

Kalinová, Eva

Article

Artificial intelligence for cluster analysis: Case study of transport companies in Czech Republic

Journal of Risk and Financial Management

Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Kalinová, Eva (2021) : Artificial intelligence for cluster analysis: Case study of transport companies in Czech Republic, Journal of Risk and Financial Management, ISSN 1911-8074, MDPI, Basel, Vol. 14, Iss. 9, pp. 1-36, <https://doi.org/10.3390/jrfm14090411>

This Version is available at:

<https://hdl.handle.net/10419/258515>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Article

Artificial Intelligence for Cluster Analysis: Case Study of Transport Companies in Czech Republic

Eva Kalinová

School of Expertness and Valuation, Institute of Technology and Business in České Budějovice, Okružní 517/10, 37001 České Budějovice, Czech Republic; kalinova@mail.vstecb.cz; Tel.: +420-380-070-218

Abstract: What is the situation of the transport sector in the Czech Republic and what is its importance for the economy of the Czech Republic? How and to what extent do businesses operating in this sector influence the sector as such, and how many businesses in this sector have such influence? Additionally, what happens if the most important businesses in the transport sector go bankrupt, and which businesses are the most important ones? Searching for the answers to these questions is a subject of this contribution, focusing primarily on the cluster analysis using artificial neural networks (ANN), specifically with Kohonen networks, which represent the main method for processing a large volume of not only accounting data on transport companies. In this research, the dataset consists of the financial statements of transport companies for the years 2015–2018. The research part of the contribution deals mainly with the issue of the transport sector's development in the years 2015–2018 with the companies operating in this sector and tries to identify the most important companies in terms of their importance for this sector. The results show that the whole transport sector is influenced mainly by the two largest companies, whose potential changes can affect companies themselves but to a great extent also the development of the whole transport sector. For the two companies, financial analysis is carried out using ratios, whose results show that despite the negative values of the important value generators of one of these companies, the company is still able to significantly influence the situation in the transport sector of the CR. This information is a clear guide for experts, development analysts, to determine the further development of the whole sector when focusing on the development of the two specific companies only. A question arises as to how the created model can be applied to other economic sectors, especially in other EU countries.

Keywords: artificial intelligence methods; Kohonen networks; cluster analysis; transport sector; business value generators; big data



Citation: Kalinová, Eva. 2021. Artificial Intelligence for Cluster Analysis: Case Study of Transport Companies in Czech Republic. *Journal of Risk and Financial Management* 14: 411. <https://doi.org/10.3390/jrfm14090411>

Academic Editor: Jakub Horak

Received: 30 June 2021

Accepted: 16 August 2021

Published: 1 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In today's world of fast-developing technologies and innovations, the need for thinking ahead and being one step ahead is becoming stronger. This can be facilitated by big datasets with collected data (sometimes also called big data), which can be analyzed and used to predict further development or to assess the current state. Data processing and evaluation can be carried out using a plethora of methods, whose application may help to find results to provide an outline for a future situation or to adopt measures that may enable users to prevent unwanted events that could affect the development either of the whole sector or one business only. Transport companies are not an exception; their data enable the analysis and assessment of the whole transport sector of the CR. However, data are currently very valued, and their collection is a long-term and complex process. Data can be analyzed using two methods: traditional statistical methods or advanced data processing methods, such as artificial intelligence or artificial neural networks. Compared to other statistical methods, a big advantage of neural networks is their ability to learn. (Cho et al. 2009) believe that artificial intelligence methods enable achieving much better results than traditional statistical methods. Very accurate and complex results are provided also by

hybrid models based on the combination of statistical and artificial intelligence methods. However, (Li et al. 2015) point to the fact that traditional statistical methods are still popular due to their easy interpretation, comprehensibility, and acceptable predictive performance. According to (Rowland and Vrbka 2016), the disadvantages of artificial neural networks (ANN) are the demand for high-quality data and possible illogical behavior of an artificial neural network. (Vochozka and Machová 2017) state that ANN differs from other traditional methods, especially in the so-called adaptation phase, where the neural network learns from appropriately selected training models that represent a given problem. The ability to learn is the biggest advantage of ANN. The volume of the data and the possibilities that the application of ANN or artificial neural networks bring make them a suitable tool for analyzing data of transport companies. The most suitable tool for cluster analysis is the so-called Kohonen networks, sometimes also referred to as self-organizing maps. These networks are used for analyzing the data of transport companies in the CR, which form a basic dataset for the network, and based on which the economic development of the transport sector in the CR can be analyzed and interpreted. Transport companies and the transport sector are a very important part of the economy, not only in the Czech Republic. The analysis of these companies provides real results, which can be used by the transport sector to improve future development and which enable focusing on certain aspects they are most affected by, paying particular attention to them, and take necessary measures. The objective of this contribution is to evaluate the development and the situation in the transport sector based on the data resulting from the analysis performed using Kohonen self-organizing maps and to determine how many companies and on what basis they most influence the transport sector. The analysis will also show the percentage of companies whose influence is low. The assumed share of companies significantly influencing the transport sector and the share of companies with small influence on the transport sector is based on the Pareto principle (80–20 rule), i.e., twenty percent of companies with a major influence on the transport sector to eighty percent of companies, which have a small influence.

2. Literature Research

Like many other infrastructure-intensive economic activities, the transport sector is an important part of the economy, which influences the development and well-being of the population. When efficient, transport systems provide economic and social opportunities and benefits that lead to positive multiplier effects, such as better access to markets, employment rate, and other investments. When inefficient in terms of their capacity or reliability, they might relate to economic costs, such as decreased or lost opportunities (Rodríguez et al. 2013). Transportation of people, information, or goods can be considered a very important component of the national economy of the CR. According to (Svobodová et al. 2013), the extraordinary importance of transport in the CR results from the location of the CR at the crossroads of trans-European routes; in the CR, of course, inland transport also plays a great role, especially in the import and export of raw materials and products and of citizens and in satisfying the need for inland and international passenger transport. As in other sectors, even in the transport sector, it is necessary to analyze and monitor its development. A tool for analyzing the data of transport companies in the CR as well as the entire transport sector can be both traditional statistical and advanced research methods.

(Rayala and Kalli 2021) state that currently, there are various methods (including advanced ones) for processing large volumes of data. The authors believe that deep learning could be a strong paradigm for analyzing large datasets; however, it requires many samples for model training, which is costly and time-consuming. This can be avoided by using a fuzzy approach. (Rayala and Kalli 2021) proposed and developed an improvised Fuzzy C-means (IFCM), which includes the model of Convolutional Neural Network (CNN) and Fuzzy C-means (FCM) for improving the clustering mechanism. In addition, a comparative analysis was performed for each dataset, which showed that IFCM surpasses the existing model. Cluster analysis (clustering) is a major problem in machine learning

without a teacher. It can be applied in many fields such as bioinformatics, gene sequencing, market research, medicine, analysis of social networks, and recommender systems. The main purpose of clustering is to arrange similar data objects from a given dataset into clusters (groups). Clusters usually represent a kind of real-world entity or meaningful abstractions (Lakhawat and Somani 2017). (Bakoben et al. 2017) used cluster analysis of behavior of credit card accounts to assess the level of credit risk. The credit card account was parametrically modeled, and subsequently, behavioral cluster analysis was applied using a recently proposed dissimilarity measure of statistical parameters of the model. (Bakoben et al. 2017) considered only the popular k-medoids clustering method; however, they state that a problem to be solved in the future is the computational complexity of the new metrics growing with the size of the sample. (Tyukhova and Sizykh 2019) studied the possibility of applying cluster analysis for creating preliminary groups of securities based on which the portfolio will be created subsequently. The research results showed that the application of the cluster analysis of shares as a preparatory phase for creating investment portfolios enables increasing its profitability and efficiency and reducing risks. It is more efficient to apply cluster analysis as a preliminary phase with a subsequent portfolio optimization using the Markowitz model. (Cahyana et al. 2020) used hybrid cluster analysis for the classification of customers of PT Pelindo I based on their satisfaction with the services offered by PT Pelindo I. Cluster analysis was performed for the purposes of grouping the research objects based on their characteristic similarities. The results showed that 72% of PT Pelindo I customers perceived the services provided by PT Pelindo I as special, while the remaining 28% perceived this service as good. (Feranecova et al. 2016) used cluster analysis to classify the companies in different financial situations operating in automotive into groups with similar characteristics, thus creating clusters of companies in a similar financial situation. Cluster analysis provided a methodology for determining the financial health of companies operating in automotive. In the study by (Vahalík and Staníčková 2016), factor analysis was carried out prior to cluster analysis. Key factors of foreign trade competitiveness were defined by means of factor analysis; subsequently, cluster analysis was used to identify the countries with similar characteristics of competitiveness factors. (Abdelkafi et al. 2018) carried out the classification of 42 developed and developing countries based on the impact of risk crises by means of Kohonen networks (SOM). The results of SOM (self-organizing maps) indicate the optimal map represented by 56 micro-classes (cells or neurons) and five areas of grouped countries, which is defined by five different economic situations. The classification was based on the level of development, development of economic ratios, and period. (Kohonen 1986) describes artificial neural networks as undetermined classification algorithms based on neural networks, consisting of competitive layers, which use Kohonen classification rules for input classification. Neurons in competitive layers learn to distinguish the groups of similar input vectors. This is a self-organizing method of mapping, which has two layers only (input and output) and each layer is made up of neurons. At the same time, the application of the Kohonen network with database clustering enables the projection of multidimensional data into two-dimensional space and the analysis of the resulting cluster system. Choosing cluster numbers is based on the calculation of cluster indicators, unlike the traditional statistical approach used by (Liashenko et al. 2018). A clear choice for the analysis and processing of the dataset containing the data of transport companies in the form of financial statements are Kohonen ANN (artificial neural networks), which are able to work with big data.

