5. International Anatolian Scientific Research Congress 21-23 JULY 2023, Hakkari, Türkiye

# **Measuring homogeneity of countries in the European Union based on similarity analyses**

Dániel Váradi (<https://orcid.org/0000-0001-9610-8566>)

László Pitlik (<https://orcid.org/0000-0001-5819-0319>),

e-Mails: danielvaradi140@gmail.com, pitlik@my-x.hu,

Kodolányi János University and MY-X research team Hungary

*Keywords: homogeneity, similarity, modelling, OAM, data science, EU, artificial intelligence*

**Abstract**

Present status: Increasing integrity plays for the European Union a significant role since it was established back in 1950. Also, it is to be found in the Single European Act where the article 130.a defines Community shall develop and pursue its actions leading to the strengthening of its economic and social cohesion (c.f. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:11986U/TXT>). To foster each country on cohesion/integrity, the leadership of the EU decides on millions of euros to be allocated and used for integration purposes. Normally, these goals are also agreed on beforehand and it aligns with the future strategy of the European Union.

To ensure that financial aid contributed accordingly across countries, there are numerous ways to calculate the degree of integration, which were developed by well-known economists. These methods are heavily relying on the GDP and the produced volume of buying and selling in a country rather than focusing on the characteristics of the raw dataset in comparison to other countries.

Goals/Tasks: The aim of this study is to demonstrate a lot of alternative ways how the homogeneity of European countries can be measured in a holistic and automated way, moving away from GDP- and volume-based measuring, but taking the characteristics of time-series statistical data into consideration in an objective way. The research needs to answer one main question among others: Can different countries/regions of the EU have the same homogeneity index but in other way? Another important goal is to validate and eliminate potential fake-news in online or written newspapers concerning especially Hungary and Germany.

Solution: To assess raw data characteristics, five time-series data layers has been “randomly” extracted from the OECD (The Organization for Economic Cooperation and Development) database and used between 1995-2021 and includes 22 countries for the EU. These five layers are average yearly wages, average yearly working hours, GDP, average life expectancy and unemployment rate. Additional calculations e.g. averages, standard deviations, relative distances from averages has been calculated for every year in order to build an OAM (Object-Attribute-Matrix) as the input database for COCO Y0 (c.f. <https://miau.my-x.hu/myx-free/index.php3?x=e091>).

Conclusions: During the naive analysis approach (where no artificial intelligence is included for optimizing the objectivity) it is close to impossible to deliver a ranking between countries due to the high number of considerable effects (over 550 graphs) of raw data characteristics. With using our artificial intelligence solution for antidiscrimination models, it is possible to form rankings between countries and regions objectively and with that our algorithm can function as decision making programme for the EU.

In terms of countries which joined the EU in 2004, we identified that Hungary homogenised the least, meanwhile Slovenia made the most efforts based on the characteristics of their statistical data. Together with all examined 22 EU countries Hungary is on the 17thplace meanwhile Slovenia is the number one of the rankings. Based on this, it can be said that those countries which joined the EU back in 2004, has made extremely different pathways objectively.

Regarding homogeneity, the essay thus presents coexisting, possible interpretations, where individual interpretation is always riskier than the application of a collective hermeneutic interpretation which aims for consistency.

Also, there is no official definition of homogeneity as per nowadays as there is no clear (mathematical/algorithmic) definition of homogeneity. Nowhere is there a “fever curve” showing the dynamics of homogeneity for different country and the EU as such, i.e., the EU has not yet reached the level of cybernetics that works with dashboards like a 'spaceship'.