+j}}, \quad (1)$$

where p_{ij} is the value in i -th row and j -th column, p_{i+} is the sum of i -th row, and p_{+j} is the sum of j -th column. Table 3 on page 49 presents overrepresentation indexes for 2004.

Overrepresentation indexes for each value may be presented in a form of an overrepresentation map. Such a map is a square divided into rows representing objects and columns representing characteristics. Heights of rows and widths of columns are proportional to sums of rows and sums of columns of the table of normalised values (see Table 2 on page 48) and colours of cells represent overrepresentation indexes of each characteristics

²GDP – Gross Domestic Product, Debt – public debt. All characteristics will be described later on.

³Life expectancy at birth for women and for men.

in each object. Fig. 2 on page 50 presents an overrepresentation map for 28 countries in 2004, and Fig. 3 explains the meanings of colours.

In its second step GCA finds such a permutation of rows and columns of the overrepresentation map which maximises Spearman's rank correlation coefficient ρ^*

$$\rho^* = 3 \sum_{i=1}^m \sum_{j=1}^n (p_{ij} (2S_{row}(i) - 1) (2S_{col}(j) - 1)), \quad (2)$$

where

$$S_{row}(i) = \left(\sum_{s=1}^{i-1} p_{s+} \right) + \frac{1}{2} p_{i+}, \quad (3)$$

$$S_{col}(j) = \left(\sum_{t=1}^{j-1} p_{+t} \right) + \frac{1}{2} p_{+j}. \quad (4)$$

After GCA we receive an overrepresentation map which is more regular; its segments with have the same degree of grey form tight areas, and the darkest ones are arranged close to the diagonal. Rows and columns which are more similar become adjacent. Such a structure of the map permits us to perform a cluster analysis – both for rows and columns; we can find groups of similar objects as well as groups of interrelated attributes. Fig. 4 on page 51 presents an overrepresentation map after GCA and the cluster analysis for 2004.⁴

GCA allows also to detect outliers, i.e. objects which are "untypical". These outliers differ from those detected by standard statistical methods. Traditional techniques usually consider as outliers those of them which have extremely low or extremely high levels of one characteristics. GCA detects outliers based on the entire set of characteristics; it counts for each row a value denoted as $AvgDistA_{row}$ which may be considered as an average distance of this row from all other rows.⁵ Objects (rows of the table) which have $AvgDistA_{row}$ much greater than others are considered as outliers. Fig. 6 on page 56 presents a graph for values $AvgDistA_{row}$ in 2004. We may notice that the most outlying country is Luxembourg.

⁴An overrepresentation map is based on overrepresentation indexes, not on values of characteristics. This is why the cell representing Life M for Slovenia is darker than the one for Sweden, though life expectancy of men in Sweden is higher than in Slovenia.

⁵A comprehensive explanation of the method for calculating this value can be found in [12].

GCA allows to perform a similar analysis for characteristics. It counts for each column a value $AvgDistA_{col}$ which is used to detect outliers among characteristics. Fig. 7 on page 57 presents a graph for values $AvgDistA_{col}$ in 2004. We may notice that the most outlying characteristic is public debt.

4 EU COUNTRIES. GDA APPLIED

As we noted in Section 2, for a few years the European Union has been going through one of the most difficult periods in its history. High unemployment, rising public debt and political instability have become reality of many EU members. The most glaring example of this is Greece, which – according to many analysts – is bankrupt.

Many academics and policy-makers say that this breakdown of the "European prosperity" could have been foreseen. They argue that alarm signals were noticeable for a long time, but have been ignored both by decision-making bodies of the Union, and by the governments of Member States. When asked for examples of those signals, they usually point at rapidly growing public debt, stagnation or decrease in GDP and rising unemployment.

It is obvious that these phenomena, characteristic for weak economies, should have been a warning of upcoming problems. However, it should be noted that a temporary rise in unemployment or short-term economic stagnation periods occur even in countries with very strong economies. Thus, when analysing the economic situation of a country, one should take into account the coexistence of these factors – as they are interrelated – and the period in which they can be observed. In such a research Grade Data Analysis seems to be a suitable tool.

In this section we will explore the economic condition of Member States of the European Union in the period 2004-2014. Our study will include all 28 countries currently making up the EU regardless of when they acceded to it. This approach may be surprising; however, each new member brought to the Union all its problems, contributing in this way to the troubles of the whole community. Thus, we believe to be reasonable to take into account also those countries, which – during the period covered by the analysis – were still negotiating their accession.

4.1 Disclaimer

In [15] we considered whether there is an objective assessment and we came to the conclusion that it does not exist. Each assessment is done by

humans who, by their very nature, are subjective; thus, each assessment is marked by the preferences and beliefs of those who perform it.

In GDA we can distinguish two areas, where the impact of subjectivism of researchers is particularly evident. These are: the choice of characteristics of objects, and the assignment of weights to groups into which these characteristics have been divided. This will also take place in our study. Though we will try to keep a maximum objectivity, we do not claim that the result of our research will be completely free from the influence of our convictions; it would be unfair on our part. The outcome of our analysis can be considered valuable insofar as its assumptions are correct. Anyway – this is a feature of most studies.

We believe that a fundamental feature of each researcher should be honesty. Thus, we honestly admit that our research is performed from the perspective of those, who – while appreciating the benefits of European integration – note many mistakes made by the EU in economic and social areas. However, our aim is not to make a final assessment of the phenomena that have occurred over the past ten years; we just want our study to be our voice in the debate on the future of the European Union. We think that in a period of crisis affecting many Member States such a discussion is necessary.

4.2 Characteristics of Countries

We will analyse the following characteristics of each country (please see Section 1):

Macroeconomic measures:

- Gross Domestic Product (GDP)⁶ per capita (in PPS)⁷ – code: GDP,
- real GDP growth rate (percentage change with respect to the previous year) – code: GDP +,
- inflation rate (annual average rate of change in %) – code: Inflation,
- general government gross debt⁸ (percentage of GDP) – code: Debt,

⁶The Organization for Economic Co-operation and Development (OECD) defines the GDP as *an aggregate measure of production equal to the sum of the gross values added of all resident institutional units engaged in production (plus any taxes, and minus any subsidies, on products not included in the value of their outputs). The sum of the final uses of goods and services (all uses except intermediate consumption) measured in purchasers' prices, less the value of imports of goods and services, or the sum of primary incomes distributed by resident producer units.*

⁷Purchasing Power Standard. An artificial currency unit which eliminates differences in price levels between countries. Theoretically, one PPS can buy the same amount of goods and services in each country.

⁸In the Maastricht Treaty it is defined as *consolidated general government gross debt at nominal (face) value, outstanding at the end of the year in the following categories of government liabilities*

- unemployment rate (in %) – code: Unempl.

Quality of life:

- life expectancy at birth for females (in years) – code: Life F,
- life expectancy at birth for males (in years) – code: Life M,
- level of internet access (percentage of households)⁹ – code: Web. We will identify this characteristics with the use of information technology in a given country.

We considered introducing the period of membership as one of characteristics covered by the analysis, but we rejected this idea. For many years the European Union has been consistently implementing "Community policy" (mostly: social and economic) which strongly influences the actions taken by individual countries, and which – by consequence – is reflected in the economic condition of EU members. Thus, the period of membership would be strongly correlated with other "parameters" already included in the research. In addition, it would be extremely difficult to assign a weight to it as in fact we would have to answer the question whether the period of membership in the EU is more important, equally important, or less important than, for example, GDP per capita or the level of unemployment.

4.3 Research Methodology

We will analyse three "snapshots" of 28 countries: for 2004, 2009, and 2014¹⁰ – i.e. three sets of the above-described characteristics for those years. We will use databases of Eurostat¹¹ – however, with some exceptions described in Section 4.4.

4.4 Missing Values

Time series available in the Eurostat portal are not complete. For 2004, there are no values of:

- real GDP growth rate in Lithuania,

(as defined in ESA2010): currency and deposits, debt securities and loans. The general government sector comprises the subsectors: central government, state government, local government and social security funds. Government debt is also known as public debt, national debt and sovereign debt. In this study, we will use these terms interchangeably.

⁹Percentage of households which have internet access at home. All forms of internet use are included. The population considered is aged 16 to 74.

¹⁰By a "snapshot" we mean the status at the end of the year.

¹¹Access: July 3rd, 2015.

- general government gross debt in Greece,
- level of internet access in Belgium, Croatia, Malta, and Sweden.

As GDA does not allow to perform analysis based on incomplete data, we had to supplement them. Real GDP growth and general government debt could be found in the Trading Economics portal,¹² but the level of internet access could not. One of imputation¹³ methods used in cluster analysis is replacing each missing value with the average of its two adjacent values, but this approach could not be used here. We had no values for 2003 for Belgium, Malta, and Sweden; for Croatia it was even worse: we had no values for whole period 2003-2006. Therefore we decided to use another approach. We drew a scatter plot for the time series for each country.¹⁴ Then we added a trendline to it.¹⁵ Finally we counted values of the trend function for 2004 and we used them to fill in the missing values.¹⁶ Fig. 1 presents a scatter plot for Belgium.

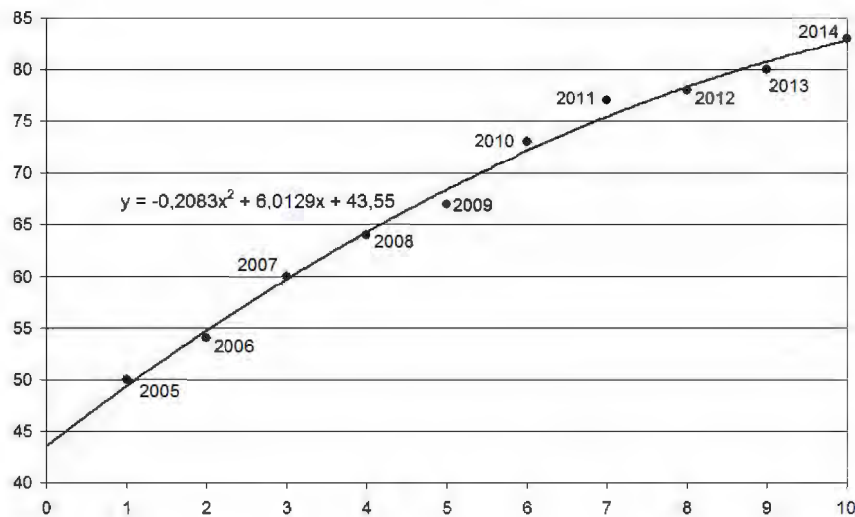


Fig. 1. Level of internet access. Scatter plot with a trend line for Belgium

¹²<http://www.tradingeconomics.com/>.

¹³The process of replacing missing data with substituted values.

¹⁴We used Excel.

¹⁵We used a polynomial function. For Sweden – of the third degree, for other countries – of the second degree.

¹⁶We assumed no seasonality.

4.5 Groups of Characteristics. Weights

We decided to put life expectancy at birth of males and of females in the same group. This way we could analyse "life expectancy of citizens" of each country. Other characteristics were placed in individual groups.

General government gross debt was given the highest weight: 1.5, all other groups – weights equal to 1. In [16] we analysed almost the same set of characteristics, and we assigned to all groups identical weights equal to 1. This time we changed our mind; the crisis in Greece, provoked mainly by its high public debt, was the reason.

4.6 Year 2004

As we said in Section 4.3, we will analyse "snapshots" of 28 countries taken during the period 2004-2014 in 5-year intervals. Table 1 contains a "snapshot" for 2004. Values marked in bold were missing in Eurostat databases; they were supplemented as described in Section 4.4. Table 2 presents values for 2004 normalised with groups and weights. Table 3 presents overrepresentation indexes for 2004.

Fig. 2 presents the overrepresentation map for 28 countries in 2004, and Fig. 3 explains the meanings of colours.

4.6.1 Year 2004. Clusters

We processed the data with GCA algorithm and we performed a cluster analysis. We grouped countries in 4 clusters, and their characteristics – in 3 clusters.¹⁷ Fig. 4 presents the resulting overrepresentation map.

We will start the analysis of results by looking at clusters of characteristics. Their grouping is surprising.

- Cluster 1. Inflation rate and real GDP growth rate.

GCA places in a common cluster characteristics which are "close" to each other. However, inflation is a negative phenomenon, while the growth of the GDP – a positive one. How to explain the fact that these two characteristics are close to each other?

Firstly, these characteristics are close to each other **in the analysed group of countries**; we cannot say, on the basis of our research, that inflation and GDP growth are always interrelated. Secondly, we can

¹⁷As we said in [15] and [16], there is no universal rule on how many clusters to search for. We decided to divide countries into 4 clusters, and their characteristics into 3 clusters after a visual analysis of the overrepresentation map. For more information please see [4].