3. Materials and Methods

3.1. Dataset

A fundamental step for achieving the objective set is the collection and subsequent modification of data. The analysis used the data obtained from financial statements, specifically the balance sheet and profit and loss account (PLA). The data source for the research sample was the data obtained from Bisnode's Albertina database. The data

were obtained from section H—Transportation and Storage according to the CZ NACE classification. All legal persons from section H for the period 2015–2018 were selected.

After the generation of the data, the dataset consisted of 21,882 data rows, where each row contained the following items:

1. Specification of the company: company registration number (IČO), name of the company;
2. Information on the company: date of resources, beginning and end of the period, number of months of the financial statement;
3. Financial statements: balance sheet, profit and loss list (PLL).

The data were modified in MS Office Excel as follows:

- a. The value “0” was added in empty cells: the table containing the generated data was selected, and “Replace” was used to replace fill in all empty cells in the table with the value 0.
- b. It is necessary to add another column “EBIT” by adding the column “EBT” and “Interest payable”.
- c. The beginning and end of the period outside the period of 2015–2018 were removed.
- d. The columns and rows containing zero entries, the columns with zero variance, and duplicate rows were removed (“Data”—“Remove duplicates”—“by company registration number”). Subsequently, the entities (in rows) for which the data for the period other than 12 months (different accounting periods) needed to be removed, as well as the non-numeric entries.
- e. ROA and ROE were calculated, where ROA is expressed as a ratio of EBIT and assets and ROE is the ratio of EAT and equity.
- f. All necessary components for calculating EVA Equity (from the point of view of shareholders) was calculated: risk-free return— r_f ; indicators characterizing the company size— r_{LA} ; indicators characterizing the production power— $r_{entrepreneurship}$; XP ; indicators characterizing the relationship between assets and liabilities— $r_{finstab}$; and weighted average costs of capital—WACC (risk-free return + indicators characterizing the company size + indicators characterizing the production power + indicators characterizing the relationship between assets and liabilities) according to (Vochozka Marek 2020):

$$WACC = r_f + r_{LA} + r_{entrepreneurship} + r_{FinStab} \tag{1}$$

Alternative costs of equity— r_e according to (MPO 2016):

$$r_e = \frac{WACC * \frac{UZ}{A} - (1 - d) * \frac{U}{BU+O} * \left(\frac{UZ}{A} * \frac{VK}{A} \right)}{\frac{VK}{A}} \tag{2}$$

where

UZ—financial resources (equity and interest-bearing debt capital);

A—assets;

VK—equity;

BU—bank loans;

O—bonds;

$\frac{U}{BU+O}$ —interest rate, can be marked i (interest);

d —income tax rate (can be marked $t-tax$).

Financial resources—UZ (equity + issued long-term bonds + issued short-term bonds + bank loans and financial assistance), rate of corporate tax— d (according to a relevant year—since 2010, the rate has been 19%).

- g. After calculating all necessary components, EVA Equity was calculated as follows:

$$EVA\ Equity = (ROE - r_e) * VK \tag{3}$$

- h. It was necessary to carry out the so-called data cleansing by removing nonsensical and extreme values; subsequently, non-numerical entries of EVA Equity needed to be removed, e.g., those with the error message claiming that it is not possible to divide by zero.

In order to keep entities (enterprises) which were further processed in the dataset, the enterprises needed to meet the following 29 conditions: positive assets, positive fixed assets, positive fixed financial assets, positive fixed intangible assets, positive current assets, positive inventories, positive long-term receivables, positive short-term receivables, positive trade receivables, positive receivables to associates, positive share capital, positive reserve funds, positive reserves, positive cash, positive sales of goods, positive consumed material, positive production–consumption, positive performance, positive costs of providing goods, positive depreciation, positive sales of fixed assets, positive sales of material, positive net book value of sold fixed assets, positive interest payable, wage costs higher than CZK 120 thousand per year, ROA in the interval $(-100\%, +100\%)$, ROE in the interval $(-100\%, +100\%)$, alternative costs of equity in the interval $(0\%, +100\%)$, and sales of goods and performance together at least CZK 120 thousand per year.

3.2. Methods

As regards the methods for processing and analyzing the data used in this work, it is necessary to focus primarily on the Kohonen networks. To clarify self-organizing processes, their operation is first demonstrated by finally simplified system models. According to (Kohonen 1982), the basic components of these systems are:

1. A set of processor units that receive coherent inputs from the event space and create simple distinguishing functions of their input signals;
2. A mechanism that compares distinguishing functions and selects the unit with the highest functional value;
3. A kind of local interaction that simultaneously activates the selected drive and its closest neighbors;
4. An adaptive process that causes the parameters of activated units to increase their distinguishing functional values related to simultaneous input.

Ordered mapping

To clarify the methodology, let us take a look at Figure 1, which describes a simple one-stage self-organized system. Information about events A_1, A_2, A_3, \dots , which take place in the outside world, is transmitted in the form of sensory signals to a set of processor units (shown here a one-dimensional field for simplicity) via a transmission network.

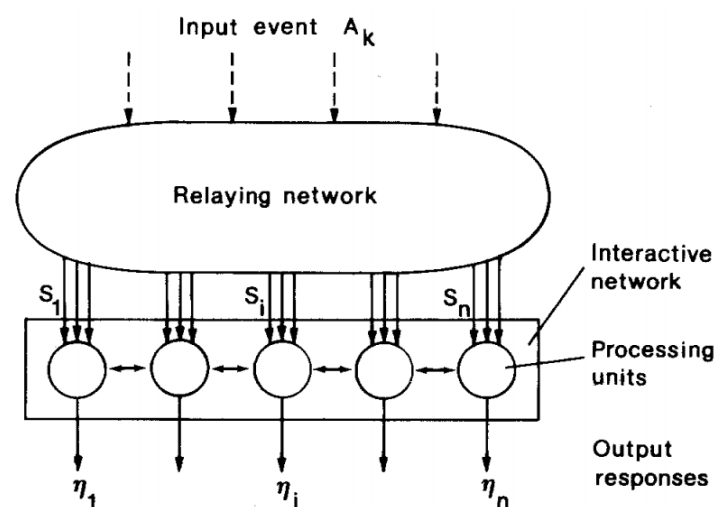


Figure 1. Illustration of a system that implements ordered mappings. Source: (Kohonen 1982).

