



## **Justyna Wilk**

Uniwersytet Ekonomiczny we Wrocławiu  
Wydział Ekonomii, Zarządzania i Turystyki  
Katedra Ekonometrii i Informatyki

*justyna.wilk@ue.wroc.pl*

# **Cluster analysis of unemployment using histogram panel data**

Panel data refers to data sets consisting of the multiple time series (years, quarters etc.) observations on multiple cases (e.g. households, countries etc.). The cross-sectional information reflects the differences between subjects, e.g. one can estimate the rate of unemployment at a particular point in time. The time-series information reflects the changes within subjects over time, e.g. repeated cross-sections can show how this proportion changes over time [1].

The situation is more complicated when we examine the changes of the unemployment structures for a set of territorial units over time. A solution can be found in the methodology of Symbolic Data Analysis. SDA uses multivariate data analysis methods to proceed untypically structured data sets. Its application in the comparative unemployment study based on histogram data set was presented in [5].

The aim of this paper is to carry out a comparative study of unemployment using SDA to find its itineracy. The scope of research covers the situations of almost 380 Polish districts in the years 2003-2014. The study examines the changes of unemployment structure regarding the periods of unemployment [4]. It is based on a histogram panel data set. Series of consecutive cluster analyses are conducted to distinguish homogeneous clusters of districts and to find regularities over the years [3].



**Keywords:** panel data, longitudinal data, cross-sectional time series data, symbolic data analysis, histogram data, unemployment, cluster analysis

#### References

1. Baltagi B.H., 2008, *Econometric analysis of panel data*, John Wiley & Sons, Chichester.
2. Bock H.-H., Diday E. (Eds.), 2000, *Analysis of symbolic data. Explanatory methods for extracting statistical information from complex data*, Springer Verlag, Berlin-Heidelberg.
3. Fred A.L.N., Jain A.K., 2005, *Combining multiple clusterings using evidence accumulation*, IEEE Transaction on Pattern Analysis and Machine Intelligence”, Vol. 27, Issue 6, pp. 835-850, DOI: 10.1109/TPAMI.2005.113.
4. Layard R., Nickell S., Jackman R., 2005, *Unemployment: Macroeconomic Performance and the Labour Market*, Second Edition, Oxford University Press, Oxford.
5. Wilk J., 2011, *Taksonomiczna analiza rynku pracy województw Polski – podejście symboliczne*, *Ekonometria* 34. Zastosowania metod ilościowych, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 200, pp. 26-37.