

A new method of identifying key industries

Dimitra Adami 1067738 Filippas Mitsos 1019913 Eleni Bonatsou 1067623

University of Patras
MSc Applied Economics and Data Analysis

This article is related to the analysis of industries and their grouping by different industries of the USA for 4 specific years with main PCA technique, i. e. principal component analysis and at the end we observe what effects they will have on the economy of the country, the data to be used are the inputs - outputs of the industries, not including the housing and household industries, so the board will have dimensions 70x70 . The purpose of all these methods is to reduce the tables, i. e. to squeeze the entire economy into fewer sectors. The statistical methods used, in addition to principal component analysis (PCA), used k-means and Silhouette procedures to group industries into similar industries. For k-means clustering, validity indices that are assumed to be independent of the clustering algorithms must be used. Pca is an efficient dimensionality reduction technique that constructs features by averaging linear combinations of the original features, where it classifies industries into distinct groups and well-defined industries. This is achieved first by dendrogram and then by cluster analysis. The presentation of the networks reveals the interconnections between industries within the clusters and the hierarchical positions of the clusters and their interconnections. A dendrogram is a hierarchical tree graph showing a grouping of industries into distinct clusters. The length of each branch in the graph measures the distance between industries in the cluster. The purpose is to decide on the appropriate number of clusters. For this, we use the aggregation method, which creates a hierarchy of industries that start from all of them as if they were completely separate and then merge them until only one cluster remains. Cluster analysis starts with the selection of a distance measurement and optimization procedure for each industry. Using principal component analysis (PCA), we separate the impact of the top two (at most three) eigenvalues, which is equivalent to changes in the rate of return. As a result, the PCA , further refines this ranking by placing industries into distinct and well-defined groups. In this way, to extend the identification of key industries in new directions. The next step is to use the top two computers perpendicular to each other, which means that their correlation is zero, so we cluster the data into three specific groups for each of the four distant years of our study. Finally, by using the statistical methods mentioned above and by using the analytical data of the inputs and outputs, the appropriate changes in the economy will be achieved and will have a better long-term effect on industrial policy.