Sets of sensory signals S_i distributed to each processor unit i may be non-identical, and the number of signals in each S_i may be different; however, these signals are considered coherent in the sense that they are clearly determined by the same events A_k . Suppose that events A_k can be arranged in some metric or topological way so that $A_1 \mathbf{R} A_2 \mathbf{R} A_3 \mathbf{R} \dots$, where \mathbf{R} stands for general arrangement relationship that is transitive (the above means e.g., $A_1 \mathbf{R} A_3$). Next, the processor units produce output responses to events with scalar values $\eta_i(A_1), \eta_i(A_2), \dots$

The system according to Figure 1 is intended to implement one-dimensional ordered mapping if for $i_1 > i_2 > i_3 > \dots$,

$$\eta_{i_1}(A_1) = \max_j \{ \eta_j(A_1) \mid j = 1, 2, \dots, \eta \}$$

$$\eta_{i_2}(A_2) = \max_j \{ \eta_j(A_2) \mid j = 1, 2, \dots, \eta \}$$

$$\eta_{i_3}(A_3) = \max_j \{ \eta_j(A_3) \mid j = 1, 2, \dots, \eta \}$$

The above definition is easily generalizable for two or more large arrays of processor units; in this case, topological order, induced by more than one ordered relationship with respect to different attributes, shall be definable for events A_k . On the other hand, the field topology is simply defined by the definition of the neighbors of each unit. If the unit with the maximum response to a specific event is treated as an image of that event, then the mapping is considered ordered if the topological relationships of images and events are similar (Kohonen 1982).

Creation of topological maps in a two-dimensional field with the same inputs to all units

Figure 2 defines a rectangular array of processor drives. In the first experiment, the transmission network was neglected, and the same set of input signals was connected to all units $\{ \xi_1, \xi_2, \dots, \xi_n \}$. In accordance with the notations used in mathematical system theory, this set of signals was expressed as a column vector $x = [\xi_1, \xi_2, \dots, \xi_n]^T \in R^n$ where T indicates transposition. Unit i shall have input balances or parameters $\mu_{i1}, \mu_{i2}, \dots, \mu_{in}$, expressible as an additional vector $m_i = [\mu_{i1}, \mu_{i2}, \dots, \mu_{in}]^T \in R^n$. The unit shall constitute a discriminatory function:

$$\eta_i = \sum_{j=1}^n \mu_{ij} \xi_j = m_i^T x \tag{4}$$

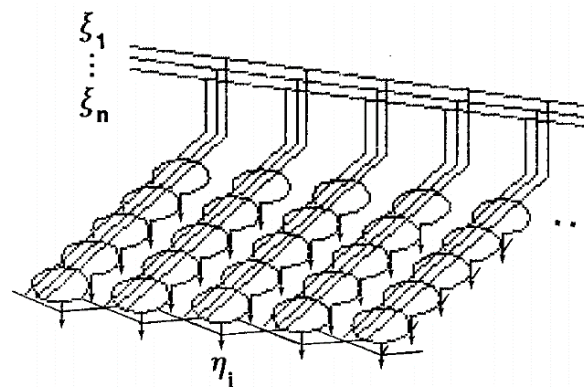


Figure 2. Two-dimensional array of processor units. Source: (Kohonen 1982).

In addition, a discriminatory mechanism works to select the maximum of η_i :

$$\eta_k = \max_i \{ \eta_i \} \tag{5}$$

For unit k and all eight of its closest neighbors (except the edges of a field where the number of neighbors was different), the following adaptive process is assumed to be active:

$$m_i(t+1) = \frac{m_i(t) + \alpha \chi(t)}{||m_i(t) + \alpha \chi(t)||_E} \tag{6}$$

where variables have been identified by the confidentiality index t (integer), α is the “adaptation parameter” in the adaptation, and the denominator is the Euclidean norm of the numerator. Equation (6) otherwise resembles the well-known perceptron learning rule, except that the direction of corrections is always the same as the direction (no decision-making process or supervision is included), and the weighing vectors are normalized. Normalization improves selectivity in discrimination and is also beneficial in maintaining “memory sources” at a certain level. Notice that the equation process (6) does not change length but only rotates. However, it is not always necessary for the Euclid χ $m_i m_i$ χ standard to be the norm, as in Equation (6).

The data were processed using the method of Kohonen networks. Figure 3 shows the self-organizing algorithm.

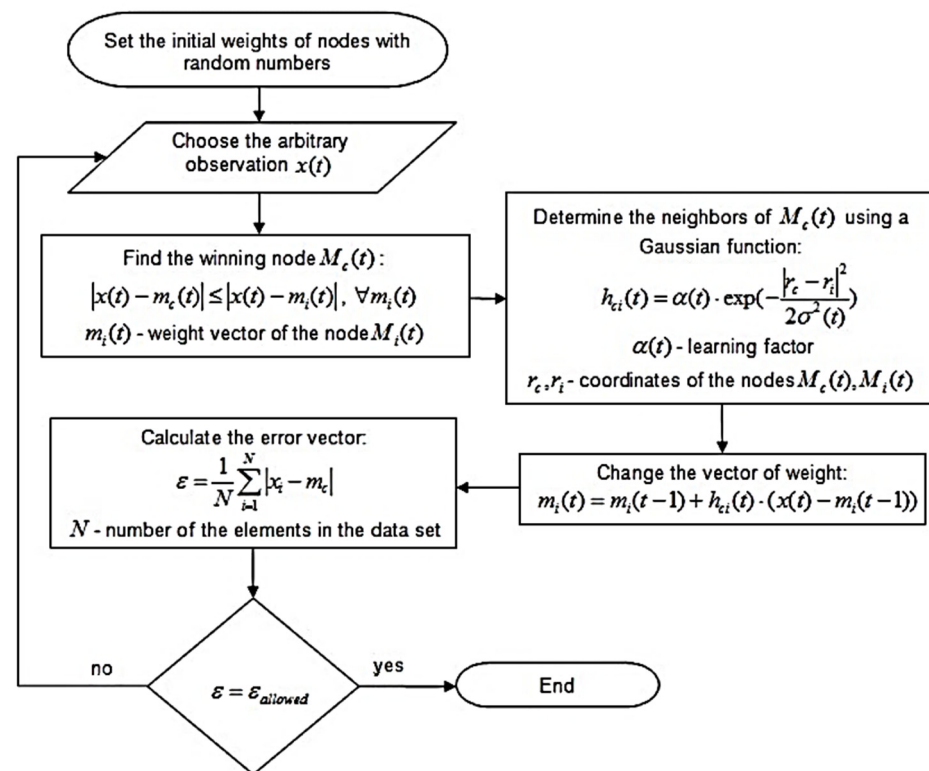


Figure 3. Kohonen self-organizing algorithm. Source: (Awrejcewicz et al. 2011).

Training—network setting and results

The data were processed using the TIBCO Statistica program. We chose “SANN: Cluster analysis”. Topological height was 10 and “Topological width” was 10. We set the training cycle to 10,000, “Learning rates” start to tenths and end to thousandths.

Procedure for evaluating the strongest cluster

i. Assignment of cluster positions to individual enterprises

As a first step, one must assign the resulting clusters to the enterprises in the original dataset in each year, as shown in Table 1. Here, the positions in the cluster are listed as “Position 2015”. The source of individual clusters is the result of neural network training.

Table 1. Original data file with assigned positions.

Company ID	Source Date	Beginning of Period	End of Period	Number of Months	Position 2015
5885	20160725	20150901	20151231	12	(1, 1)
28244532	20160725	20130901	20151231	12	(2, 2)
50993531	20160725	20130001	20151231	12	(2, 1)
80753811	20150725	20150101	20151231	12	(1, 2)
24622191	20170724	20150901	20151231	12	(1, 3)
25663135	20170115	20150901	20951231	12	(1, 1)
27092077	20150620	20130901	20151231	12	(2, 1)
25438307	20160704	20150001	20151231	12	(5, 1)
49710571	20150815	20130001	20151231	12	(2, 1)
28396678	20150604	20150901	20951231	12	(1, 1)
25506881	20190704	20130901	20151231	12	(2, 1)
25317075	20150101	20150901	20951231	12	(2, 10)
47134983	20150505	20150901	20151231	12	(1, 1)
28202375	20170227	20130901	20151231	12	(2, 9)

Source: Author.

Determination of enterprise value generators as variables

Position assignment is followed by the determination of variables that have been specified as enterprise value generator entries and then the creation of Pivot Table reports.

The variables are defined as generators of the value of the enterprise:

- total assets;
- fixed assets;
- sales;
- operating profit or loss.

ii. Determination of cluster ranking and number of their points

The individual variables mentioned above are filtered out of the Pivot Tables, along with the positions in the clusters, and these data are transferred to new tables where orders 1–10 are delivered; see Table 2. Aftersorting was determined in such a way that the first is always the cluster where the variable acquired the highest value. This was done for all years and for all variables. The aim is to find out which cluster is the strongest and then find and compare, using financial analysis, the companies in it for 2015, 2016, 2017, and 2018.

Table 2. Cluster tables and variables including order.

