



Zeszyty Naukowe  
Wyższej Szkoły Bankowej w Poznaniu  
2013, t. 49, nr 4

---

**Emília Zimková**

Wyższa Szkoła Bankowa w Poznaniu  
Wydział Zamiejscowy w Chorzowie, Katedra Finansów i Bankowości  
e-mail: emilia.zimkova@chorzow.wsb.pl  
tel. 32 349 84 74

## **Reliability of the European Aims till 2020: a comparative cluster analysis**

***Abstract.** The European Commission introduced a Strategy for Europe till 2020, which is focused on smart, sustainable and inclusive growth. In our contribution we would like to answer two research questions: which countries will probably be able to meet the aims proposed by the European Commission by 2020, and which is the best cluster analysis method to be applied in this research from a selection of the most used clustering methods. Single linkage (nearest neighbor), complete linkage (furthest neighbor), and Ward's minimum variance method of cluster analysis had been applied to the issue at hand. Ward's minimum variance method had gotten the best results from the point of view of statistical and economic interpretation. From the 28 European countries, only 8 have a real chance to meet the main criteria of the European Strategy for 2020.*

***Keywords:** cluster analysis, single linkage, complete linkage, Ward's minimum variance method, Strategy for Europe 2020, growth strategy*

### **Introduction**

The European Commission introduced a Strategy for Europe till 2020, which is focused on three reinforcing priorities: the first is smart growth, which demands developing an economy based on knowledge and innovation; the second is sustainable growth, which is based on promoting a more resource efficient, greener and more competitive economy; the third priority is inclusive growth, which fos-



ters a high-employment economy delivering social and territorial cohesion.<sup>1</sup> Are those plans reliable for the majority of European Countries?

In our contribution we would like to answer two research questions: which countries will probably be able to meet the aims proposed by the European Commission by 2020 and which is the best cluster analysis method to be applied in this research from a selection of the most used clustering methods. As such, the paper is a natural continuation of the research conducted by the author who, in a previous article of hers,<sup>2</sup> among other topics, evaluated reliability of the Strategy for Europe till 2020 from the perspective of the Ward's clustering method.

In Europe, twenty seven countries of the European Union prepared their National Reform Programmes as part of the implementation of the European 2020 strategy. Although the main endeavour has to be done in the member states of the European Union, steps towards supporting the European 2020 strategy are needed as well.

The paper is organized in 5 sections, the first of which is an introduction and the last is a conclusion. The second part provides the headline targets of the European 2020 strategy, as well as a comparison to China's National Plan for Medium and Long-term Education Reform and Development for the same period. The third section brings up characteristics of the various cluster analyses which are applied in this contribution. The fourth part contains the application of the chosen cluster analyses.

## 1. The European Aims till 2020 and their Reliability

The Strategy for Europe till 2020 had put forward three mutually reinforcing priorities: smart, sustainable and inclusive growth. Based on these priorities, the European Union defined where it wants to be by 2020. With this purpose in mind, the Commission proposed the following headline targets:

- 75% of the population aged 20-64 should be employed;
- 3% of the EU's gross domestic product should be invested in research and development (R&D);
- the percentage of early school leavers should be under 10%, and at least 40% of the young generation should have a tertiary degree;
- ambitious aims in energy policies;
- 20 million less people should be at risk of poverty.

<sup>1</sup> European Commission, *EUROPE 2020. A strategy for smart, sustainable and inclusive growth*, Brussels 2010, <http://ec.europa.eu/eu2020/pdf/COMPLET%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf> [10.04.2013].

<sup>2</sup> E. Zimková, *Quo Vadis, Euro?*, 7<sup>th</sup> Chorzowska Conference of Banking and Finance, submitted for publishing.