Table 1. Characteristics of countries in 2004

Country	GDP	GDP +	Inflation	Debt	Unempl	Life F	Life M	Web
Austria	128	270	2.00	64.80	5.40	82.10	76.40	45
Belgium	121	340	1.90	96.60	8.10	81.90	76.00	44
Bulgaria	34	660	6.10	36.10	11.50	76.20	69.00	10
Croatia	57	410	2.10	39.80	13.40	78.80	71.80	17
Cyprus	97	440	1.90	64.70	5.20	81.80	76.50	53
Czech Republic	79	490	2.60	28.50	8.00	79.10	72.50	19
Denmark	125	260	0.90	44.20	5.30	80.20	75.40	69
Estonia	55	650	3.00	5.10	9.50	78.00	66.70	31
Finland	117	390	0.10	42.70	8.70	82.50	75.40	51
France	110	280	2.30	65.70	8.90	83.80	76.70	34
Germany	116	120	1.80	64.90	10.70	81.90	76.50	60
Greece	95	500	3.00	98.60	10.30	82.00	76.60	17
Hungary	62	480	6.80	58.80	6.50	77.20	68.70	14
Ireland	143	460	2.30	28.30	4.40	81.10	76.10	40
Italy	108	160	2.30	100.00	7.80	83.70	78.00	34
Latvia	48	890	6.20	14.20	11.90	76.00	65.60	15
Lithuania	50	740	1.20	18.70	10.30	77.70	66.20	12
Luxembourg	246	490	3.20	6.50	4.90	82.40	76.00	59
Malta	81	040	2.70	72.00	6.90	81.20	77.40	36
Netherlands	133	200	1.40	50.00	5.90	81.50	76.90	65
Poland	49	510	3.60	45.30	18.30	79.20	70.60	26
Portugal	77	180	2.50	62.00	8.30	81.80	75.00	26
Romania	34	840	11.90	18.60	7.80	75.10	67.80	6
Slovakia	57	520	7.50	40.60	17.60	78.00	70.30	23
Slovenia	86	440	3.70	26.80	6.60	80.80	73.50	47
Spain	100	320	3.10	45.30	10.30	83.70	77.00	34
Sweden	129	430	1.00	47.90	7.50	82.80	78.40	71
United Kingdom	125	250	1.30	40.20	4.70	81.10	76.80	56

Table 2. Characteristics of countries in 2004. Values normalised with groups and weights

Country	GDP	GDP +	Inflation	Debt	Unempl	Life F	Life M	Web	Marginal
Austria	0.006411	0.003061	0.003017	0.009767	0.002942	0.002537	0.002361	0.005917	0.036013
Belgium	0.006061	0.003855	0.002866	0.014560	0.004414	0.002530	0.002348	0.005786	0.042419
Bulgaria	0.001703	0.007483	0.009201	0.005441	0.006266	0.002354	0.002132	0.001315	0.035895
Croatia	0.002855	0.004649	0.003167	0.005999	0.007301	0.002435	0.002218	0.002235	0.030860
Cyprus	0.004859	0.004989	0.002866	0.009752	0.002833	0.002527	0.002364	0.006969	0.037158
Czech Republic	0.003957	0.005556	0.003922	0.004296	0.004359	0.002444	0.002240	0.002498	0.029271
Denmark	0.006261	0.002948	0.001357	0.006662	0.002888	0.002478	0.002330	0.009073	0.033997
Estonia	0.002755	0.007370	0.004525	0.000769	0.005176	0.002410	0.002061	0.004076	0.029142
Finland	0.005860	0.004422	0.000151	0.006436	0.004740	0.002549	0.002330	0.006706	0.033194
France	0.005510	0.003175	0.003469	0.009903	0.004849	0.002589	0.002370	0.004471	0.036335
Germany	0.005810	0.001361	0.002715	0.009782	0.005830	0.002530	0.002364	0.007890	0.038282
Greece	0.004758	0.005669	0.004525	0.014862	0.005612	0.002534	0.002367	0.002235	0.042562
Hungary	0.003105	0.005442	0.010256	0.008863	0.003542	0.002385	0.002123	0.001841	0.037557
Ireland	0.007163	0.005215	0.003469	0.004266	0.002397	0.002506	0.002351	0.005260	0.032627
Italy	0.005409	0.001814	0.003469	0.015073	0.004250	0.002586	0.002410	0.004471	0.039482
Latvia	0.002404	0.010091	0.009351	0.002140	0.006484	0.002348	0.002027	0.001972	0.036818
Lithuania	0.002504	0.008390	0.001810	0.002819	0.005612	0.002401	0.002045	0.001578	0.027159
Luxembourg	0.012322	0.005556	0.004827	0.000980	0.002670	0.002546	0.002348	0.007758	0.039005
Malta	0.004057	0.000454	0.004072	0.010852	0.003760	0.002509	0.002391	0.004734	0.032829
Netherlands	0.006662	0.002268	0.002112	0.007536	0.003215	0.002518	0.002376	0.008547	0.035233
Poland	0.002454	0.005782	0.005430	0.006828	0.009971	0.002447	0.002181	0.003419	0.038513
Portugal	0.003857	0.002041	0.003771	0.009345	0.004523	0.002527	0.002317	0.003419	0.031799
Romania	0.001703	0.009524	0.017949	0.002804	0.004250	0.002320	0.002095	0.000789	0.041433
Slovakia	0.002855	0.005896	0.011312	0.006120	0.009590	0.002410	0.002172	0.003024	0.043379
Slovenia	0.004308	0.004989	0.005581	0.004039	0.003596	0.002496	0.002271	0.006180	0.033460
Spain	0.005009	0.003628	0.004676	0.006828	0.005612	0.002586	0.002379	0.004471	0.035189
Sweden	0.006461	0.004875	0.001508	0.007220	0.004087	0.002558	0.002422	0.009336	0.038468
United Kingdom	0.006261	0.002834	0.001961	0.006059	0.002561	0.002506	0.002373	0.007364	0.031919
Marginal	0.133333	0.133333	0.133333	0.200000	0.133333	0.069568	0.063765	0.133333	1.000000

Table 3. Characteristics of countries in 2004. Overrepresentation indexes

Country	GDP	GDP +	Inflation	Debt	Unempl	Life F	Life M	Web
Austria	1.335193	0.637527	0.628232	1.356060	0.612776	1.012497	1.027940	1.232300
Belgium	1.071551	0.681565	0.506684	1.716226	0.780345	0.857489	0.868124	1.022941
Bulgaria	0.355823	1.563509	1.922389	0.757936	1.309265	0.942817	0.931419	0.274742
Croatia	0.693862	1.129752	0.769792	0.971967	1.774506	1.134075	1.127362	0.543272
Cyprus	0.980631	1.006902	0.578420	1.312224	0.571888	0.977696	0.997552	1.406630
Czech Republic	1.013861	1.423469	1.004802	0.733780	1.116903	1.200175	1.200134	0.640140
Denmark	1.381221	0.650320	0.299469	0.979817	0.637093	1.047717	1.074644	2.001577
Estonia	0.708994	1.896679	1.164547	0.131892	1.332225	1.188752	1.109037	1.049087
Finland	1.324085	0.999069	0.034079	0.969455	1.071084	1.103827	1.100631	1.515201
France	1.137250	0.652272	0.716056	1.362694	1.000984	1.024292	1.022819	0.922810
Germany	1.138303	0.266552	0.531898	1.277658	1.142242	0.950169	0.968282	1.545687
Greece	0.838485	0.998947	0.797350	1.745897	0.988970	0.855662	0.872049	0.393904
Hungary	0.620139	1.086774	2.048152	1.179898	0.707269	0.912917	0.886328	0.367617
Ireland	1.646467	1.198880	0.797445	0.653692	0.551117	1.103964	1.130166	1.209058
Italy	1.027576	0.344597	0.658982	1.908798	0.807344	0.941526	0.957250	0.849257
Latvia	0.489745	2.055512	1.904920	0.290661	1.320840	0.916768	0.863323	0.401782
Lithuania	0.691582	2.316894	0.499817	0.518902	1.549834	1.270608	1.181061	0.435738
Luxembourg	2.369198	1.068226	0.928051	0.125588	0.513377	0.938232	0.944105	1.491724
Malta	0.926869	0.103608	0.930364	1.652858	0.858927	1.098515	1.142390	1.081448
Netherlands	1.418052	0.482693	0.449495	1.069499	0.684332	1.027342	1.057566	1.819383
Poland	0.477948	1.126044	1.057409	0.886447	1.941824	0.913328	0.888239	0.665776
Portugal	0.909628	0.481334	0.889342	1.469381	1.066658	1.142465	1.142812	0.806337
Romania	0.308263	1.723942	3.248967	0.338318	0.769327	0.805006	0.792890	0.142812
Slovakia	0.493616	1.019339	1.955833	0.705360	1.658065	0.798594	0.785255	0.522893
Slovenia	0.965522	1.118193	1.250896	0.603626	0.806086	1.072486	1.064367	1.385261
Spain	1.067549	0.773283	0.996564	0.970188	1.196187	1.056404	1.060276	0.952877
Sweden	1.259745	0.950522	0.294069	0.938423	0.796762	0.955961	0.987528	1.820210
United Kingdom	1.471153	0.666022	0.460730	0.949169	0.601755	1.128458	1.165868	1.730238

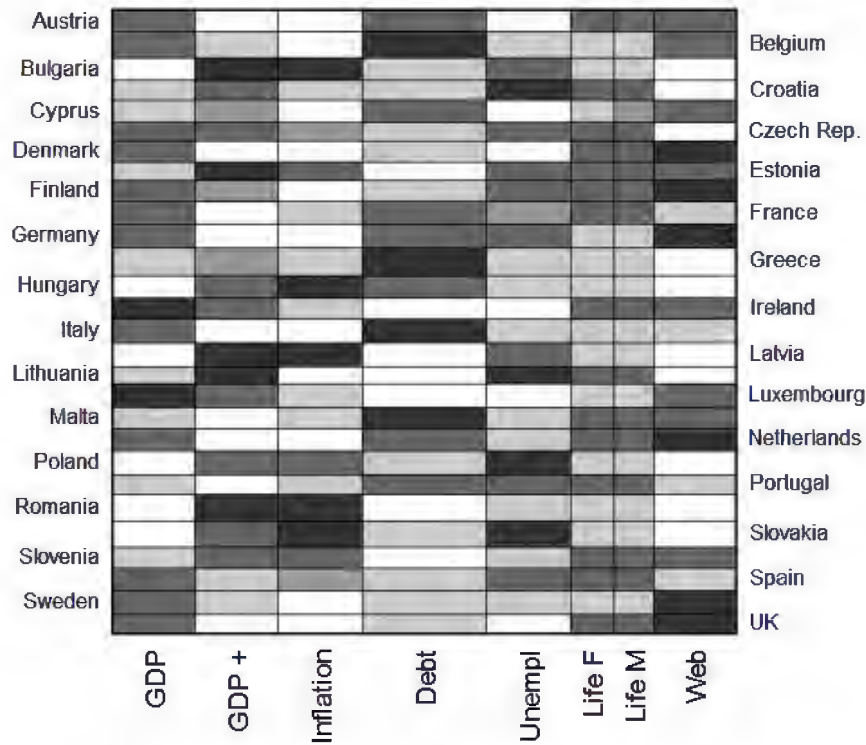


Fig. 2. Overrepresentation map. 28 countries in 2004



Fig. 3. Code of colours

analyse them only in the context of the entire study. The fact that inflation and GDP growth have been placed in the same cluster means that they are more closely related to each other than to other characteristics included in the research.¹⁸

¹⁸Discussion about the relationship between inflation and economic growth has been going on for many years. It is dominated by two beliefs. The first one is the "inflation-buys-growth" idea – an opinion that low inflation is one of the factors stimulating GDP growth (and, on the other hand,

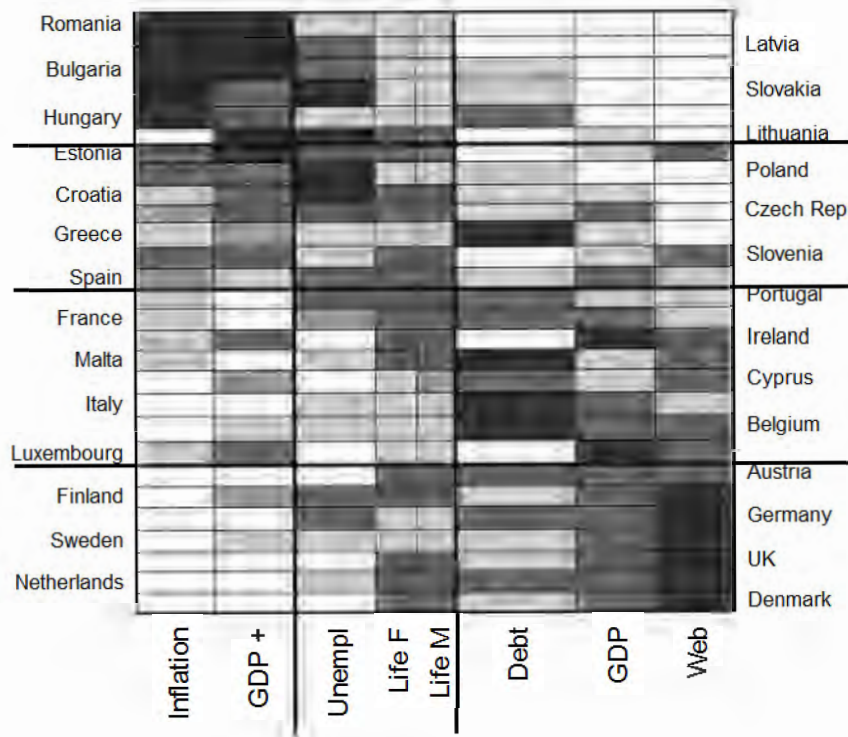


Fig. 4. Overrepresentation map after GCA and the cluster analysis. 28 countries in 2004

- Cluster 2. Unemployment rate, life expectancy at birth for females, and life expectancy at birth for males.