2015				2016				2017				2018			
Cluster	Company ID	Assets (in Thousand CZK)	Rank	Cluster	Company ID	Assets (in Thousand CZK)	Rank	Cluster	Company ID	Assets (in Thousand CZK)	Rank	Cluster	Company ID	Assets (in Thousand CZK)	Rank
(1, 1)	5	193,983,526	1	(1, 1)	4	171,374,111	1	(10, 10)	5	204,003,403	1	(1, 1)	3	154,281,306	1
(2, 1)	4	49,520,740	2	(1, 3)	4	34,370,036	2	(9, 10)	5	41,379,275	2	(1, 2)	5	56,310,851	2
(3, 1)	6	13,980,772	3	(2, 1)	2	23,429,904	3	(2, 10)	4	13,061,112	3	(3, 1)	3	43,923,622	3
(2, 10)	12	10,575,537	4	(1, 10)	13	17,644,871	4	(8, 10)	5	13,044,534	4	(5, 1)	4	32,299,614	4
(2, 2)	2	8,418,137	5	(1, 4)	5	10,231,255	5	(3, 10)	2	9,053,831	5	(1, 3)	9	12,570,680	5
(1, 2)	1	8,083,333	6	(3, 2)	6	8,736,490	6	(10, 9)	2	8,510,567	6	(1, 10)	7	9,475,552	6
(5, 1)	15	7,906,670	7	(1, 2)	2	8,415,907	7	(1, 8)	13	8,065,330	7	(4, 2)	1	6,992,077	7
(1, 3)	1	6,402,191	8	(1, 5)	11	7,725,963	8	(8, 8)	10	6,946,246	8	(1, 9)	11	5,517,543	8
(2, 6)	12	5,468,181	9	(3, 3)	14	7,274,867	9	(1, 10)	3	6,841,701	9	(6, 2)	7	4,918,045	9
(1, 4)	3	5,308,183	10	(2, 3)	11	6,764,269	10	(1, 9)	6	6,641,100	10	(2, 2)	2	4,274,520	10

Source: Author.

From the tables shown in Table 2, pivot tables were created (see Table 3), and 1–10 point scales were added to the table, with 10 points being added to the table if the cluster took first place in the order. In this way, the scoring continued up to the tenth place, where the cluster in this place received only one point for its position. The location of the cluster depends on the amount of assets, fixed assets, sales, and operating income.

Table 3. Pivot Table with assigned points.

Assets 2015		
Row Labels	Sum of Order	Number of Ranking Points
(1, 1)	1	10
(1, 3)	8	3
(1, 4)	10	1
(2, 1)	2	9
(2, 10)	4	7
(2, 2)	5	6
(3, 1)	3	8
(5, 1)	7	4
(1, 2)	6	5
(2, 6)	9	2

Source: Author.

The “Order Points Count” column is added to each of the Pivot Tables, and using the IF(B7 = 1;” function. 10”; IF(B7 = 2;” 9”; IF(B7 = 3;” 8”; IF(B7 = 4;” 7”; IF(B7 = 5;” 6”; IF(B7 = 6;” 5”; IF(B7 = 7;” 4”; IF(B7 = 8;” 3”; IF(B7 = 9;” 2”; IF(B7 = 10;”)), the number of points is assigned based on the order in the table. This is done for all variables in all years.

iii. Assembling a cluster leaderboard and determining the strongest cluster

To select the strongest class, the results from all Pivot Tables below must be copied and a single Pivot Table report must be created in which the clusters can be sorted based on all points earned for all years and for all variables. Subsequently, financial analysis was carried out, which is limited to the ratio indicators.

4. Results

4.1. Dataset

The basis for the research was a dataset including the data from 21,882 companies, whose number was reduced after modifications to 1868 companies. The data were obtained from Bisnode’s Albertina database from section H—Transportation and Storage according to the classification CZ NACE, where all legal persons from Section H for the years 2015–2018 were selected.

Each data row contains:

1. Specification of the company: company registration number (IČO), name of the company;
2. Information on the company: date of resources, beginning and end of the period, and number of months of the financial statement;
3. Financial statements: balance sheet, profit, and loss account (PLL).

The dataset was divided into four other sheets, where the information about the companies and the variables were presented for the individual years of the monitored period (2015–2018).

4.2. Training of Neural Networks and Its Results

Network training was performed in the Statistica programme using the cluster analysis function. The basis for the training was the dataset containing the data on transport companies obtained from the financial statements for the years 2015–2018. Important information for the research is the location and number of companies in individual clusters, as presented in Table 4 and 3D Graph, which can be generated directly in the Statistica

programme. The number of companies in individual clusters for the year 2015 is shown in Table 1.

Table 4. Number of companies in individual clusters for the year 2015.

	1	2	3	4	5	6	7	8	9	10
1	5	4	6	5	15	6	2	8	8	16
2	1	2	2	3	9	7	20	11	8	14
3	1	4	2	6	7	8	12	16	20	23
4	3	2	8	10	7	13	9	8	24	18
5	7	4	7	6	11	13	11	28	13	24
6	3	12	9	8	11	14	28	22	23	26
7	5	4	5	19	30	22	24	35	50	24
8	11	10	4	7	27	32	57	25	10	10
9	7	5	15	13	42	69	8	34	17	15
10	3	12	18	56	103	292	46	35	17	18

Source: Author.

Tables 4–8 need to be read according to the numbers horizontally or vertically located in the Table 4. The numbers indicate the precise location of the clusters and the number of companies in clusters. When finding out the number of companies situated, e.g., in the cluster (2, 1), start reading from the first number of the horizontal row to the number of the vertical column. You can see that in 2015, the cluster (2, 1) contained four companies.

Table 5. The number of companies in individual clusters for the year 2016.

	1	2	3	4	5	6	7	8	9	10
1	4	2	1	7	4	7	6	10	1	13
2	2	4	6	4	10	4	3	10	5	7
3	4	11	14	14	15	11	12	17	4	8
4	5	18	13	19	25	16	6	22	16	13
5	11	16	12	26	16	19	11	18	35	7
6	8	10	18	13	37	32	20	22	23	29
7	12	22	16	11	10	32	36	47	24	33
8	12	6	17	10	12	13	16	54	38	38
9	7	4	15	22	23	31	7	79	31	5
10	13	8	13	13	26	34	32	220	12	10

Source: Author.

It can be seen from Table 4 that the strongest cluster in terms of the number of companies significantly exceeds the average number of companies per cluster (1). The average number of companies per cluster in 2015 was 18.69.

It can be noted that the high number of companies in a cluster does not mean it is the strongest cluster in the overall analysis, since if a cluster contains too many companies, it indicates that the cluster consists of more companies that are weaker but show the most similarities based on which the network included them in the cluster. Experience shows that the strongest cluster usually contains fewer than the average number of companies.

Table 6. Number of companies in individual clusters for the year 2017.

	1	2	3	4	5	6	7	8	9	10
1	23	20	20	24	26	44	89	289	44	30
2	5	22	10	16	28	57	39	79	11	45
3	20	18	22	29	25	26	30	34	34	22
4	11	14	16	6	24	12	24	21	21	15
5	14	7	8	21	25	18	25	21	12	10
6	3	15	15	11	13	15	13	12	15	11
7	9	6	10	17	9	17	9	11	16	13
8	13	4	6	9	13	7	5	10	3	14
9	6	2	4	4	8	11	11	5	1	2
10	3	4	2	3	6	8	6	5	5	5

Source: Author.

Table 7. Number of companies in individual clusters for the year 2018.