These targets are interrelated and critical to European overall success. Nonetheless, it is doubtful that this can be achieved by the current procedures and “economic management” of Europe. In their research, Roth and Thum<sup>3</sup> state that the Strategy for Europe till 2020 has taken an important step forward by setting the target for tertiary graduation rates at an ambitious 40%. However, many European countries will not be able to meet this benchmark by 2020. On the other hand, China’s education system already produces the same number of graduates with a tertiary education as all the “old 15 countries of the European Union.” Given the large output of graduates, which is the key to productive spending on research and development (R&D), China is likely to soon become a growing power in innovation. The outline of China’s National Plan for Medium and Long-term Education Reform and Development till 2020 is the strategic document which aims to close the gap with the US and the EU15 in terms of innovation potential.<sup>4</sup>

The headline targets of the Strategy for Europe till 2020 might be, in view of many economists, counterproductive in attaining the goals outlined. Bogliacino<sup>5</sup> states that since technical change in economic growth has a different impact on different groups of workers (for example, on skilled and unskilled workers), it may be likely that differences in innovative activities carried out by companies accentuate differences in the labour market. It is necessary to understand which kind of innovations increase, and which decrease, unemployment in economy. The most recent consensus is focused on the traditional Schumpeterian distinction between product and process innovation. The innovation of products is characterized by the discovery of new opportunities, opening up new sectors or market niches, or transforming mature industries. Process innovation is characterized by the long term process of labour-saving technological change. There is evidence that this distinction fits the data suitably at the industrial level.<sup>6</sup>

Over the past decade, member states experienced divergent economic trends, which exacerbated competitiveness gaps and led to macro-economic imbalances within the European Union. To avoid this happening in the future and to strengthen the aims of the Europe 2020, the Commission proposed a new surveillance mechanism. A complementary agenda with additional reforms – called the Euro Plus Pact – has been agreed to among euro area member states, as well as, six

<sup>3</sup> F. Roth, A. E. Thum, *The Key Role of Education in the Europe 2020 Strategy*, CEPS Working Document no. 338, October 2010.

<sup>4</sup> Government of China, *Outline of China’s National Plan for Medium and Long-term Education Reform and Development (2010-2020)*, State Council of the Peoples’ Republic of China, Beijing 2010, [https://www.aei.gov.au/news/newsarchive/2010/documents/china\\_education\\_reform\\_pdf](https://www.aei.gov.au/news/newsarchive/2010/documents/china_education_reform_pdf) [10.04.2013].

<sup>5</sup> F. Bogliacino, *Is smart growth employment friendly?*, Joint Research Centre, 2011, [http://ftp.jrc.es/EURdoc/JRC63123\\_TN.pdf](http://ftp.jrc.es/EURdoc/JRC63123_TN.pdf) [5.04.2013].

<sup>6</sup> F. Bogliacino, M. Pianta, *Innovation and Employment. A Reinvestigation using Revised Pavitt Classes*, “Research Policy” 2010, vol. 39, no. 6, pp. 799-809.

non euro area countries which have decided to sign: Bulgaria, Denmark, Latvia, Lithuania, Poland and Romania. It focuses on four areas: competitiveness, employment, sustainability of public finances, and reinforcing financial stability. The pact was signed by European leaders in March 2011.<sup>7</sup> All 23 signatories are committed to implementing the reforms. The remaining four member states of the European Union are free to sign up if they wish. Moreover, the Fiscal Compact with new fiscal targets for the European countries has been set.<sup>8</sup>

In our research we would like to identify which countries may reach the goals of the Europe 2020, which is why we would like to find the most homogeneous group of countries in respect to the headline indicators of the Europe 2020 Strategy. We use several methods of the cluster analyses for this purpose.

## 2. Cluster analyses

Cluster analyses are based on information found in data that describe the objects and their relationships. The goal is to make the objects within a group similar (or related) to one another, and different from (or unrelated to) the objects in another group. The greater the similarity (or homogeneity) within a group, and the greater the differences between the groups, the better or more distinct the clustering is.

In our research we distinguish various types of clustering: hierarchical (nested) versus partial (unnested), exclusive versus overlapping versus fuzzy, and complete versus partial.

**Hierarchical versus Partial Clustering.** If we permit clusters to have sub-clusters, then we obtain a hierarchical clustering, which is a set of nested clusters that are organized as a tree. Each node (cluster) in the tree is the union of its children (sub-clusters), and the root of the tree is the cluster containing all the objects. For our analysis we have chosen hierarchical clustering.