As the "middle" cluster, it contains the characteristics which differentiate the countries to a lesser degree than those grouped in clusters "on the edges of the map". However, we cannot say that unemployment and life expectancy are similar in each country; they are not – let's compare for instance unemployment rates for Ireland and Slovakia.

- Cluster 3. General government gross debt, Gross Domestic Product, and level of internet access.

In this cluster we may see that GDP and the level of internet access are closed to each other; it is not surprising: countries with high GDP are

low inflation is an inevitable consequence of that growth). Opposing view is that inflation is always an evil and has a negative effect on the economy.

In this paper, we have no ambition to give a clear answer why – in the countries covered by our research – we may observe some relationship between economic growth and inflation; we just note that it exists. Considering whether it is an "objective law" or an effect of some actions taken by the EU is beyond the scope of our study.

usually technology-advanced. However, the relationship between public debt and GDP per capita can be astonishing, at least at first glance – usually we think that the rich (in this case: wealthy countries) do not need to have recourse to debt.¹⁹

We have 4 clusters of countries:

- Cluster 1. Romania, Latvia, Bulgaria, Slovakia, Hungary, Lithuania. "New" countries which either have just joined the EU (Latvia, Slovakia, Hungary, Lithuania),²⁰ or still negotiate their accession (Romania and Bulgaria).²¹

This cluster is characterised by high inflation (strong overrepresentation except Lithuania), but also by high real GDP growth rate (strong or slight overrepresentation). On the other hand, countries of this cluster have low GDP (strong or slight underrepresentation) and use information technologies to a much lesser degree than the European average (strong underrepresentation). We may also observe a relatively high level of unemployment (strong or slight overrepresentation in 4 countries), as well as low life expectancy at birth (slight underrepresentation except Lithuania).

This is a cluster of emerging countries. They have low, but quickly growing GDP. They struggle with inflation and unemployment, and try to improve their low standard of living. They only introduce the use of the IT. However, what may be surprising, their public debt is relatively low (strong or slight underrepresentation except Hungary). We can risk the conclusion that countries of the first cluster are trying to bridge the gap separating them from the rest "without extravagance": they base mainly on their own resources and use them moderately.

- Cluster 2. Estonia, Poland, Croatia, Czech Republic, Greece, Slovenia, Spain. "New" countries which either have just joined the EU (Estonia, Poland, Czech Republic, Slovenia),²² or still negotiate their accession (Croatia),²³ plus two "old" countries (Greece and Spain).²⁴

¹⁹As in the case of inflation, the impact of public debt on economic growth (and consequently: on the level of GDP per capita) is the subject of many discussions. Quite common is the belief that public debt, as long as it does not exceed a reasonable level, stimulates economic growth. The problem is, however, how to determine that "reasonable level". More information may be found in [3], [6], [7] and [10].

²⁰They joined the EU on May 1st, 2004.

²¹They joined the EU on January 1st, 2007.

²²They joined the EU on May 1st, 2004.

²³Croatia joined the EU on July 1st, 2013.

²⁴Greece joined the EU on January 1st, 1981, and Spain – on January 1st, 1986.

Countries of this cluster have lower inflation, but also lower real GDP growth rate. Their GDP is slightly higher than that of countries in cluster 1, but still lower than that of countries grouped in clusters 3 and 4 (Greece is an exception). They are not IT usage leaders (except Estonia and Slovenia). Their unemployment rate is comparable with that in cluster 1 and much higher than that in clusters 3 and 4. Life expectancy is higher than that for cluster 1. Public debt is moderate (except Greece).

It is hard to find proper words to classify this cluster. It comprises countries which are still emerging (new or future EU members), but also Greece and Spain. These two "old" EU members have probably been included here because of their problems: unemployment for Spain and public debt for Greece.

- Cluster 3. Portugal, France, Ireland, Malta, Cyprus, Italy, Belgium, Luxembourg. Four Founding Member States²⁵ (France, Italy, Belgium and Luxembourg), two "old" EU members (Portugal and Ireland),²⁶ and two "new": Malta and Cyprus.²⁷

Countries of this cluster may be characterised by low inflation and low GDP growth rate (except Ireland and Luxembourg). They have also high GDP and high level of IT usage. Their unemployment is lower, and life expectancy – higher than in previous clusters. However, on this positive image there is one crack: public debt in nearly all countries is high. Only Ireland and Luxembourg have an underrepresentation of it; all other countries have a slight or even strong overrepresentation.

This cluster groups developed countries. They are rich, and they offer a high standard of living. However, they must remember that some part of their success is "borrowed" and one day will have to be paid back.

- Cluster 4. Austria, Finland, Germany, Sweden, United Kingdom, the Netherlands, Denmark. Two Founding Member States (Germany and the Netherlands), and five "old" EU members.²⁸

All countries in this cluster have low inflation (strong underrepresentation), but also low GDP growth rate (ideal representation at best). Their GDP is high (slight overrepresentation in each country), and their usage of IT – very high. Life expectancy of their people is also high. These

²⁵Countries who in 1957 signed the Treaty of Rome, which created the European Economic Community.

²⁶Portugal joined the EU on January 1st, 1986, and Ireland – on January 1st, 1973.

²⁷They joined the EU on May 1st, 2004.

²⁸Austria, Finland and Sweden joined the EU on January 1st, 1995; United Kingdom and Denmark – on January 1st, 1973.

countries are not, though, free of trouble; two of them have high unemployment and three of them – high public debt.

This last cluster groups EU leaders: very rich countries with high standard of living and relatively low financial liabilities.

Across the 28 countries surveyed, we can see some interesting patterns.

- GDP growth rate and inflation are correlated (that is why these two characteristics have been put in the same cluster), and it is a positive correlation:²⁹ when inflation decreases, also decreases the rate of growth of GDP and vice-versa (see Table 6 on page 79). This relationship may explain why some governments allow for an increase in inflation in order to stimulate economic growth. However, there is the question of what long-term consequences of such a policy may be.
- High public debt is characteristic not for poorer countries of the community, but for wealthier. This phenomenon raises one question: isn't a high standard of wealthy countries – at least in part – illusory? However, public debt is the most overrepresented in cluster 3 (developed countries), not in cluster 4 (leaders). So maybe public debt is a factor stimulating development, and when a country reaches a certain standard, it loses its importance, thus: decreases?
- GDP per capita and IT usage are strongly positively correlated³⁰ (see Table 6 on page 79). High GDP makes IT usage cheaper, so it becomes more available. High utilization of IT stimulates economic development, and influences on GDP growth. Thus, economic development and technological advances stimulate each other.

If we perform a cluster analysis for countries dividing them only into two clusters, we get a very interesting result. GCA creates one cluster for old EU members, Malta and Cyprus, and another one for new members plus Greece (see Fig. 5). Thus, according to the results of our research, the condition of Greece – already in 2004 – made it more similar to countries newly joining the EU than to those who were its members for many years.³¹

²⁹ $\rho = 0.548356$.

³⁰ $\rho = 0.852034$.

³¹ It should be also noted that all new EU members in the first cluster are countries of eastern and central Europe. Malta and Cyprus have always been close to the EU, so it seems natural that GCA placed them in the second cluster. However, the presence of Greece in the first cluster is surprising. It shows that Greece already in 2004 – by criteria adopted in our research – was more similar to countries of the former Eastern bloc than to those of the "old Union".

4.6.2 Year 2004. Outliers

Outliers detection shows that the most outlying country in 2004 is Luxembourg³² (see Fig. 6). We can explain it after a careful analysis of the overrepresentation map. Luxembourg belongs to cluster 3 which groups developed countries. They are characterised, among others, by low (or relatively low) GDP growth rate and high public debt. Values of these two characteristics for Luxembourg are different: it has high GDP growth rate and low public debt.

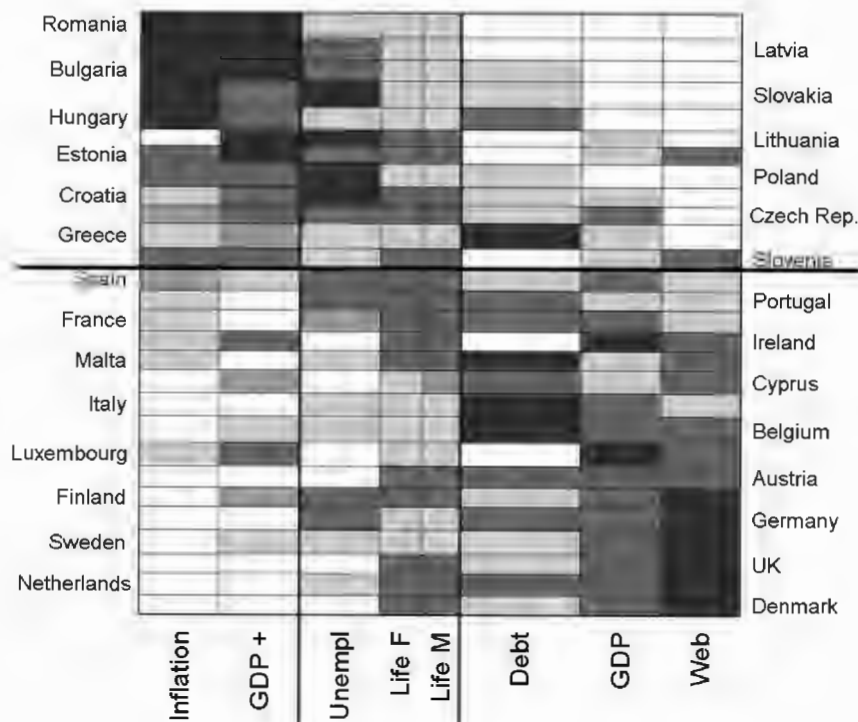


Fig. 5. Overrepresentation map after GCA and the cluster analysis. 28 countries in 2004, only 2 clusters of countries

In Fig. 4 we may observe that Ireland is very similar to Luxembourg: it has the same colours in all cells except those corresponding to life expectancy. Moreover, the overrepresentation of life expectancy is higher in

³² $AvgDistA = 0.183754$.

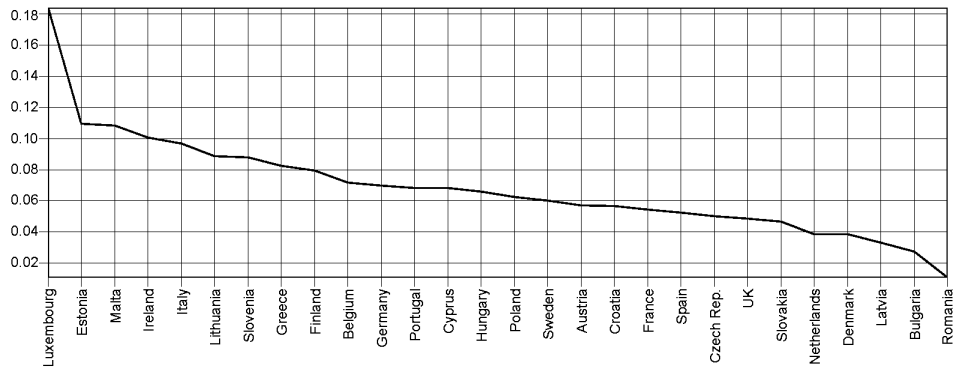


Fig. 6. Outliers for countries. 28 countries in 2004

Ireland than in Luxembourg. We may wonder why, then, GCA indicates Luxembourg, and not Ireland, as the most outlying country.

Unfortunately, we cannot explain it based only on the overrepresentation map. The code of colours we use (see Fig. 3) does not permit to observe some differences between objects and their characteristics. We have to go back to Table 1 and analyse rows representing both countries. GDP in Ireland is nearly 42% lower, and Irish public debt is over 4 times higher than those of Luxembourg.

Surprisingly, Greece and Italy are not outlying countries in spite of their very high public debt. This drawback of these countries is compensated with their low unemployment and low inflation rates. Greece has also relatively high GDP growth rate.