	1	2	3	4	5	6	7	8	9	10
1	3	0	3	1	4	11	6	9	11	4
2	5	2	0	1	3	7	6	14	14	8
3	9	1	2	9	5	11	7	3	10	18
4	5	4	8	5	10	6	13	9	5	10
5	5	7	8	10	16	14	15	9	15	32
6	6	9	6	10	19	17	19	28	14	42
7	5	6	6	9	18	23	44	20	25	41
8	1	7	12	5	3	14	29	33	61	38
9	11	4	9	9	8	5	7	15	14	57
10	7	5	9	11	13	25	23	5	38	222

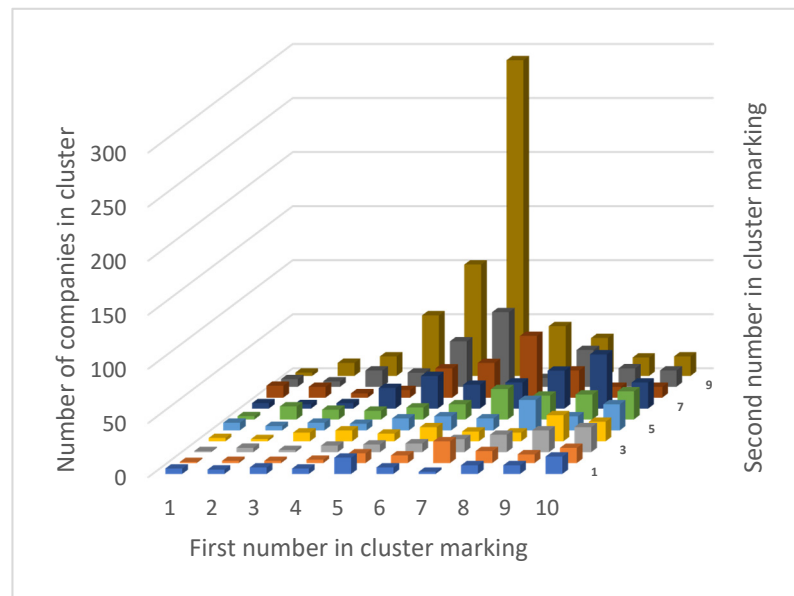
Source: Author.

Table 8. Table of ten clusters with the highest sum of assets for the year 2015.

Clusters	Number of Companies	Sum of Total Assets—Thousand CZK
(1, 1)	5	193,983,526
(2, 1)	4	49,520,740
(3, 1)	6	13,980,772
(2, 10)	12	10,575,537
(2, 2)	2	8,418,137
(1, 2)	1	8,083,333
(5, 1)	15	7,906,670
(1, 3)	1	6,402,191
(2, 6)	12	5,468,181

Source: Author.

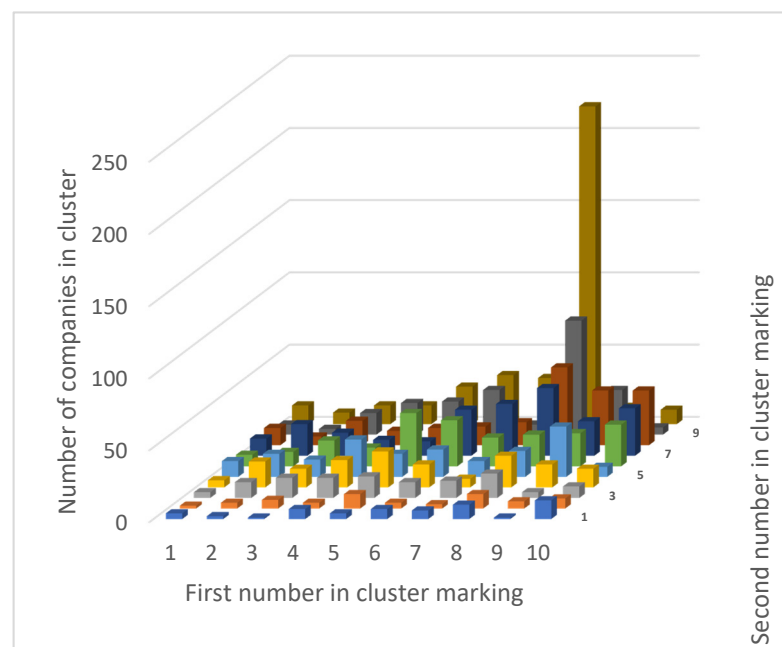
Scheme 1 shows the number of companies in clusters; i.e., it shows how many companies are in the given cluster in a specific year. The schema also shows that the biggest cluster in terms of the number of companies is in the position (6, 10), as seen in Table 4 based on which the Scheme 1 was created.



Scheme 1. Graphical representation of the number of companies in individual clusters in the year 2015. Source: Author.

Table 5 shows the number of companies in clusters for the year 2016.

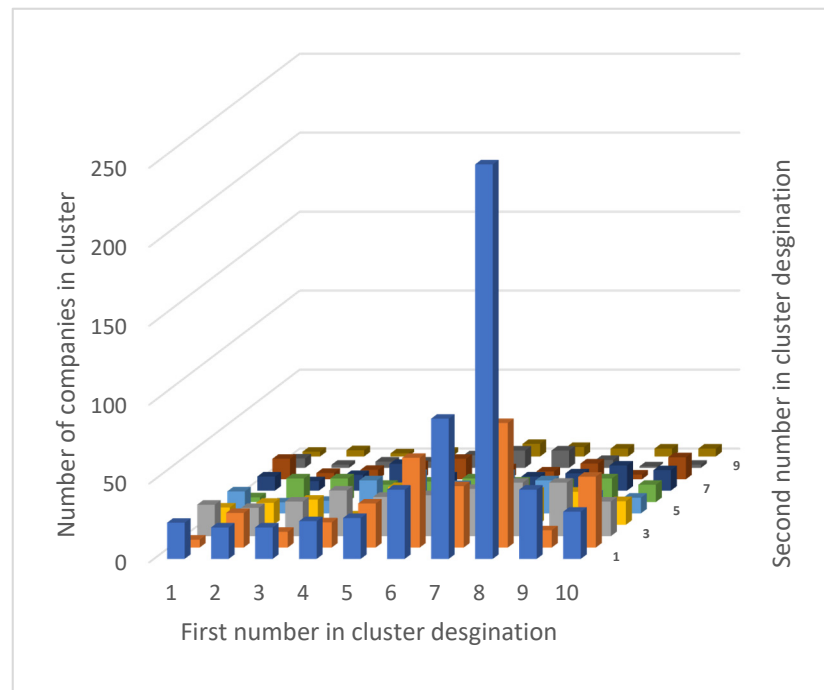
Scheme 2 is a graphical representation of the location and number of companies in clusters based on the information contained in Table 5 above. The highest number of companies (220) was in 2016 in the cluster (8, 10). The average number of companies per one cluster in 2016 was 18, which is slightly fewer than in 2015.



Scheme 2. Graphical representation of the number of companies in individual clusters for the year 2016. Source: Author.

Table 6 shows the number of companies in individual clusters in the year 2017.

Scheme 3 also indicates the changes in the number of companies in clusters and the location of clusters. In 2017, the number of companies in the first clusters increased, while the more distant clusters contain fewer companies, as seen in Scheme 3.

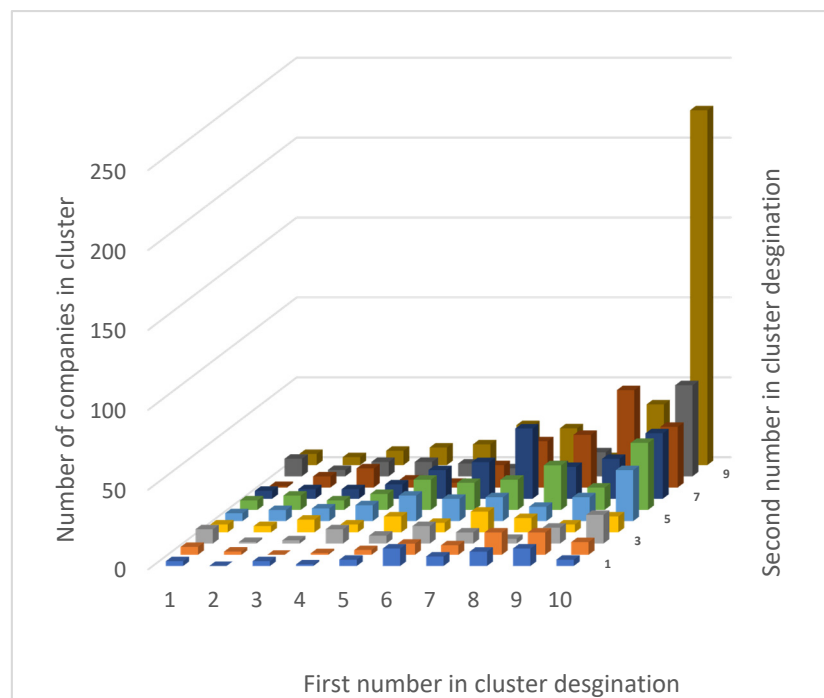


Scheme 3. Graphical representation of the number of companies in individual clusters in 2017. Source: Author.

The average number of companies per one cluster in 2017 is 19. As follows from the graph, the cluster with the highest number of companies is (8, 1) with 289 companies.