**Exclusive versus Overlapping versus Fuzzy Clustering.** Exclusive clustering assigns each object to a single cluster. An overlapping or non-exclusive clustering is used to reflect the fact that an object can simultaneously belong to more than one group (class). For example a person at a university can be both an enrolled student and an employee of the university. In a fuzzy clustering, every object

<sup>7</sup> European Council, *Conclusions of the Heads of State or Governments of the Euro Area of 11 March 2011*, Brussels 2011, [http://www.consilium.europa.eu/uedocs/cms\\_data/docs/pressdata/en/ec/119809.pdf](http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/119809.pdf) [10.04.2013]; European Council, *European Council – 24/25 March 2011 – Conclusions*, General Secretariat of the Council, European Council, Brussels 2011, [http://www.consilium.europa.eu/uedocs/cms\\_data/docs/pressdata/en/ec/120296.pdf](http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/120296.pdf) [5.04.2013].

<sup>8</sup> European Council, *Treaty on Stability, Coordination and Governance in the Economic and Monetary Union*, Brussels 2012, [http://european-council.europa.eu/media/639235/st00tscg26\\_en12.pdf](http://european-council.europa.eu/media/639235/st00tscg26_en12.pdf) [10.04.2013].

belongs to every cluster with a membership weight that is between 0 (absolutely does not belong) and 1 (absolutely belongs). For our analysis we have chosen exclusive clustering.

**Complete versus partial clustering.** A complete clustering assigns every object to a cluster, whereas a partial clustering does not. The motivation for a partial clustering is that some objects in a data set may not belong to well-defined groups. Many times objects in the data represent statistical noise, outliers, etc. For our analysis we have used complete clustering.

The clustering methods define the distance between two clusters at each stage of the procedure. In our analysis we have applied single linkage (nearest neighbor), complete linkage (furthest neighbor) and Ward's method.

**Single Linkage (nearest neighbor).** In single linkage the distance between two clusters is the minimum distance between an observation in one cluster and an observation in the other cluster. At every step of clustering, the distance between two clusters is taken to be the distance between their two closest members.

**Complete Linkage (furthest neighbor).** In complete linkage, the distance between two clusters is the maximum distance between an observation in one cluster and an observation in the other cluster. Complete linkage is strongly biased toward producing clusters with approximately equal diameters and can be severely distorted by moderate outliers.

**Ward's minimum variance method.** Ward's minimum variance method is based on the minimalization of the variance of observations within a cluster and on the maximalization of the distance between two clusters. At each generation, the within-cluster sum of squares is minimized over all partitions obtainable by merging two clusters from the previous generation. The sums of squares are easier to interpret when they are divided by the total sum of squares to give the proportions of variance (squared semipartial correlations).

### 3. Methodology of research and data

In our analysis we have compared single linkage (nearest neighbor), complete linkage (furthest neighbor), and Ward's minimum variance method of cluster analysis, which were applied on the database of 28 countries of the European Union to examine the most homogenous groups in respect to the aims which were set by the Europe 2020 Strategy.

We had analyzed a group of 28 countries of the European Union: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovenia, Slovakia, Spain, Sweden, the United Kingdom and Norway. Our aim was to identify the most ho-

mogenous groups in regard to these available indicators: gross domestic product in purchasing power parity, research and development (R&D) expenditure of countries as percent of GDP, unemployment rate, and the unemployment rate of people less than 25 years old. The data was obtained from Eurostat. The pre-crisis period is characterized by data from the year 2007, the crises period is characterized by data from the year 2011. There is no more up-to date data in regard to research and development (R&D) expenditure of countries as percent of GDP. Greece did not report this data in 2011, and that is the reason why this country is excluded from the analysis in the crisis period.