Outliers detection for characteristics shows that the most outlying one is public debt³³ (see Fig. 7). It becomes clear when we realise how diversified this characteristics is: from 5.10% of GDP (in Estonia) to 100.00% of GDP (in Italy).³⁴ We may observe it on the overrepresentation map – column Debt is very "colour-diversified".

4.6.3 Year 2004. Outliers Eliminated. New Analysis

As we said in Section 3, outliers may hinder the perception of certain relationships in the data. Those "hidden" relationships may become visible if we remove outliers from the data, proceed the data again with GCA and perform another outliers detection.

³³ $AvgDistA = 0.120861$.

³⁴ See Table 1.

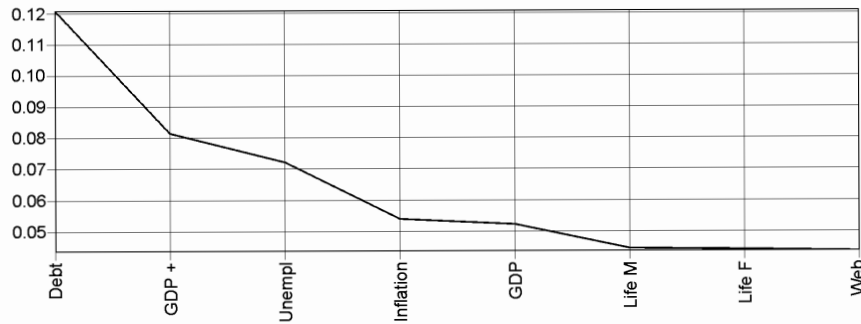


Fig. 7. Outliers for characteristics. 28 countries in 2004

Fig. 8 presents the overrepresentation map for 27 countries (without Luxembourg) in 2004 after GCA and the cluster analysis. As we may notice, the only difference between this map and the one presented in Fig. 4 is the absence of Luxembourg. Neither the order of countries, nor the clusters have changed.

Fig. 9 presents the result of outliers detection for 27 countries (without Luxembourg). There is no evident outlier; the graph is fairly smooth, and the first country (Estonia) stands out from the others in a relatively small extent. Thus, Luxembourg is a "European leader" and its position is not endangered.

Fig. 10 presents the overrepresentation map for 28 countries in 2004 after GCA and the cluster analysis with public debt removed. As we may notice, it is only the order of countries in clusters what differs this map from one presented in Fig. 4.

Fig. 11 presents the result of outliers detection for characteristics without public debt taken into account. There is no evident outlier; the graph is fairly smooth, and the first characteristics (GDP growth rate) stands out from the others in a relatively small extent.

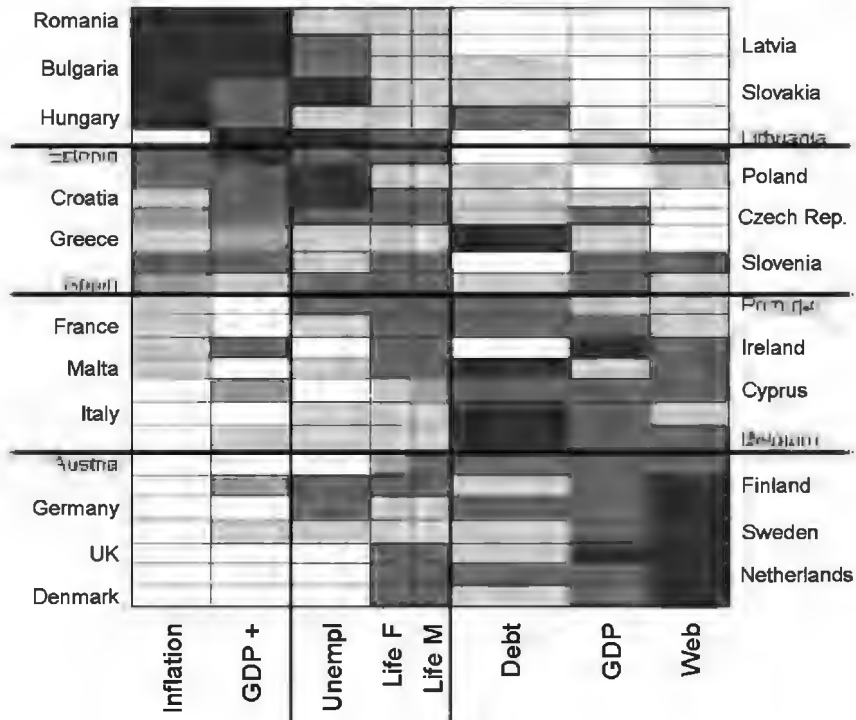


Fig. 8. Overrepresentation map after GCA and the cluster analysis. 27 countries in 2004 (without Luxembourg)

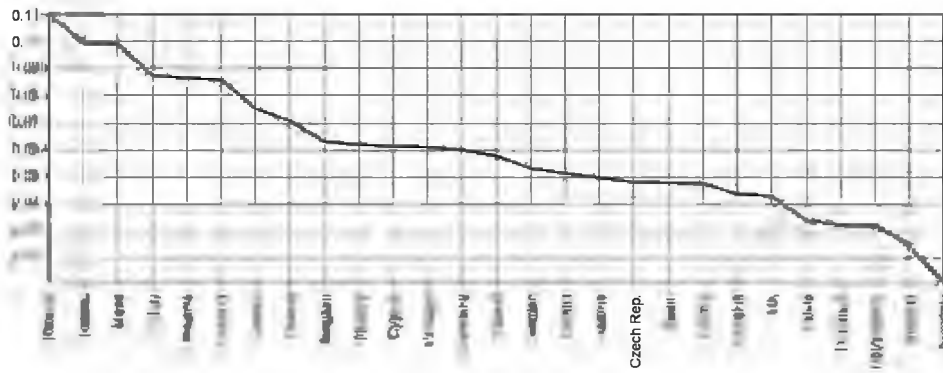


Fig. 9. Outliers for countries. 27 countries in 2004 (without Luxembourg)

4.7 Year 2009

Table 4 contains a "snapshot" of 28 countries for 2009. Values marked in bold were missing in Eurostat databases. They were supplemented as described in Section 4.4.

Table 4. Characteristics of countries in 2009

Country	GDP	GDP +	Inflation	Debt	Unempl	Life F	Life M	Web
Austria	126	-3.80	0.40	79.70	5.20	83.20	77.60	70
Belgium	117	-2.60	0.00	99.20	8.20	82.80	77.30	67
Bulgaria	44	-5.00	2.50	14.20	8.40	77.40	70.20	30
Croatia	62	-7.40	2.20	48.00	9.80	79.70	72.80	50
Cyprus	105	-2.00	0.20	54.10	6.50	83.50	78.50	53
Czech Republic	83	-4.80	0.60	34.10	7.50	80.50	74.30	54
Denmark	123	-5.10	1.10	40.40	6.90	81.10	76.90	83
Estonia	62	-14.70	0.20	7.00	16.80	80.30	70.00	62
Finland	116	-8.30	1.60	41.70	8.70	83.50	76.60	78
France	108	-2.90	0.10	79.00	9.50	85.00	78.00	69
Germany	115	-5.60	0.20	72.60	7.50	82.80	77.80	79
Greece	94	-4.40	1.30	129.70	10.70	83.30	77.50	38
Hungary	64	-6.60	4.00	78.20	10.70	78.40	70.30	55
Ireland	128	-6.40	-1.70	62.30	13.20	82.70	77.80	67
Italy	105	-5.50	0.80	112.50	8.40	84.30	79.10	53
Latvia	53	-14.20	3.30	36.40	20.10	77.70	67.50	58
Lithuania	57	-14.80	4.20	29.00	16.60	78.70	67.10	60
Luxembourg	247	-5.30	0.00	15.50	4.80	83.30	78.10	87
Malta	84	-2.50	1.80	67.80	6.90	82.70	77.90	64
Netherlands	137	-3.80	1.00	56.50	5.00	82.90	78.70	90
Poland	59	2.60	4.00	49.80	9.00	80.10	71.60	59
Portugal	81	-3.00	-0.90	83.60	11.30	82.80	76.50	48
Romania	49	-7.10	5.60	23.20	7.00	77.70	69.80	38
Slovakia	71	-5.30	0.90	36.00	14.40	79.10	71.40	62
Slovenia	85	-7.80	0.90	34.50	6.50	82.70	75.90	64
Spain	101	-3.60	-0.20	52.70	19.00	85.00	78.80	53
Sweden	123	-5.20	1.90	40.30	9.00	83.50	79.40	86
United Kingdom	112	-4.30	2.20	65.80	7.60	82.40	78.30	77

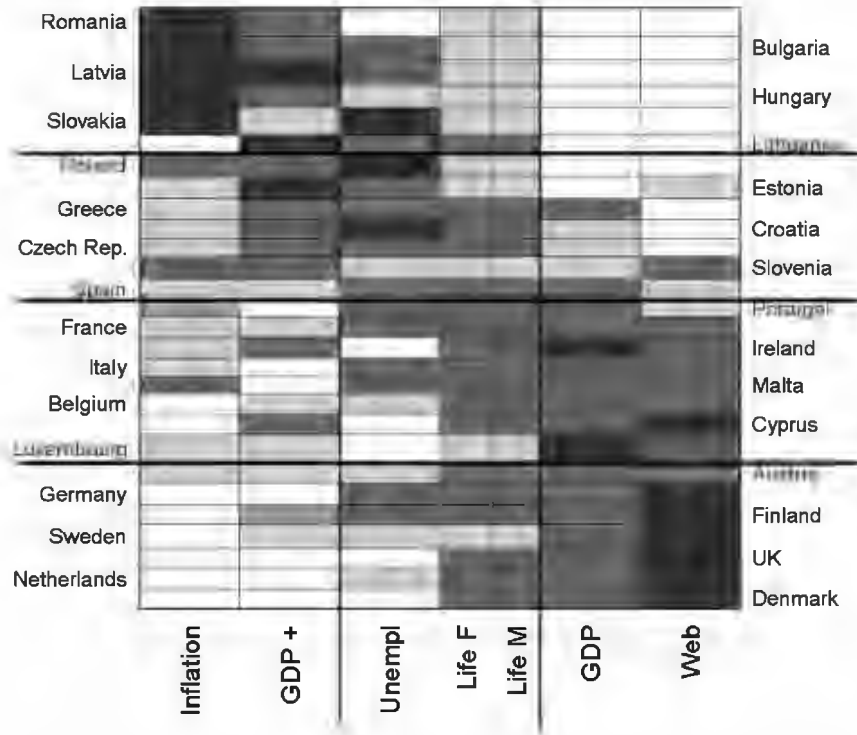


Fig. 10. Overrepresentation map after GCA and the cluster analysis. 28 countries in 2004 (without public debt)

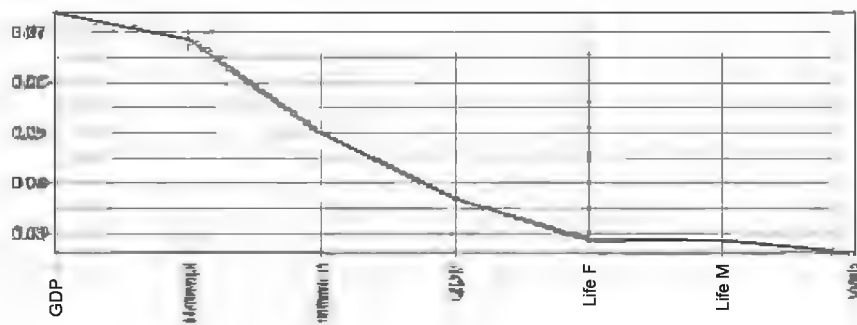


Fig. 11. Outliers for characteristics. 28 countries in 2004 (without public debt)

4.7.1 Year 2009. Problem with negative values

Some values in Table 4 are negative: GDP growth rate for all countries except Poland, and inflation rate for Ireland, Portugal, and Spain. As GCA

cannot analyse negative values, we had to eliminate them. We considered several approaches:

- To remove columns with negative values from the analysis.
Because GDP growth rate and inflation are important measures of economic condition of each country, removing them would make our research useless.
- To introduce two new columns: real GDP decline rate (code: GDP –) and deflation rate (code: Deflation); to replace negative values in column GDP + by zeros and put their negations in column GDP –, and to proceed the same way with columns Inflation and Deflation.
This approach seemed attractive. It would preserve all rules used for 2004 – we may say, in fact, that these two columns were also present previously, but filled exclusively with zeros. However, they would dominate whole analysis, making it also useless. Let's consider Poland – the only country with positive GDP growth rate. After normalisation of columns Poland would have value 1 in column GDP + and values much smaller in other columns. Thus, GDP + would dominate the entire analysis of Poland, and Poland would dominate the entire analysis of GDP + (see Fig. 12). Such an analysis would also be useless.
- To "shift all values to the right", i.e. to subtract from each value in columns GDP + and Inflation the minimum value for each column: –14.80 for GDP + and –1.70 for Inflation.
This approach would preserve "proportions of distances" between pairs of objects "measured in a certain characteristics" during normalisation of columns (see Formula (5)).

$$\frac{x_{is} - x_{js}}{x_{ks} - x_{ls}} = \frac{\frac{x_{is} - x_{min,s}}{\sum_{p=1}^n (x_{ps} - x_{min,s})} - \frac{x_{js} - x_{min,s}}{\sum_{p=1}^n (x_{ps} - x_{min,s})}}{\frac{x_{ks} - x_{min,s}}{\sum_{p=1}^n (x_{ps} - x_{min,s})} - \frac{x_{ls} - x_{min,s}}{\sum_{p=1}^n (x_{ps} - x_{min,s})}}, \quad (5)$$

where

x_{rs} – value in r -th row and s -th column (value of s -th characteristics in r -th country),

$x_{min,s}$ – minimum value in s -th column (minimum value of s -th characteristics in all countries),

n – number of rows (number of countries).