Table 7 below shows the data from the year 2018, which indicate that the strongest cluster in terms of the number of companies is the cluster (10, 10), with its 222 companies. It is also the first year when clusters with zero companies appear—clusters (2, 1) and (3, 2).

Scheme 4 is a graphical representation of the number of companies in individual clusters.



Scheme 4. Graphical representation of the number of companies in individual clusters in 2018. Source: Author.

4.3. Analysis of Strongest Cluster

The variables of business value generators are Assets, Fixed assets, Sales, and Operating result. These variables are used for the evaluation of the individual clusters. For a general comparison of clusters, other items are selected, specifically indebtedness, services, personnel costs, and interest payable.

First, the assets for the year 2015 were evaluated and ten clusters with the highest sum of company assets (in descending order) in each cluster were selected. The column “Number of companies from company ID” in Table 8 indicates the number of companies in each cluster. “Sum of Assets in total” indicates the sum of assets of individual companies in each cluster.

Table 8 indicates that the strongest cluster in terms of the total assets (193,983,526 thousand CZK) is the cluster (1, 1), which contains five companies. The second strongest cluster is (2, 1) containing four companies with a total sum of assets of 49,520,740 thousand CZK. The cluster with the third-highest sum of assets is the cluster (3, 1) containing six companies with a total sum of assets being 13,980,772 thousand CZK.

The analysis by assets is as follows.

The cluster (1, 1) is also the strongest one in the year 2016 (see Table 9). Compared to the year 2015, the cluster contains four companies only. The cluster (2, 1) dropped to the third position and contained only two companies. The second place is occupied by the cluster (1, 3), which was on the 9th position in 2015.

Table 9. Table of ten clusters with the highest sum of assets for the year 2016.

Clusters	Number of Companies	Sum of Total Assets—In Thousand CZK
(1, 1)	4	171,374,111
(1, 3)	4	34,370,036
(2, 1)	2	23,429,904
(1, 10)	13	17,644,871
(1, 4)	5	10,231,255
(3, 2)	6	8,736,490
(1, 2)	2	8,415,907
(1, 5)	11	7,725,963
(3, 3)	14	7,274,867
(2, 3)	11	6,764,269

Source: Author.

The year 2017 brought a major change in the form of the cluster (10, 10) (see Table 10), which was replaced the cluster (1, 1). However, this does not mean that the companies in the cluster (1, 1) lost a major part of their assets in this period; the cluster (1, 1) was replaced by the cluster (10, 10) in the network training.

Table 10. Table of ten clusters with the highest total sum of assets for the year 2017.

Clusters	Number of Companies	Sum of Total Assets—Thousand CZK
(10, 10)	5	204,003,403
(9, 10)	5	41,379,275
(2, 10)	4	13,061,112
(8, 10)	5	13,044,534
(3, 10)	2	9,053,821
(10, 9)	2	8,510,567
(1, 8)	13	8,065,330
(8, 8)	10	6,946,246
(1, 10)	3	6,841,701
(1, 9)	6	6,641,100

Source: Author.

The cluster (10, 10) is considered to be the cluster (1, 1) since a detailed examination of the companies showed that the companies from the cluster (1, 1) are now in the cluster (10, 10) (see Table 11).

Table 11. Change of clusters—(1, 1) replaced by (10, 10).

Company ID	Name of the Company	Source Date	Beginning of the Period	End of the Period	Number of Months of Financial Statements	Position 2015	Position 2016	Position 2017	Position 2018
70994234	SŽDC	20170502	20150101	20151231	12	(1, 1)	(1, 1)	(10, 10)	(1, 1)
5886	Dopavní podnik hl. m. Prahy	20160725	20150101	20151231	12	(1, 1)	(1, 1)	(10, 10)	(1, 1)

Source: Author.

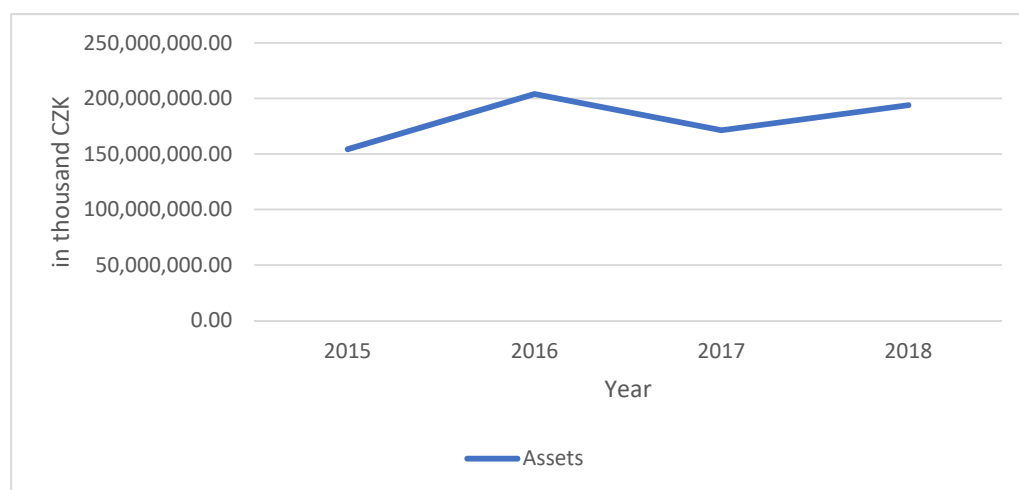
As seen from Table 12, in 2018, the cluster (1, 1) was again in the first position, with the total assets being 154,281,306 thousand CZK. In 2018, it contained three companies. The second position was occupied by the cluster (1, 2) with five companies, while the third position was occupied by the cluster (3, 1) with three companies.

Table 12. Table of ten clusters with the highest total sum of assets for the year 2018.

Clusters	Number of Companies	Sum of Total Assets—Thousand CZK
(1, 1)	3	154,281,306
(1, 2)	5	56,310,851
(3, 1)	3	43,923,622
(5, 1)	4	32,299,614
(1, 3)	9	12,570,680
(1, 10)	7	9,475,552
(4, 2)	1	6,992,077
(1, 9)	11	5,517,543
(6, 2)	7	4,918,045
(2, 2)	2	4,274,520

Source: Author.

Scheme 5 below represents the development of assets in the cluster (1, 1), which was replaced by the cluster (10, 10) in 2017.



Scheme 5. Development of assets of companies in cluster (1, 1) in the years 2015–2018 (in thousand CZK). Source: Author.

The analysis by fixed assets is as follows.

In terms of the sum of fixed assets, the strongest cluster for the year 2015 is (1, 1) (see Table 13), which occupied the highest positions in the evaluation of clusters by the highest sum of total assets. The cluster (1, 1) is represented by five companies, with the total sum of fixed assets being 160,176,338 thousand CZK.

Table 13. Table of ten clusters with the highest total sum of fixed assets for the year 2015.

Clusters	Number of Companies	Sum of Fixed Assets—In Thousand CZK
(1, 1)	5	160,176,338
(2, 1)	4	30,832,486
(3, 1)	6	11,397,642
(1, 2)	1	6,361,529
(1, 3)	1	5,993,084
(5, 1)	15	5,775,456
(2, 6)	12	4,593,348
(2, 2)	2	3,926,090
(2, 10)	12	3,692,457
(1, 4)	3	3,531,139

Source: Author.

In 2016, the strongest cluster was again (1, 1) represented by four companies only, with the total sum of fixed assets being 151,923,525 thousand CZK (see Table 14 below).

Table 14. Table of ten clusters with the highest total sum of fixed assets for the year 2016.

Clusters	Number of Companies	Sum of Fixed Assets—In Thousand CZK
(1, 1)	4	151,923,525
(1, 3)	4	27,258,661
(1, 4)	5	7,207,819
(2, 1)	2	6,711,638
(3, 2)	6	6,678,759
(3, 1)	1	5,968,467
(3, 3)	14	5,125,938
(1, 5)	11	4,853,276
(1, 2)	2	4,022,048
(2, 3)	11	3,509,271

Source: Author.

Table 15 shows that the cluster (1, 1) was replaced again by the cluster (10, 10). This cluster is thus again considered the cluster (1, 1); in 2017, it was represented by five companies, with the total sum of fixed assets being 163,426,753 thousand CZK.

Table 15. Table of ten clusters with the highest total sum of fixed assets for the year 2017.

