UE Uniwersytet Ekonomiczny
we Wrocławiu

Instytut Informatyki Ekonomicznej
Prof. dr hab. Jerzy KORCZAK

IT in Management
Assignment 1

Submission deadline and oral presentation: May 13, 2018

The aim of the Assignment is to master the skills of the use of financial and economic databases in data analysis projects. Your task is to discover the clusters of the EU countries based on selected macroeconomic indicators and comment the position of Poland in 2015-2017.

To reduce the work the data are provided and available on my Web page. But if you prefer the data can be found in the Passport database (http://www.bg.ue.wroc.pl). The selection of the macroeconomic indicators (named Categories in the Passport database) should be argued. Please note that to do this part of the work, you should use the computers or your laptop ONLY in the UE library.
To download the statistics you may also use the EU Open Data Portal (http://open-data.europa.eu/en/apps).
The second part of the work should be done using the Orange data mining platform. To process the extracted data you should first transform them and store into native Orange's data format ".tab". Native format starts with feature (attribute) names, their type (continuous, discrete, string). Note that Orange can read files in native and other data formats (.txt, .csv,…).
To cluster the data use one of the available algorithms such as Hierarchical clustering, k-Means, SOM. Visualize and comment your results of clustering.

Instructions and hints:

A. *Become familiar with the use of the Passport database*
B. *Use Excel to transform the raw data extracted from the Passport database*
C. *Install and become familiar with the use of the Orange software* (http://orange.biolab.si/).
D. *Submit a report-presentation as zipped file named* **ITM-your-name.zip** by email to:
   jerzy.korczak@ue.wroc.pl
   This is a comprehensive presentation of your work using MS PowerPoint. You should provide the following information and data: experiment description, description of indicators (categories), data sets used, data mining schema, visual representations of the results, conclusion of the experiments.

References:
– Lecture notes "IT in Management" (http://www.korczak-leliwa.pl)
– Passport User Guide, docs and tutorials
– Orange Tutorial and documentations