We chose the last solution.

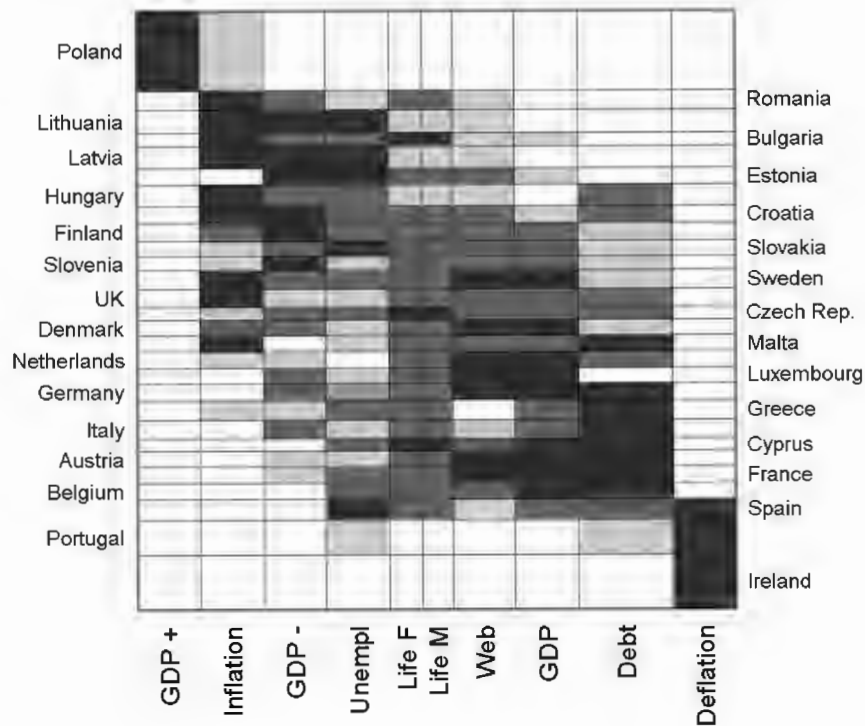


Fig. 12. Overrepresentation map after GCA. 28 countries in 2009. Two new columns: GDP – and Deflation

4.7.2 Year 2009. Clusters

We processed the data with GCA algorithm and we performed a cluster analysis. We used the same groups and weights as for 2004 (see Section 4.5). As before, we identified 4 clusters of countries and 3 clusters of characteristics. Fig. 13 presents the resulting overrepresentation map.

We may observe that clusters of characteristics have changed.

- Cluster 1. Inflation rate and unemployment rate.
As previously, inflation rate was placed "on the edge" of the overrepresentation map, which means: it is strongly differentiating countries in the study. Unlike in 2004, it was grouped with unemployment rate, not with GDP growth rate. Thus, two features considered negative phenomena form a common cluster, what seems to be natural.
- Cluster 2. Life expectancy at birth for females, level of internet access, life expectancy at birth for males, and Gross Domestic Product.
Once again we have a more natural cluster than that for 2004. Life ex-

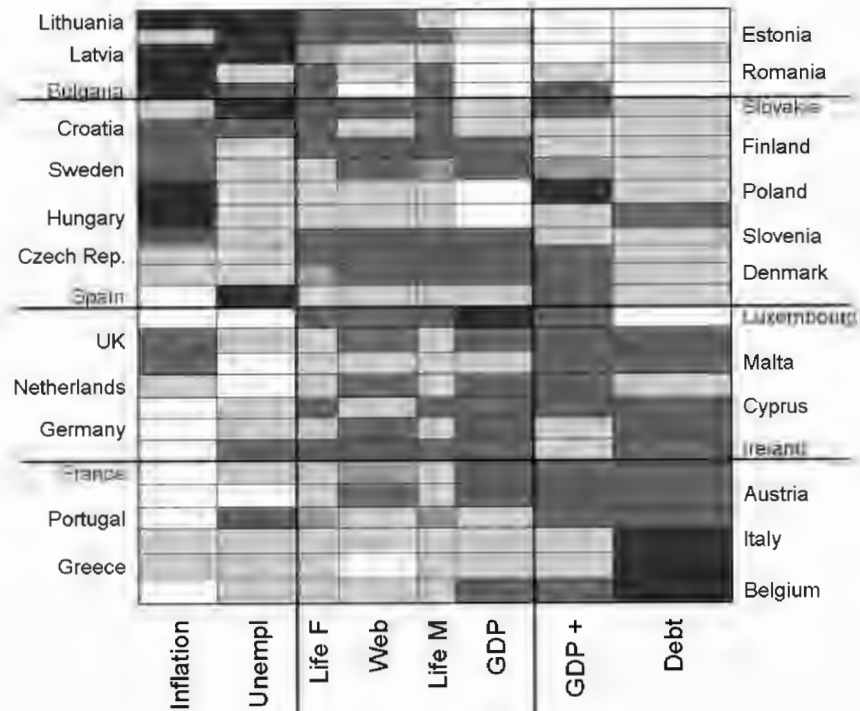


Fig. 13. Overrepresentation map after GCA and the cluster analysis. 28 countries in 2009

pectancy, GDP per capita, and IT usage can be considered measures of quality of life. What is surprising, life expectancies for females and males are not adjacent. It may signal that in the analysed countries life expectancies for women and men do not approach and in the period 2004-2009 their differentiation increased. However, to verify it, we should also analyse changes observed in the same period between other pairs of characteristics.

- Cluster 3. GDP growth rate and general government gross debt. As in 2004, we have a puzzling combination of GDP growth rate with a characteristics that awakens negative connotations. In 2004 it was inflation rate, and now it is public debt. This makes us wonder whether the development of the EU countries (at least of some of them) was not development "on credit".

Clusters of countries also changed. Now these are:

- Cluster 1. Lithuania, Estonia, Latvia, Romania, Bulgaria.

- Cluster 2. Slovakia, Croatia, Finland, Sweden, Poland, Hungary, Slovenia, Czech Republic, Denmark, Spain.
- Cluster 3. Luxembourg, United Kingdom, Malta, the Netherlands, Cyprus, Germany, Ireland.
- Cluster 4. France, Austria, Portugal, Italy, Greece, Belgium.

We may feel the temptation to compare these clusters with those received for 2004, to indicate which countries remained in their "original" clusters, and which moved to the others. However, such an approach would be totally wrong. The explanation why may be found above. Clusters of characteristics for 2004 and 2009 vary considerably, some characteristics significantly shifted; the new identification of clusters of countries is based on rules different from those used 2004. This means that **the situation in whole EU significantly changed**. Thus, in 2009 it is necessary to re-examine the EU as a whole, not its individual subgroups identified five years earlier.

- Cluster 1. It groups five new EU members. This cluster is characterised by high inflation (strong overrepresentation except Estonia) and high unemployment (strong or slight overrepresentation except Romania). Countries of this cluster have low GDP per capita (strong or slight underrepresentation) and low GDP growth rate (strong or slight underrepresentation except Bulgaria). IT usage is differentiated: it is low in three countries (Latvia, Romania, and Bulgaria), but high in two (Lithuania and Estonia).
This is a cluster of emerging countries. They have low and slowly growing GDP. They struggle with inflation and unemployment. However, they try to develop based mainly on their own resources and use them moderately – their public debt is low (strong or slight underrepresentation). We may risk the conclusion: they are not rich, but reasonable.
- Cluster 2. It groups four old EU members and six new. This cluster may be distinguished from the previous one by much lower unemployment (slight underrepresentation in seven countries out of ten) and slightly lower inflation (only two countries with high overrepresentation). It has also higher GDP per capita which is growing faster. Unfortunately, countries of the second cluster are less thrifty; nine of them have slight underrepresentation of public debt, and one: slight overrepresentation, while in the first cluster it is strong underrepresentation which dominates. IT usage is moderate: from slight underrepresentation to slight overrepresentation.

It is the most diversified cluster. It is hard to say why it unites so different countries. Perhaps what is joining them is a period of dynamic changes they are going through? To verify it, however, further in-depth study would be needed.

- Cluster 3. It groups five old EU members and two new. This cluster may be distinguished from the previous one by lower inflation and lower unemployment (in both characteristics this is slight or strong underrepresentation which dominates). Countries of this cluster have high GDP per capita and its growth rate is high (domination of slight overrepresentation). Like in previous cluster, IT usage is moderate: from slight underrepresentation to slight overrepresentation.

Unfortunately, there is also one important difference between this cluster and both previous ones: high public debt. Five countries (out of seven) have strong overrepresentation of characteristic; it makes us wonder whether the development of these countries is not a development on credit, at the expense of future generations.

Some more attention should be devoted to two countries: Luxembourg and the Netherlands. Luxembourg, "the European banker", is the only country with strong overrepresentation of GDP and it has also strong underrepresentation of public debt. The Netherlands have slight overrepresentation of GDP and slight underrepresentation of public debt. Both countries have slight overrepresentation of GDP growth rate and strong underrepresentation of unemployment. They could be called "reasonable cousins" of the others belonging to the same cluster.

- Cluster 4. It groups six old EU members. This cluster may be characterised by low inflation (strong or slight underrepresentation) and relatively low unemployment (slightly lower than in the second cluster, but slightly higher than in the third one). IT usage in this cluster, as in the first one, is diversified: from strong underrepresentation to slight overrepresentation.

GDP per capita and its growth rate are lower than in the third cluster, though higher than in the remaining two. Public debt is much higher than in any other cluster: all countries have overrepresentation of this characteristics (three of them: slight, and three others: strong).

A comparison of this cluster with the other ones (especially with the third) raises a pessimistic hypothesis. Perhaps countries of the last cluster have crossed the "security border", abroad which the increase in public debt is not accompanied by economic development? The analysis for 2014 will verify this assumption.

When comparing overrepresentation maps for 2004 and 2009 we can observe that:

- Inflation is still one of the most important traits of the countries.
- GDP growth rate and inflation are correlated, but in 2004 it was a positive correlation, and in 2009 it is negative³⁵ (see Table 6 on page 79).
- In 2004 IT usage was one of two most important traits of the countries (the other one was inflation). In 2009 it is not. Column Web for 2009 is much less diverse than that for 2004. None of the countries has strong overrepresentation and only two of them have strong underrepresentation. Thus, in the period 2004-2009 the use of IT has been greatly expanded in new EU members.
- Public debt is still one of the basic ways of stimulating economic development. There are already first signs of the danger linked with this.

It is also worth noting that:

- Still very few cells of the overrepresentation map have an average level of grey which symbolizes an ideal representation. 28 countries under our analysis are still very diversified.
- Both in 2004 and 2009 columns Life F and Life M are not very diversified. None of 28 countries has strong under- or overrepresentation. We can say that when it comes to life expectancy, EU countries have reached a certain standard.

4.7.2 Year 2009. Outliers

Outliers detection for countries in 2009 is surprising (see Fig. 14). We may observe that:

- There is no clear one outlier. There is a group of outliers: Romania, Lithuania, Bulgaria, and Luxembourg.
- The "economic diversity" of the EU countries in the period 2004-2009 has increased, not decreased. In 2004 $AvgDistA$ for the most outlying country (Luxembourg) was equal to 0.183754, in 2009 it is equal to 0.201077 (for Romania). This makes us wonder about the effectiveness of European integration processes.

The first three outliers: Romania, Lithuania, and Bulgaria are considered as poor countries, while Luxembourg – as the richest one. For sure they have been indicated as outliers for different reasons.

³⁵In 2004 $\rho = 0.548356$, in 2009 $\rho = -0.308577$.