Clusters	Number of Companies	Sum of Fixed Assets—In Thousand CZK
(10, 10)	5	163,426,753
(9, 10)	5	29,441,318
(8, 10)	5	8,776,917
(2, 10)	4	8,669,415
(3, 10)	2	8,580,520
(8, 9)	5	5,531,644
(8, 8)	10	5,310,814
(10, 9)	2	2,887,681
(1, 8)	13	2,241,491
(9, 7)	16	2,125,259

Source: Author.

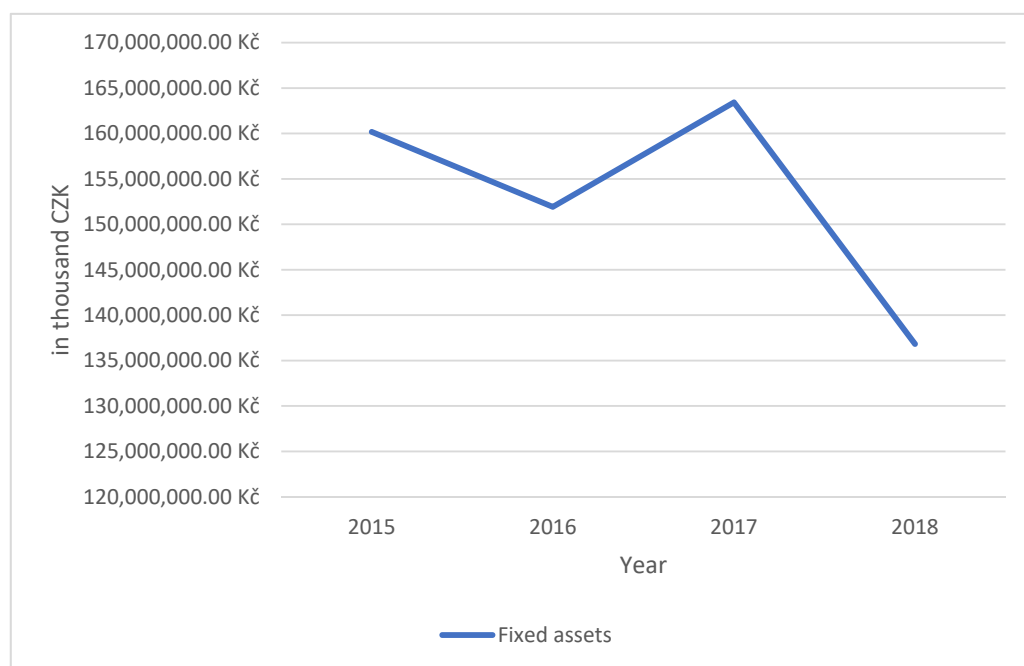
In 2018, the strongest cluster in terms of the total sum of fixed assets was (1, 1), which so far appears to be the strongest cluster overall. As seen in Table 16, the second strongest cluster is (1, 2), with a total of five companies; the third strongest cluster (1, 2) contains four companies.

Table 16. Table of ten clusters with the highest total sum of fixed assets for the year 2018.

Clusters	Number of Companies	Sum of Fixed Assets—In Thousand CZK
(1, 1)	3	136,822,241
(1, 2)	5	50,126,671
(5, 1)	4	26,546,120
(3, 1)	3	16,249,671
(1, 3)	9	7,026,354
(4, 2)	1	4,845,555
(6, 2)	7	3,338,331
(5, 3)	5	2,702,552
(1, 5)	5	2,474,129
(1, 10)	7	2,119,050

Source: Author.

The development of the sum of fixed assets in the strongest cluster (1, 1) (cluster (10, 10) for the year 2017) is shown in Scheme 6. In 2018, the curve goes down, which indicates the drop in the total sum of fixed assets of all companies in the cluster (1, 1).



Scheme 6. Fixed assets of companies in cluster (1, 1) in the years 2015–2018 (in thousand CZK). Source: Author.

The analysis by sales is as follows.

As seen from Table 17, the first position in the year 2015 is occupied again by the cluster (1, 1) with the total sum of sales of 69,370,150 thousand CZK. In 2015, the cluster includes five companies. The second position is occupied by the cluster (2, 10) with twelve companies, and the third position, the cluster (1, 10), containing three companies.

Table 17. Table of ten clusters with the highest total sum of fixed assets for the year 2015.

Clusters	Number of Companies	Sum of Sales of Products and Services—In Thousand CZK
(1, 1)	5	69,370,150
(2, 10)	12	26,868,472
(1, 10)	3	16,289,368
(2, 1)	4	11,123,704
(2, 2)	2	9,894,425
(1, 9)	7	9,840,673
(1, 8)	11	8,364,450
(3, 4)	8	7,140,208
(2, 9)	5	6,380,794
(1, 2)	1	5,927,997

Source: Author.

As in the previous cases, even in the year 2016, the strongest cluster turned out to be cluster (1, 1), which contained four companies with the total sum of sales being 51,118,919 thousand CZK. The second strongest cluster, which contained thirteen companies, was the cluster (1, 10) (see Table 18).

Table 18. Table of ten clusters with the highest total sum of fixed assets for the year 2016.

Clusters	Number of Companies	Sum of Sales of Products and Services—In Thousand CZK
(1, 1)	4	51,118,919
(1, 10)	13	40,328,857
(2, 1)	2	12,232,248
(1, 3)	4	10,824,087
(1, 2)	2	10,407,169
(2, 3)	11	10,080,753
(1, 9)	7	9,220,336
(2, 2)	4	7,888,608
(2, 4)	18	7,720,079
(1, 8)	12	6,668,351

Source: Author.

In the year 2017, the cluster (1, 1) was typically replaced by the cluster (10, 10). As seen from Table 19, the cluster contains five companies with the sum of sales being 59,077,349 thousand CZK. The second position is occupied by the cluster (10, 9); the third strongest cluster was (1, 8), for the first time being one of the three strongest clusters.

Table 19. Table of ten clusters with the highest total sum of fixed assets for the year 2017.

Clusters	Number of Companies	Sum of Sales of Products and Services—Thousand CZK
(10, 10)	5	59,077,349
(10, 9)	2	25,728,590
(1, 8)	13	20,212,601
(9, 10)	5	20,043,831
(1, 10)	3	19,716,830
(1, 9)	6	15,135,957
(2, 10)	4	14,996,516
(2, 6)	15	7,661,997
(6, 10)	8	7,547,896
(1, 7)	9	7,451,702

Source: Author.

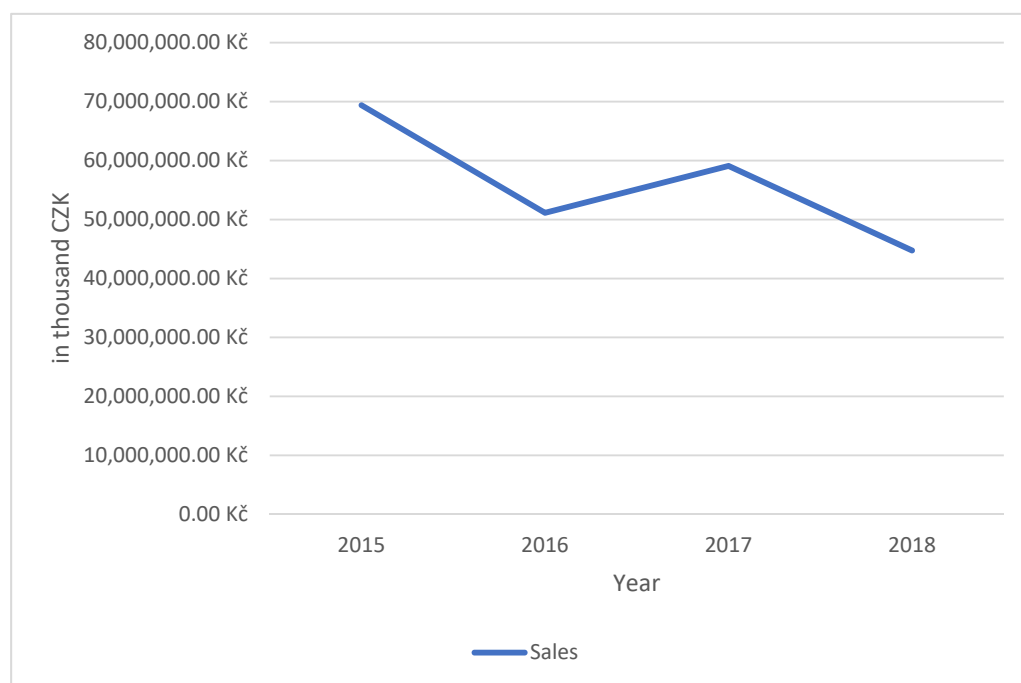
The year 2018 was no exception; the first position was occupied by the cluster (1, 1) (see Table 20). The cluster (1, 1) consists of three companies, with the total sum of sales being 44,758,784 thousand CZK.