#### 4. Results of cluster analyses

**Ward's minimum variance method.** Ward's minimum variance method applied to the data from 2007 (pre-crisis period) divided the 29 analyzed countries into two main clusters. The first big homogeneous (more prominent) group (in regard to gross domestic product, expenditures to research as percentage of GDP, unemployment rate, and unemployment rate of people younger than 25) consisted of 9 countries: Denmark, Austria, United Kingdom, Ireland, Netherlands, Norway, Luxembourg, Finland, and Sweden. The most homogeneous were found to be Finland – Sweden, Norway – Netherlands – Ireland, and Denmark – Austria. The second big homogeneous (but less prominent) group consisted of 20 countries, and the most homogeneous (from the analyzed point of view) were found to be: Estonia – Slovenia – Cyprus – Lithuania, Latvia – Malta – Bulgaria – Romania, Poland – Croatia – Greece – Slovakia, Spain – Portugal – Czech Republic – Hungary – Italy, Belgium – France – Germany.

Ward's minimum variance method applied to the data from 2011 (crisis period) divided the analyzed countries (without Greece due to the unavailability of data) into two main clusters as well. The first (most prominent) group was joined by Germany and it was left by the United Kingdom and Ireland in comparison with the results from 2007. Under the influence of the crisis the prominent group dropped to 8 countries: Luxembourg, Norway, Germany, Austria, Netherlands, Finland, Sweden, and Denmark. The most homogeneous were Germany – Austria – Netherlands, and Finland – Sweden. In reference to the 2011 indicators, the second biggest homogeneous (but less prominent) group consists of 20 countries, and the most homogeneous (from the analyzed point of view) were found to be: Latvia – Lithuania – Slovakia – Croatia, Spain, Hungary – Poland – Bulgaria, Cyprus – Romania, Malta, Belgium – United Kingdom – France, Czech Republic – Slovenia, Italy – Portugal – Ireland, Estonia.

The cluster analysis proved that the group of eight prominent economies underwent structural changes and fought against the crisis successfully, experi-

Chart 1. Dendrogram of 29 European countries (2007), Ward Linkage

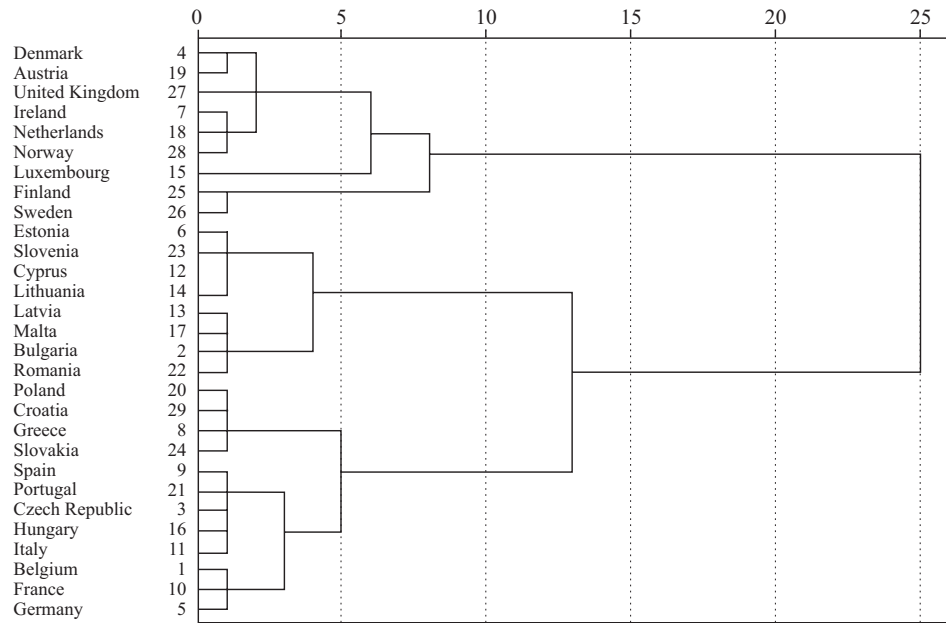
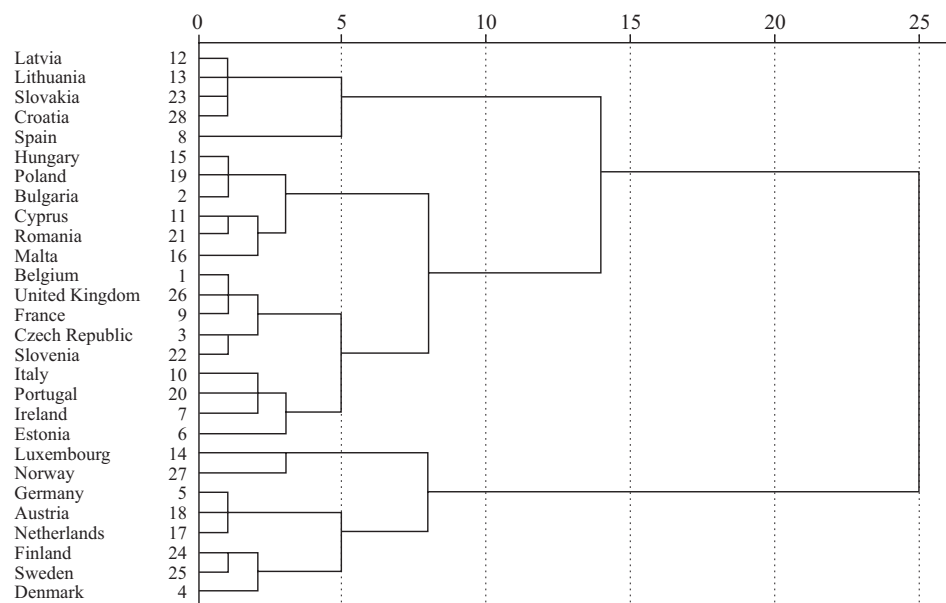