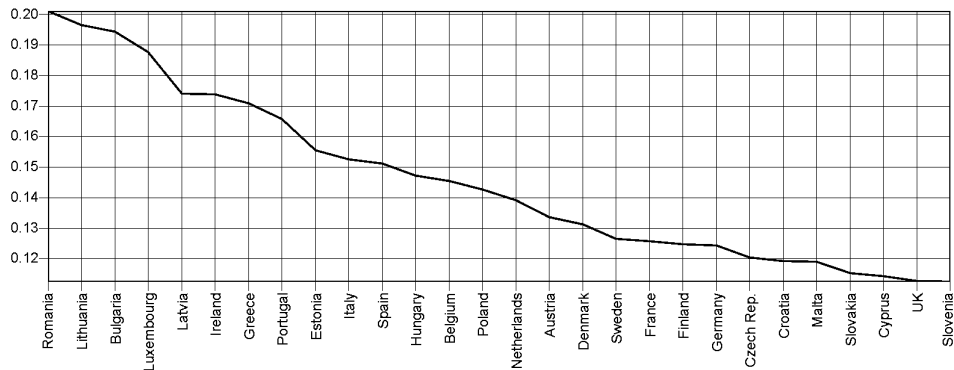


Fig. 14. Outliers for countries. 28 countries in 2009

The situation of Luxembourg is clear. It is the only country with strong overrepresentation of the GDP, and the only country with high GDP accompanied by low public debt. As for Romania, Lithuania, and Bulgaria, the problem is deeper. While analysing the overrepresentation map (see Fig. 13) we may notice that they have high inflation (strong overrepresentation) and low GDP (strong underrepresentation). The same may be observed for Latvia, Poland, and Hungary, but there is one difference: Romania, Lithuania, and Bulgaria have very low public debt (strong underrepresentation) which distinguishes them from Latvia, Poland, and Hungary. Thus, we may say that Romania, Lithuania, and Bulgaria are outliers because they are "poor, but reasonable".

Outliers detection for characteristics shows that the most outlying one is inflation³⁶ (see Fig. 15). We may easily observe that its column is the most "colour-diversified" in the overrepresentation map (see Fig. 13).

4.7.4 Year 2009. Outliers Eliminated. New Analysis

Fig. 16 presents the overrepresentation map for 25 countries (without Romania, Lithuania, Bulgaria, and Luxembourg) in 2009 after GCA and the cluster analysis. We may notice only some minor differences between this map and the one presented in Fig. 13: columns Life F and Web have been swapped, some countries have also been swapped, and 5 of them have moved to lower clusters.

³⁶ $AvgDistA = 0.227459$.

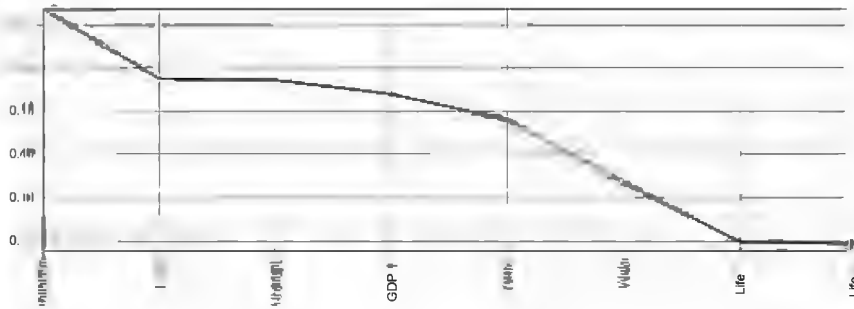


Fig. 15. Outliers for characteristics. 28 countries in 2009

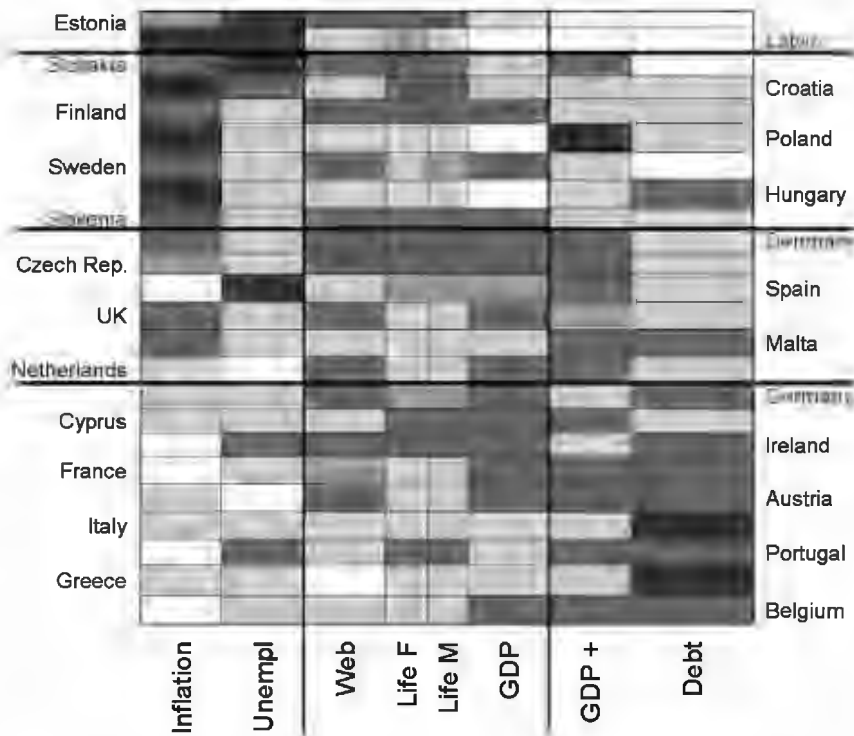


Fig. 16. Overrepresentation map after GCA and the cluster analysis. 25 countries (without Romania, Lithuania, Bulgaria, and Luxembourg) in 2009

Fig. 17 presents the result of outliers detection for 25 countries (without Romania, Lithuania, Bulgaria, and Luxembourg). The graph is interesting; we may say that there is a new outlier: Poland, but we may also say that there is a new group of outliers: Poland, Ireland, Hungary, and Spain.

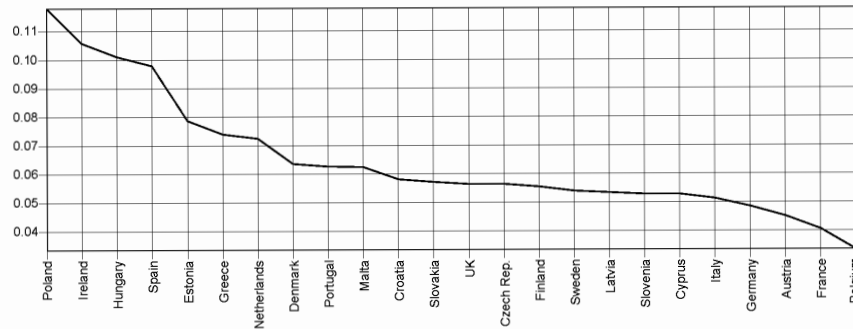


Fig. 17. Outliers for countries. 25 countries (without Romania, Lithuania, Bulgaria, and Luxembourg) in 2009

Fig. 18 presents the overrepresentation map for 28 countries in 2009 after GCA and the cluster analysis with inflation removed from characteristics. Once again we may notice only some minor differences between this map and the one presented in Fig. 13.

Fig. 19 presents the result of outliers detection for characteristics without inflation taken into account. The new outlier is GDP per capita, but it stands out from other characteristics in much lower extent³⁷ than inflation. Though we have a new outlier, we will not perform another analysis with this characteristics removed. We think that an analysis without GDP per capita taken into account would have no sense.

4.8 Year 2014

Table 5 contains a "snapshot" of 28 countries for 2014. GDP growth rate for Luxembourg was missing; we supplemented it as described in Section 4.4. Data for life expectancy at birth were also missing (for all countries), so we supplemented them by data for 2013.

We transformed columns GPD + and Inflation to eliminate negative values using the same procedure as in Section 4.7. Then we analysed the table in GradeStat using the same groups and weights as for 2004 and 2009 (see Section 4.5).

4.8.1 Year 2014. Clusters

³⁷ $AvgDistA = 0.099908$.

Table 5. Characteristics of countries in 2014

Country	GDP	GDP +	Inflation	Debt	Unempl	Life F ^a	Life M ^b	Web
Austria	128	0.30	1.50	84.50	5.60	83.80	78.60	81
Belgium	119	1.10	0.50	106.50	8.60	83.20	78.10	83
Bulgaria	45	1.70	-1.60	27.60	10.20	78.60	71.30	57
Croatia	59	-0.40	0.20	85.00	17.70	81.00	74.50	68
Cyprus	85	-2.30	-0.30	107.50	16.50	85.00	80.10	69
Czech Republic	84	2.00	0.40	42.60	5.80	81.30	75.20	78
Denmark	124	1.10	0.30	45.20	6.10	82.40	78.30	93
Estonia	73	2.10	0.50	10.60	6.40	81.70	72.80	83
Finland	110	-0.10	1.20	59.30	9.00	84.10	78.00	90
France	107	0.20	0.60	95.00	10.40	85.60	79.00	83
Germany	124	1.60	0.80	74.70	4.80	83.20	78.60	89
Greece	72	0.80	-1.40	177.10	25.90	84.00	78.70	66
Hungary	68	3.60	0.00	76.90	7.30	79.10	72.20	75
Ireland	132	4.80	0.30	109.70	10.20	83.10	79.00	82
Italy	97	-0.40	0.20	132.10	12.40	85.20	80.30	73
Latvia	64	2.40	0.70	40.00	10.30	78.90	69.30	73
Lithuania	74	2.90	0.20	40.90	10.00	79.60	68.50	66
Luxembourg	263	2.90	0.70	23.60	5.90	83.90	79.80	96
Malta	85	3.50	0.80	68.00	5.90	84.00	79.60	81
Netherlands	130	1.00	0.30	68.80	7.20	83.20	79.50	96
Poland	68	3.40	0.10	50.10	8.20	81.20	73.00	75
Portugal	78	0.90	-0.20	130.20	13.60	84.00	77.60	65
Romania	54	2.80	1.40	39.80	6.60	78.70	71.60	61
Slovakia	76	2.40	-0.10	53.60	12.40	80.10	72.90	78
Slovenia	83	2.60	0.40	80.90	9.30	83.60	77.20	77
Spain	93	1.40	-0.20	97.70	23.60	86.10	80.20	74
Sweden	124	2.30	0.20	43.90	7.60	83.80	80.20	90
United Kingdom	108	3.00	1.50	89.40	5.50	82.90	79.20	90

^aFor 2013.^bFor 2013.

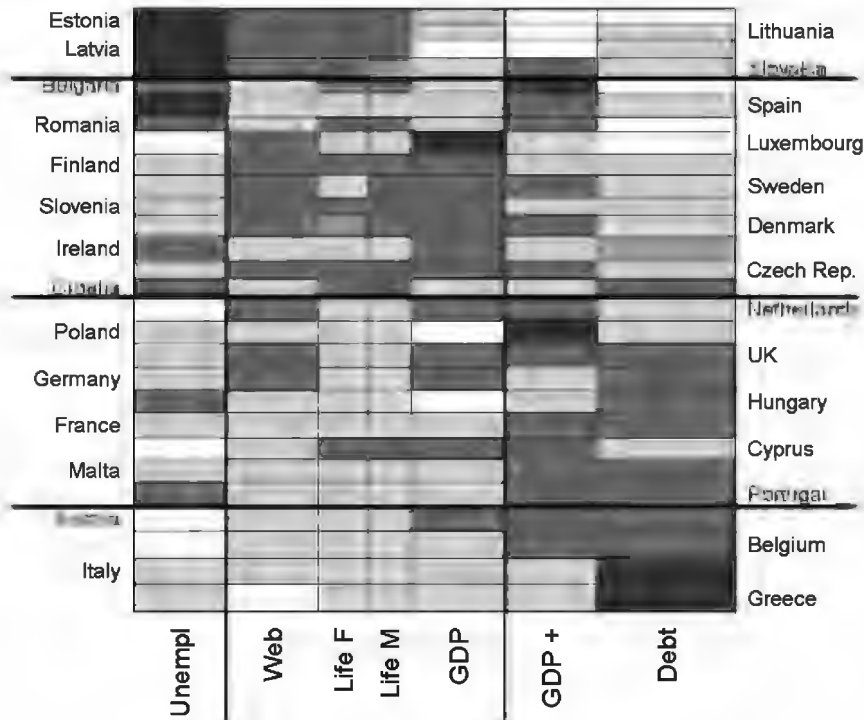


Fig. 18. Overrepresentation map after GCA and the cluster analysis. 28 countries in 2009 (without inflation)

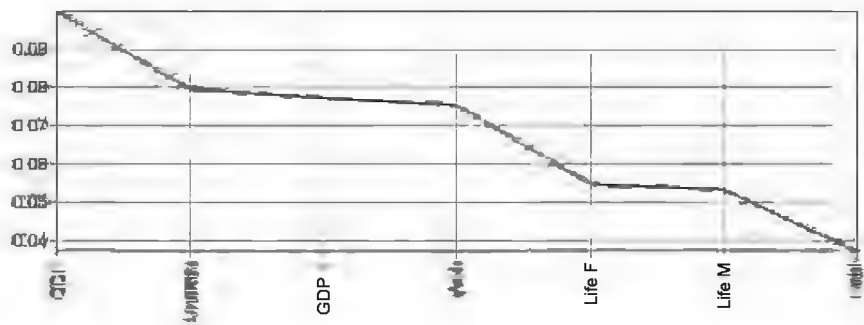


Fig. 19. Outliers for characteristics. 28 countries in 2009 (without inflation)

We processed the data with GCA algorithm and we performed a cluster analysis. We grouped countries in 4 clusters, and their characteristics – in 3 clusters. Fig. 20 presents the resulting overrepresentation map.