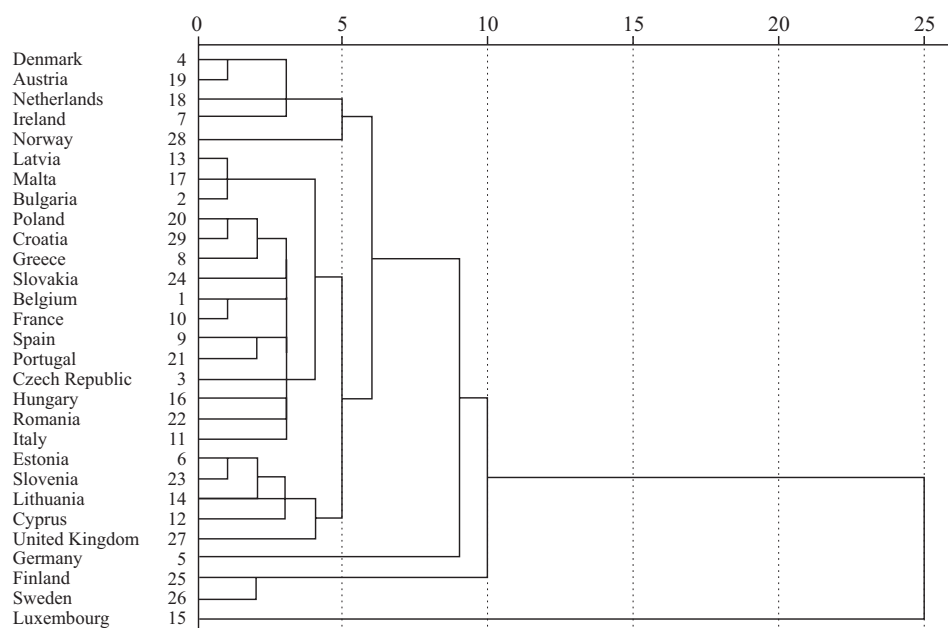
Chart 2. Dendrogram of 28 European countries (2011), Ward Linkage



encing economic growth, low inflation, and a low rate of unemployment. These countries might be able to reach the aims proposed by the European Commission by 2020. Based on the cluster analysis, the core group of the euro zone could consist of Germany, Austria, Netherlands, and Finland. The less prominent group of countries (especially the euro zone members) should continue in implementation of structural changes and strict macroeconomic policies.

**The single linkage method.** The single linkage method applied to the data from 2007 (pre-crisis period) appeared not to be a proper cluster method for our analysis. It highlights the outlier position of Luxembourg. The clustering based on the proximity between the closest two points, that are in different clusters, yields non-homogeneous groups from the view point of the Europe 2020 Strategy. It shows that the closest countries in the view-point of the analyzed indicators are Denmark & Austria, Latvia, Malta & Bulgaria, Poland & Croatia, Belgium & France, Estonia & Slovenia.

Chart 3. Dendrogram of 29 European countries in Pre-Crisis time (2007), Single Linkage



The single linkage method applied to the data from 2011 (crisis period) proves that this method is not appropriate for our research. The worst performance in the Europe 2020 Strategy using this method was indicated to be Spain.

The single linkage method does not result in a greater difference between groups, which is the reason why it is not appropriate for our research.



Chart 4. Dendrogram of 28 European countries (2011), Single Linkage

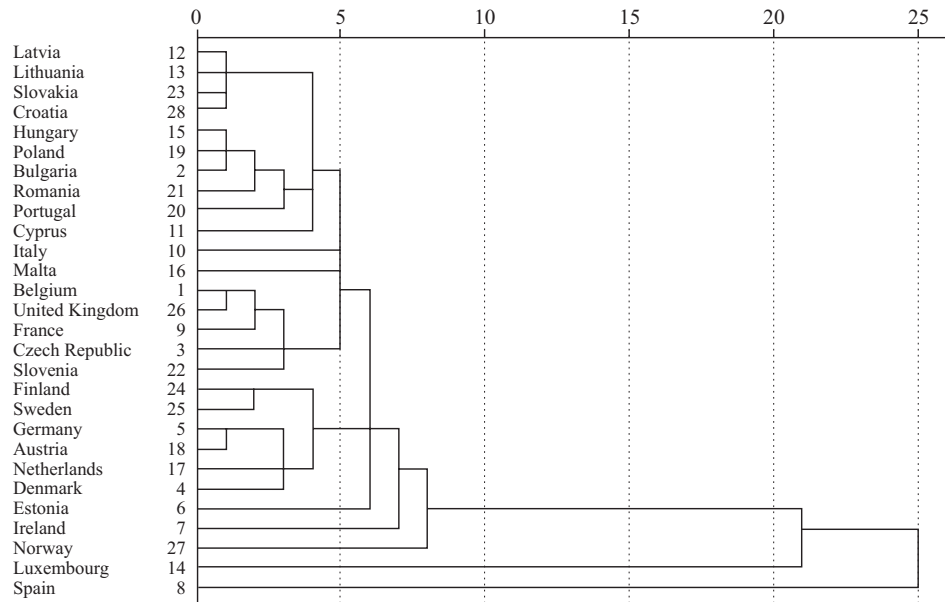
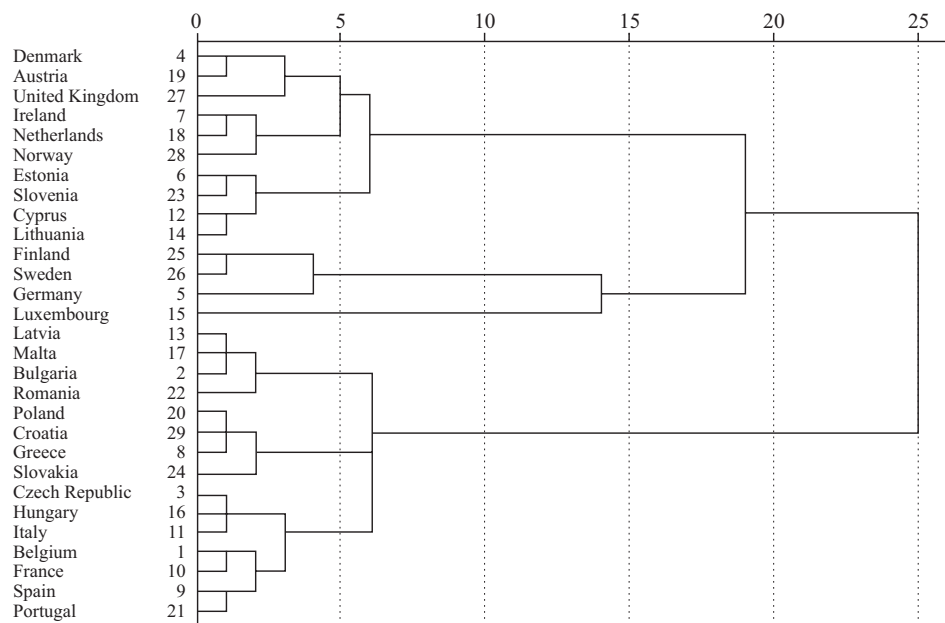