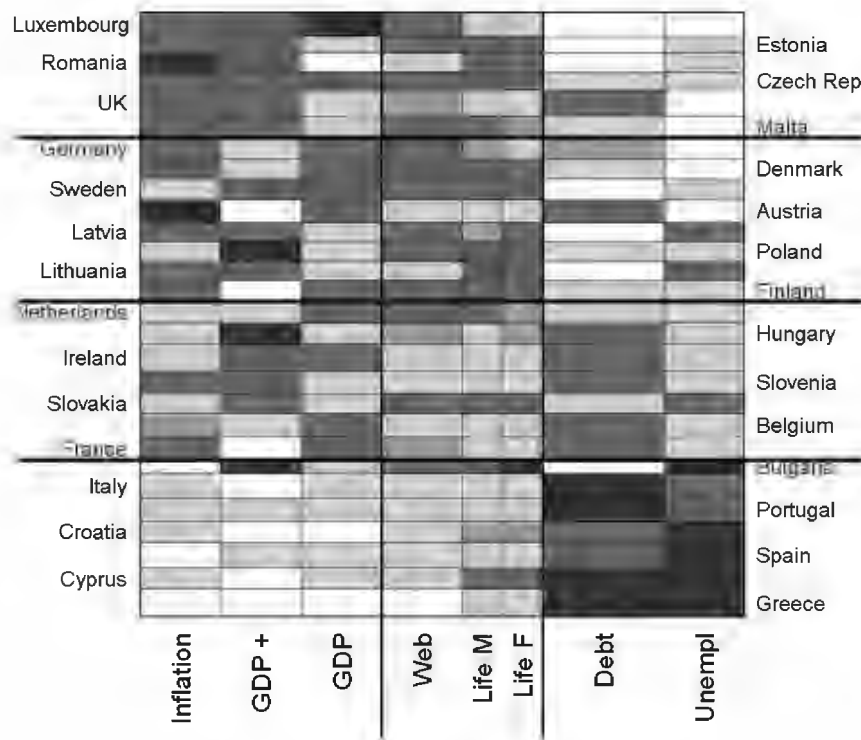


Fig. 20. Overrepresentation map after GCA and the cluster analysis. 28 countries in 2014

We have the following clusters of characteristics:

- Cluster 1. Inflation, GDP growth rate, and GDP per capita.
Once again, inflation is "at the heart" of the first cluster. It is grouped with GDP growth rate and GDP per capita. This combination requires a deeper reflection.
First, we should notice that in 2004 GDP per capita and GDP growth rate were located in non-adjacent clusters (see Fig. 4). These two characteristics were negatively correlated;³⁸ countries with lower GDP per capita had higher GDP growth rate. This means that poorer countries developed faster, which was in line with the policy of eliminating disparities of development between richer and poorer EU members. In 2009 this correlation was inverted.³⁹ In 2014 is again negative, but the principle "poorer countries develop faster" is much less visible.⁴⁰ The

³⁸ $\rho = -0.594276$.

³⁹ $\rho = 0.266191$.

⁴⁰ $\rho = -0.110715$.

question is whether this is due to the fact that disparities of development of individual countries decreased during the analysed ten years, or it signals a lack of a real "economic integration" policy.

Secondly, we should consider whether the combination of these three characteristics in one cluster is good. We will not give here a categorical answer, but we believe that it is an alarm signal (see our remarks on page 54).⁴¹

- Cluster 2. IT usage, life expectancy at birth for males, and life expectancy at birth for females.

These are characteristics which may be considered as measures of quality of life. They are grouped in one cluster which seems natural.

- Cluster 3. Public debt and unemployment.

These two characteristics are defects which each country tries to eliminate, or at least: to minimize. It seems natural they are forming a common cluster.

A comparison of the overrepresentation map for 2014 with those for years 2004 (see Fig. 4) and 2009 (see Fig. 13) leads to an interesting insight. In 2004 and 2009 public debt and unemployment were separated in two clusters, and their correlation was negative.⁴² Now they form their own cluster and are positively correlated.⁴³ We may venture the conclusion that excessive indebtedness of several countries finally triggered the intensification of negative economic phenomena.

Public debt initially may support development (e.g. cause an increase in employment), but when it is too high it has negative consequences such as unemployment. This relationship is not direct; when governments need to service the debt, they impose additional burdens on employers, and those – looking for savings – reduce the workforce.

Clusters of countries are the following:

- Cluster 1. Luxembourg, Estonia, Romania, Czech Republic, United Kingdom, Malta.
- Cluster 2. Germany, Denmark, Sweden, Austria, Latvia, Poland, Lithuania, Finland.
- Cluster 3. The Netherlands, Hungary, Ireland, Slovenia, Slovakia, Belgium, France.

⁴¹The correlation of inflation and GDP growth rate is equal to 0.166392. It is positive, but lower than in 2004, when it was equal to 0.548356. However, the correlation of inflation and GDP per capita is equal to 0.376682, while in 2004 and 2009 it was strongly negative (–0.606384 and –0.558722, respectively).

⁴²In 2004 $\rho = -0.036976$, in 2009 $\rho = -0.025198$.

⁴³ $\rho = 0.475017$.

- Cluster 4. Bulgaria, Italy, Portugal, Croatia, Spain, Cyprus, Greece.

In each cluster we can observe a "mix of countries": both "old" and "new" EU members. This is a significant change compared to previous years. In 2004, old and new Member States formed two dense blocks (although there were three exceptions – see our remarks on page 54). In 2009, this regularity was still noticeable. In 2014 it disappeared. We believe this is the most important change that took place during the period 2004-2014.

- Cluster 1. Its main characteristics is high GDP growth rate (slight overrepresentation) accompanied by high inflation (slight or strong overrepresentation). We may also notice low unemployment (slight or strong underrepresentation) and low public debt (slight or strong underrepresentation except the United Kingdom). In this cluster IT is commonly used (Romania is the only exception).

We may say countries of this cluster have avoided problems currently afflicting the EU: high public debt and lack of jobs. However, this was at the expense of high inflation.

- Cluster 2. This cluster is a real mixture. It is difficult to identify basic characteristics of countries which belong to it. Each column is diversified, and one of them: GDP + contains nearly all shades of grey,⁴⁴ which means: this cluster unites countries with very low GDP growth rate as well as those with very high.

We can understand the rule, according to which the GCA identified clusters of countries, when we look at the whole map. We may observe that above the middle horizontal line there is much light on the right; below this line there is much more light on the left. "Right" and "left" are designated by the second vertical line. Clusters 1 and 2 group countries with lower public debt and lower unemployment, but with higher inflation and – usually – higher GDP growth rate. In clusters 3 and 4 this relation is inverted. In clusters 1 and 2 we may also notice higher IT usage and higher life expectancy.

Thus, cluster 2 groups countries which have relatively low public debt (except Austria) and relatively low unemployment (except Latvia and Lithuania), but they are not as trouble-free as are countries of cluster 1. They have slightly lower inflation and higher GDP per capita than countries in cluster 1, but also much lower GDP growth rate (except Poland). IT usage and life expectancy are comparable in both clusters.

⁴⁴The only missing colour in the one representing ideal representation, but it is almost absent from the entire map.

- Cluster 3. In this cluster high public debt is a common concern (slight overrepresentation except the Netherlands and Slovakia). Unemployment is not very high (slight underrepresentation except Slovakia), but soon may become a serious problem. Inflation is lower than in cluster 2, and GDP per capita and GDP growth rate – comparable to those in cluster 2. IT usage and life expectancy are lower than in clusters 1 and 2 (except the Netherlands and Slovakia).

This cluster groups countries at risk of crisis. They have still relatively good economic results, but their high public debt is a serious burden for them. For now, there are reasons for concern for the Netherlands and Slovakia, but other countries should find a way out of the loop of debt.

- Cluster 4. This cluster stands out on the map: there is a large light area in the left part of it, and a small dark area in the right. That means: it groups countries with low GDP per capita and low GDP growth rate, and with high public debt and high unemployment. Surprisingly, despite their debts, those countries have low inflation. We may also notice low IT usage and low life expectancy.

However, there is one exception: Bulgaria. It has low public debt, high GDP growth rate, high IT usage and high life expectancy. These are characteristics that do not fit in this cluster. We can assume that Bulgaria has been put in this cluster because of its high unemployment, low GDP per capita and low inflation which are typical for cluster 4. As Bulgaria is "50% in cluster 4, and 50% outside it", we may expect it is an outlier.

4.8.2 Year 2014. Outliers

Fig. 21 proves that Bulgaria is really the most outlying country. However, the differentiation of countries in 2014 is lower than it was in 2004 and 2009.⁴⁵

Fig. 22 shows that there is not one outlying characteristics, and there is not a group of outlying characteristics, either. The graph is smooth and the value of *AvgDistA* for the first characteristics (GDP +) is lower than it was in 2004 and 2009.⁴⁶

⁴⁵ *AvgDistA* = 0.128539.

⁴⁶ *AvgDistA* = 0.104057.

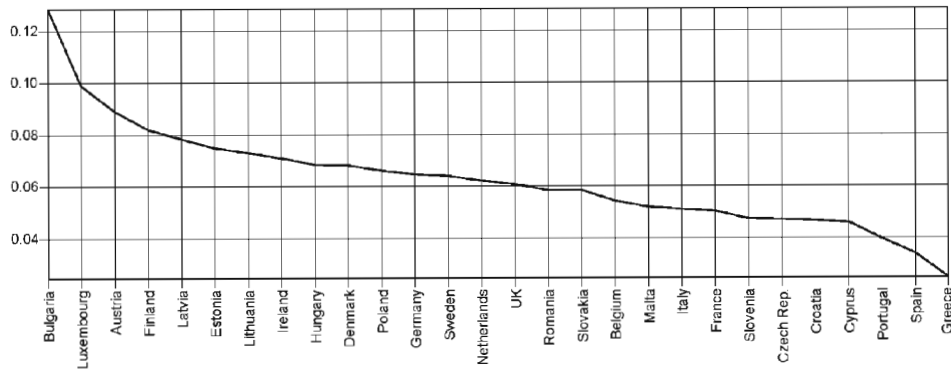


Fig. 21. Outliers for countries. 28 countries in 2014

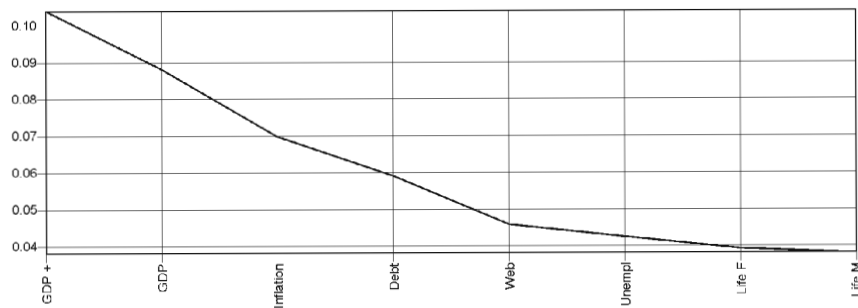


Fig. 22. Outliers for characteristics. 28 countries in 2014

4.8.3 Year 2014. Outliers Eliminated. New Analysis

Fig. 23 presents the overrepresentation map for 27 countries (without Bulgaria).

We may observe the following changes in clusters of characteristics:

- Columns Inflation and GDP + have been swapped.
- Column GDP has been shifted from cluster 1 to cluster 2.
- Columns Life M and Life F have been swapped.
- Columns Debt and Unemployment have been swapped.

As a result, clusters of countries also changed:

- The United Kingdom, a former member of cluster 1, is now located in cluster 2, so there is no country with high public debt in cluster 1. The rule used by GCA: separating countries with high public debt from those without it is now even more visible.

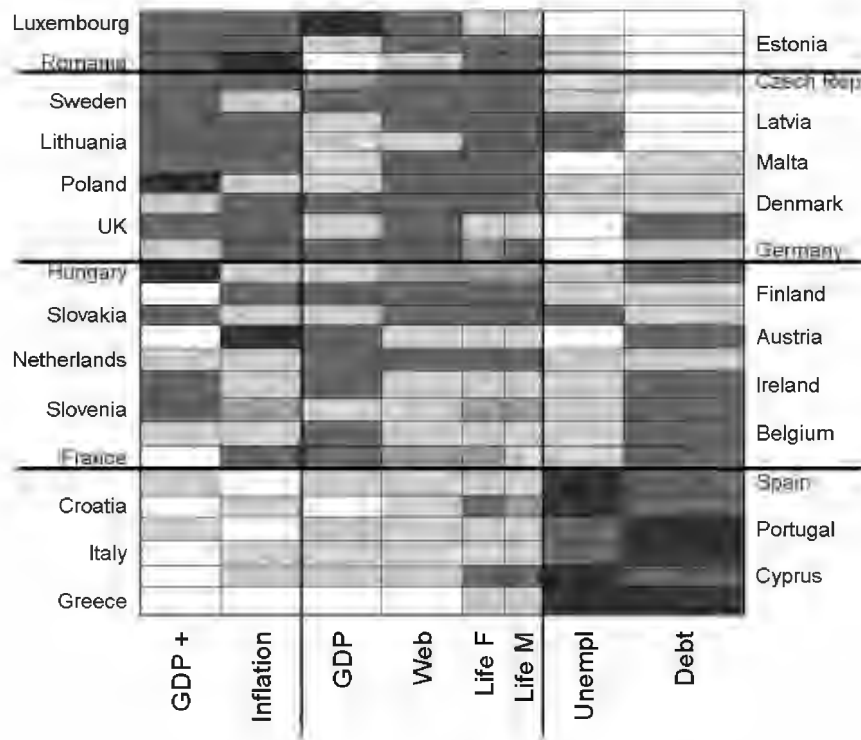


Fig. 23. Overrepresentation map after GCA and the cluster analysis. 27 countries in 2014 (without Bulgaria)

- Czech Republic, which was a member of cluster 1, is now the first country in cluster 2.
- Malta, a former member of cluster 1, is now located in the middle of cluster 2.
- Finland and Austria, former members of cluster 2, dropped to cluster 3.

All countries from former clusters 3 and 4 retained their positions.

The new overrepresentation map is more regular, and characteristics being a base of the identification of clusters are more visible. These are: public debt and unemployment (cluster 3 of characteristics), and GDP growth rate and inflation (cluster 1 of characteristics). Fig. 24 shows that among 27 countries (without Bulgaria) the most outlying country is Luxembourg (as in 2004).

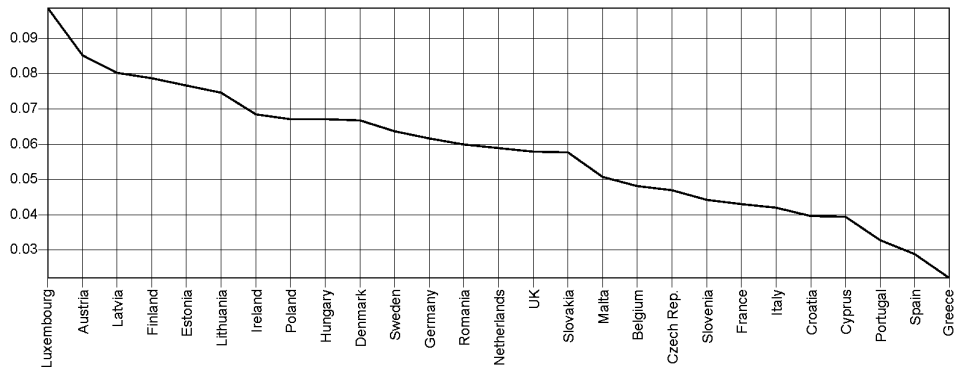


Fig. 24. Outliers for countries. 27 countries in 2014 (without Bulgaria)

5 Conclusion

As we can see, over the past 10 years the European Union has undergone rapid evolution. The main changes are the following:

- The division into old and new members – clearly visible in 2004 – considerably blurred.
- An increasing problem of the EU is high public debt of many Member States. It entails further problems, one of which is unemployment.
- The undisputed EU leader is Luxembourg. It is the only country with high and still growing GDP and low public debt.
- Profile of Greece in 2004 was more similar to profiles of new EU members than to those of old Member States. Thus, current problems of this country could have been foreseen... if there were available reliable data on its economy, and these – as we know – were hidden by Greek governments.

Once again, we emphasize that our research is only our voice in the discussion on the present and the future of the European Union. We are aware of our own limitations and of the limitations of the research method we used. We know that our assumptions (choice of characteristics, their grouping and their weights) have influenced the results obtained. We are also aware that of the utmost importance is the credibility of the data; the example of Greece shows that it is not always guaranteed.

Table 6. Spearman's rank correlation coefficients for characteristics of countries

	GDP	GDP +	Inflation	Debt	Unempl	Life F	Life M	Web
2004								
GDP	1.000000	-0.594276	-0.606384	0.305038	-0.667260	0.716438	0.722404	0.852034
GDP +	-0.594276	1.000000	0.548356	-0.663381	0.353376	-0.634844	-0.750376	-0.625753
Inflation	-0.606384	0.548356	1.000000	-0.291193	0.304234	-0.455392	-0.462529	-0.636376
Debt	0.305038	-0.663381	-0.291193	1.000000	-0.036976	0.605998	0.656626	0.241747
Unempl	-0.667260	0.353376	0.304234	-0.036976	1.000000	-0.235269	-0.406575	-0.565163
Life F	0.716438	-0.634844	-0.455392	0.605998	-0.235269	1.000000	0.829452	0.609154
Life M	0.722404	-0.750376	-0.462529	0.656626	-0.406575	0.829452	1.000000	0.657350
Web	0.852034	-0.625753	-0.636376	0.241747	-0.565163	0.609154	0.657350	1.000000
2009								
GDP	1.000000	0.266191	-0.558722	0.370157	-0.477325	0.706448	0.772179	0.760997
GDP +	0.266191	1.000000	-0.308577	0.500410	-0.315342	0.408860	0.507049	0.005617
Inflation	-0.558722	-0.308577	1.000000	-0.301917	0.057444	-0.614550	-0.488286	-0.208310
Debt	0.370157	0.500410	-0.301917	1.000000	-0.025198	0.500206	0.485835	0.008767
Unempl	-0.477325	-0.315342	0.057444	-0.025198	1.000000	-0.223274	-0.379263	-0.392788
Life F	0.706448	0.408860	-0.614550	0.500206	-0.223274	1.000000	0.848877	0.327522
Life M	0.772179	0.507049	-0.488286	0.485835	-0.379263	0.848877	1.000000	0.420468
Web	0.760997	0.005617	-0.208310	0.008767	-0.392788	0.327522	0.420468	1.000000
2014								
GDP	1.000000	-0.110715	0.376682	0.220487	-0.383203	0.541066	0.709412	0.796789
GDP +	-0.110715	1.000000	0.166392	-0.417853	-0.404739	-0.494585	-0.279139	0.086282
Inflation	0.376682	0.166392	1.000000	-0.253842	-0.701031	0.006319	0.034866	0.523920
Debt	0.220487	-0.417853	-0.253842	1.000000	0.475017	0.595150	0.504585	-0.145050
Unempl	-0.383203	-0.404739	-0.701031	0.475017	1.000000	0.224506	0.005067	-0.595803
Life F	0.541066	-0.494585	0.006319	0.595150	0.224506	1.000000	0.819079	0.249416
Life M	0.709412	-0.279139	0.034866	0.504585	0.005067	0.819079	1.000000	0.436840
Web	0.796789	0.086282	0.523920	-0.145050	-0.595803	0.249416	0.436840	1.000000

References

1. Aggarwal C.C., Reddy C.K. (ed.) (2014) *Data Clustering. Algorithms and Applications*, Chapman & Hall, Boca Raton.
2. Aldenderfer M.S., Blashfield R.K. (1984) *Cluster Analysis, Quantitative Applications in the Social Sciences*, **44**, Sage University Papers, Newbury Park.
3. Checherita C., Rother Ph. (2010) The impact of high and growing government debt on economic growth – an empirical investigation for the euro area, *Working Paper Series, No 1237*, European Central Bank, (<https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1237.pdf>, access: August 26th, 2015).
4. Ciok A. (2004) *On the number of clusters – a grade approach*, Instytut Podstaw Informatyki PAN, Warszawa.
5. Ciok A., Kowalczyk T., Pleszczyńska E., Szczesny W. (1995) Algorithms of grade correspondence – cluster analysis, *The Collected Papers of Theoretical and Applied Computer Science*, **1-4**, Vol. 6, 3-20.
6. Clements B., Bhattacharya R., Nguyen T.Q. (2003) External Debt, Public Investment, and Growth in Low-Income Countries, IMF Working Paper, **WP/03/249**, International Monetary Fund (<https://www.imf.org/external/pubs/ft/wp/2003/wp03249.pdf>, access: August 26th, 2015).
7. Eberhardt M., Presbitero A.F. (2013) Public debt and economic growth: There is no 'tipping point', VOX CEPR's Policy Portal (<http://www.voxeu.org/article/debt-and-growth-no-tipping-point>, access: August 26th, 2015).
8. Fayyad U., Piatetsky-Shapiro G., Smyth P. (1997) From Data Mining to Knowledge Discovery in Databases, in *Advances in knowledge discovery and data mining*, American Association for Artificial Intelligence, Menlo Park, 1-34.
9. Grabowska G., Wiech M. (2009) Grade analysis of data from the European Economic Survey 2005 on Economic Climate in Polish Servicing Sector, in *Control and Cybernetics*, **3**, Vol. 28, 783-810.
10. Greiner A., Fincke B. (2015) *Public Debt, Sustainability and Economic Growth: Theory and Empirics*, Springer International Publishing, Cham – Heidelberg – New York – Dordrecht – London.
11. Jarochońska E., Grzegorek M., Hirny J., Maryja O., Wiech M. (2005) *Analiza danych medycznych i demograficznych przy użyciu programu GradeStat*, Instytut Podstaw Informatyki PAN oraz Instytut Pomnik – Centrum Zdrowia Dziecka, Warszawa.
12. Kowalczyk T., Pleszczyńska E., Rulad F. (2004) Grade Models and Methods for Data Analysis with Applications for the Analysis of Data Populations, *Studies in Fuzziness and Soft Computing*, **151**, Springer Verlag, Berlin – Heidelberg – New York.
13. Korzeniewski J. (2012) *Metody selekcji zmiennych w analizie skupień. Nowe procedury*, Wydawnictwo Naukowe Uniwersytetu Łódzkiego, Łódź.
14. Lenkiewicz St. (2014) Impact of the Common Agricultural Policy on regional differences in Polish agriculture. Attempt to use the Grade Data Analysis, in *"Agricultural budget" and the competitiveness of the Polish agriculture*, **111.1**, 38-66.
15. Lenkiewicz St. (2014) Grade Data Analysis Applied to the European Agriculture, in *Techniki Informacyjne. Teoria i Zastosowania*, **Tom 4(16)**, 56-98.
16. Lenkiewicz St. (2012) Gradacyjna analiza danych – idea i przykład zastosowania, in *Współczesne Problemy Zarządzania*, **1/2012**, 63-98.
17. Romesburg H.Ch. (2004) *Cluster Analysis for Researchers*, Lulu Press, North Carolina.
18. Szczesny W., Kowalczyk T., Wolińska-Welcz A., Wiech M., Dunicz-Sokołowska A., Grabowska G., Pleszczyńska E. (2012) *Models and Methods of Grade Data Analysis: Recent Developments*, Institute of Computer Science, Polish Academy of Sciences.

19. Ząbkowski T., Szczesny W. (2012) Badanie atrakcyjności oferty dostępu do Internetu za pomocą analizy gradacyjnej, in *Metody ilościowe w badaniach ekonomicznych*, Vol. XIII/3, 276-287.

GOSPODARKI KRAJÓW UNII EUROPEJSKIEJ W ŚWIETLE GRADACYJNEJ ANALIZY DANYCH

Streszczenie. W niniejszym artykule dokonujemy analizy zmian, jakie zaszły w Unii Europejskiej w latach 2004-2014, to jest w okresie jej intensywnego poszerzania. Na podstawie zestawu ośmiu cech dokonujemy podziału Unii na skupienia krajów o podobnym profilu, jak również podziału cech na skupienia cech powiązanych. Badaniem obejmujemy dwadzieścia osiem krajów wchodzących obecnie w skład Unii bez względu na to, kiedy stały się jej członkami. Jako metodę badawczą wykorzystujemy gradacyjną analizę danych, a narzędziem przetwarzania danych jest dla nas program GradeStat.

Słowa kluczowe: Unia Europejska, kondycja gospodarcza, analiza skupień, gradacyjna analiza danych, nadreprezentacja, element odstający

IBS PAN *serie*

47 789

Bibl. podręczna

ISBN 83-894-7558-8