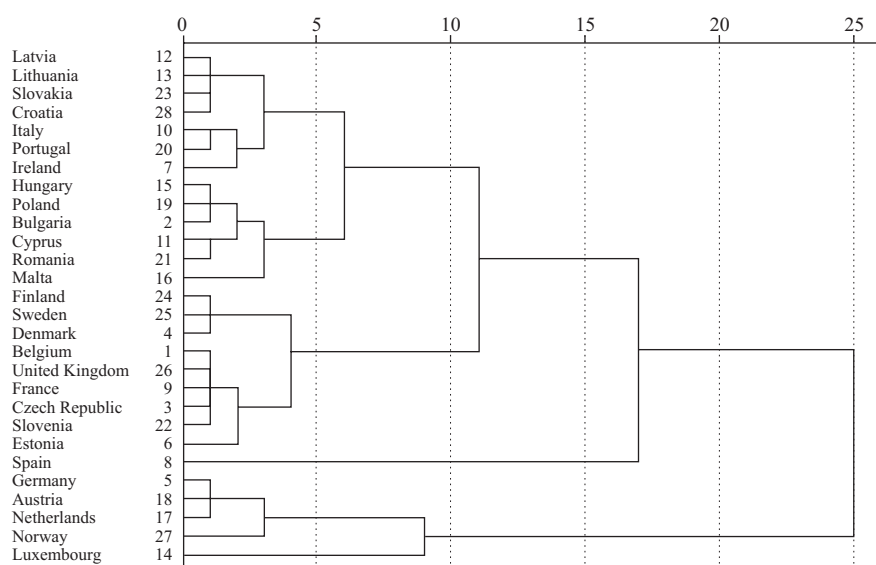
Chart 5. Dendrogram of 29 European countries (2007), Complete Linkage



**The complete linkage method.** The complete linkage method applied to the data from 2007 (pre-crisis period) divided countries into two main groups. Both groups consist of a mixture of the new and old countries of Europe. From the clustering point of view, this method is better than the single linkage. From the point of view of economic interpretation, this method is not better than the Ward's clustering method.

The complete linkage method applied to the data from 2011 (crisis period) loses its statistical information (visualisation of the 28 analyzed countries into two main clusters).

Chart 6. Dendrogram of 28 European countries (2011), Complete Linkage



It is proved that the best statistical and analytical information can be achieved by applying the Ward's cluster method.

## Conclusion

In our research we distinguish between various types of clustering: hierarchical (nested) versus partial (unnested), exclusive versus overlapping versus fuzzy, and complete versus partial. Based on the character of the research data, we have applied hierarchical, exclusive, and complete types of clustering.

As mentioned above, the greater the similarity (or homogeneity) within a group and the greater the difference between groups, the better or more distinct the clustering is. Having this main rule in mind we can say that when it comes to single

linkage (nearest neighbor), complete linkage (furthest neighbor), and Ward's minimum variance method, the best explanatory results were achieved by the Ward's minimum variance method of clustering for the data analyzed in this paper.

The dendrogram created from Ward's minimum variance method of clustering – which takes into consideration gross domestic product in purchasing power parity, expenditures to research and development, unemployment rate, and unemployment rate of people less than 25 years – divides 28 European countries into two big clusters. The better group consists of 8 countries where the best results are achieved by Luxembourg and Norway, followed by the group consisting of Germany, Austria and the Netherlands. Also found in this better group of analyzed countries are the Scandinavian countries of Finland, Sweden and Denmark. Countries in this better group pay more attention to education, and research and development in the long term. Their economies underwent structural changes and they are capable to fight with the crisis successfully, thus experiencing economic growth, low inflation and a low rate of unemployment.

The second group of European countries is more heterogeneous. It includes the Baltic states, Visegrad countries, and the countries of Middle and Southern Europe. The majority of them experienced a high level of unemployment and slow economic growth or depression during the crisis time of 2011.

## Literature

- Bogliacino F., *Is smart growth employment friendly?*, Joint Research Centre, 2011, [http://ftp.jrc.es/EURdoc/JRC63123\\_TN.pdf](http://ftp.jrc.es/EURdoc/JRC63123_TN.pdf) [5.04.2013].
- Bogliacino F., Pianta M., *Innovation and Employment. A Reinvestigation using Revised Pavitt Classes*, "Research Policy" 2010, vol. 39, no. 6.
- European Commission, *EUROPE 2020. A strategy for smart, sustainable and inclusive growth*, Brussels 2010, <http://ec.europa.eu/eu2020/pdf/COMPLET%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf> [10.04.2013].
- European Council, *Conclusions of the Heads of State or Governments of the Euro Area of 11 March 2011*, Brussels 2011, [http://www.consilium.europa.eu/uedocs/cms\\_data/docs/pressdata/en/ec/119809.pdf](http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/119809.pdf) [10.04.2013].
- European Council, *European Council – 24/25 March 2011 – Conclusions*, General Secretariat of the Council, European Council, Brussels 2011, [http://www.consilium.europa.eu/uedocs/cms\\_data/docs/pressdata/en/ec/120296.pdf](http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/120296.pdf) [5.04.2013].
- European Council, *Treaty on Stability, Coordination and Governance in the Economic and Monetary Union*, Brussels 2012, [http://european-council.europa.eu/media/639235/st00tscg26\\_en12.pdf](http://european-council.europa.eu/media/639235/st00tscg26_en12.pdf) [10.04.2013].
- Government of China, *Outline of China's National Plan for Medium and Long-term Education Reform and Development (2010-2020)*, State Council of the Peoples' Republic of China, Beijing 2010, [https://www.aei.gov.au/news/newsarchive/2010/documents/china\\_education\\_reform\\_pdf](https://www.aei.gov.au/news/newsarchive/2010/documents/china_education_reform_pdf) [10.04.2013].
- Roth F., Thum A.E., *The Key Role of Education in the Europe 2020 Strategy*, CEPS Working Document no. 338, October 2010.

- Tan P.-N., Steinbach M., Kumar V., *Introduction to Data Mining*, Pearson Education, London 2005.
- Ward J. H. Jr., *Hierarchical Grouping to Optimize an Objective Function*, "Journal of the American Statistical Association" 1963, no. 58.
- Zimková E. *Quo Vadis, Euro?* 7<sup>th</sup> Chorzowska Conference of Banking and Finance, submitted for publishing.

### Osiągalność celów europejskich do 2020 r. – porównawcza analiza skupień

**Streszczenie.** Komisja Europejska wprowadziła długookresową strategię Europa 2020 na rzecz inteligentnego i zrównoważonego rozwoju, sprzyjającego społecznemu włączeniu. W niniejszym artykule podjęto próbę odpowiedzi na dwa pytania: które kraje europejskie prawdopodobnie będą w stanie osiągnąć do roku 2020 cele założone przez Komisję Europejską oraz która spośród najpowszechniej stosowanych metod analizy skupień najlepiej posłuży do zbadania tej kwestii. Sięgnięto po trzy metody grupowania hierarchicznego: pojedynczego wiązania (najbliższego sąsiedztwa), pełnego wiązania (najdalszego sąsiedztwa) oraz metodę minimalnej wariancji Warda. Okazało się, że z punktu widzenia interpretacji statystycznej i ekonomicznej najlepsze rezultaty daje metoda Warda. Za jej pomocą ustalono, że realne szanse osiągnięcia głównych kryteriów strategii Europa 2020 ma tylko 8 na 28 państw europejskich poddanych analizie.