Table 20. Table of ten clusters with the highest total sum of fixed assets for the year 2018.

Clusters	Number of Companies	Sum of Sales of Products and Services—In Thousand CZK
(1, 1)	3	44,758,784
(3, 1)	3	31,104,767
(1, 10)	7	21,032,127
(1, 2)	5	19,486,667
(1, 9)	11	16,032,084
(1, 3)	9	15,311,278
(2, 2)	2	13,777,707
(3, 4)	8	6,205,826
(3, 10)	9	4,795,071
(6, 1)	11	4,498,169

Source: Author.

The overall development of sales in the years 2015–2018 is shown in Scheme 7. The course and the shape of the curve are very similar to the curve representing the development of company assets in the years 2015–2018.



Scheme 7. Development of sales of companies in cluster (1, 1) in the years 2015–2018 (in thousand CZK). Source: Author.

The analysis by operating results is as follows.

In the case of the variable “operating profit” (hereinafter also referred to as “OR”), a significant change was recorded, which concerned the strongest clusters. The strongest cluster appears to be cluster (2, 1) represented by four companies, whose sum of OR is 2,486,859 thousand CZK. The cluster (1, 1), represented by five companies in the previous cases, occupies the second position in 2017; the third strongest cluster appears to be the cluster (2, 10), with its twelve companies. Table 21 shows ten clusters with the highest sum of OR for the year 2015.

Table 21. Table of ten clusters with the highest sum of OR for the year 2015.

Clusters	Number of Companies	Sum of Sales of Products and Services—In Thousand CZK
(1, 1)	3	44,758,784
(3, 1)	3	31,104,767
(1, 10)	7	21,032,127
(1, 2)	5	19,486,667
(1, 9)	11	16,032,084
(1, 3)	9	15,311,278
(2, 2)	2	13,777,707
(3, 4)	8	6,205,826
(3, 10)	9	4,795,071
(6, 1)	11	4,498,169

Source: Author.

As for the results for the year 2016, Table 22 indicates that the cluster (1, 2) still occupies the first position. The cluster (1, 1) is not among the three best ones; the second position is occupied by the cluster (1, 10), while the third strongest one is the cluster (2, 1). The cluster (1, 10) consists of thirteen companies, while the cluster (2, 1) includes two companies.

Table 22. Table of ten clusters with the highest sum of OR for the year 2016.

Clusters	Number of Companies	Sum of Operating Results—In Thousand CZK
(1, 2)	2	1,813,347
(1, 10)	13	1,368,837
(2, 1)	2	1,368,284
(1, 3)	4	1,233,860
(1, 5)	11	825,693
(1, 9)	7	506,888
(2, 2)	4	492,859
(3, 1)	1	440,739
(2, 3)	11	431,646
(2, 4)	18	312,215

Source: Author.

Table 23 shows that the first position is occupied by the cluster (9, 10) with a total sum of OR being 3,556,518 thousand CZK. The second position in 2017 is occupied by the cluster (10, 10) containing five companies while the third strongest cluster is the cluster (2, 10) with a total of four companies.

Table 23. Table of ten clusters with the highest sum of OR for the year 2017.

Clusters	Number of Companies	Sum of Operating Results—Thousand CZK
(9, 10)	5	3,556,518
(10, 10)	5	2,753,954
(2, 10)	4	1,508,917
(1, 9)	6	850,809
(1, 10)	3	841,200
(1, 8)	13	621,475
(10, 9)	2	615,782
(8, 9)	5	467,929
(2, 6)	15	386,044
(3, 10)	2	292,488

Source: Author.

In 2018, cluster (1, 1) is again among the strongest clusters, specifically, it occupies the second position with its three companies. As in the year 2016, the strongest cluster is the cluster (1, 2), with the total sum of OR being 5,138,259 thousand CZK (see Table 24).

Table 24. Table of ten clusters with the highest sum of OR for the year 2018.

Clusters	Number of Companies	Sum of Operating Result—In Thousand CZK
(1, 2)	5	5,138,259
(1, 1)	3	2,000,371
(1, 10)	7	1,690,826
(3, 1)	3	1,332,033
(1, 9)	11	599,075
(5, 1)	4	505,988
(4, 2)	1	406,254
(2, 9)	4	387,415
(5, 3)	5	318,877
(2, 2)	2	299,555

Source: Author.

The development of the sum of the highest OR for the years 2015–2018 is shown in Scheme 8 below. In this case, it does not show the companies in the same cluster; for each year, the strongest cluster is different.

**Scheme 8.** Development of the highest sum of OR in the years 2015–2018 (in thousand CZK). Source: Author.

4.4. Order of Clusters and Their Scoring

To identify the strongest clusters, it was necessary to determine their order based on the sum of the given variable (e.g., the Table 25 shows the sum of assets).

Subsequently, the clusters were assigned points in such a way that clusters occupying the first position were assigned 10 points, clusters occupying the second position 9 points; for the third position, clusters are assigned 8 points; 7 points were assigned for the fourth position; 6 points for the fifth position; 5 points for the sixth position; 4 points for the seventh position; 3 points for the eighth position; 2 points for the ninth position; and 1 point for the cluster occupying the tenth position.

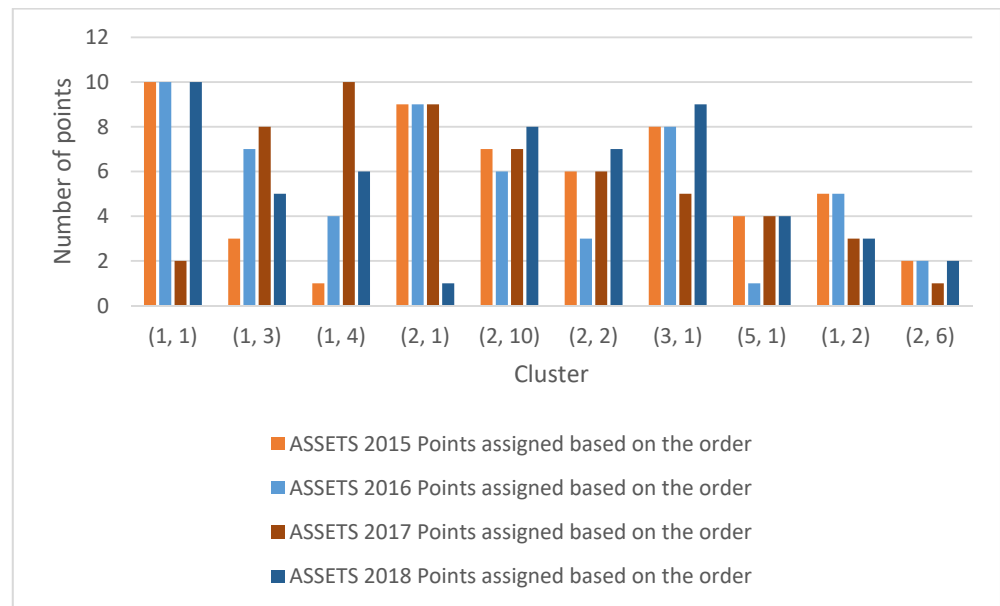
As seen from Table 25 below, the highest number of points was achieved by the cluster (1, 1). Even in this case, in 2017, the cluster (10, 10) represents the cluster (1, 1); however, this was not considered when assigning points.

Table 25. Evaluation of clusters based on their order—Assets in 2015–2018.

Assets 2015			Assets 2016			Assets 2017			Assets 2018		
Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order
(1, 1)	1	10	(1, 1)	1	10	(1, 10)	9	2	(1, 1)	1	10
(1, 3)	8	3	(1, 10)	4	7	(2, 10)	3	8	(1, 10)	6	5
(1, 4)	10	1	(1, 2)	7	4	(10, 10)	1	10	(1, 3)	5	6
(2, 1)	2	9	(1, 3)	2	9	(9, 10)	2	9	(2, 2)	10	1
(2, 10)	4	7	(1, 4)	5	6	(8, 10)	4	7	(3, 1)	3	8
(2, 2)	5	6	(1, 5)	8	3	(3, 10)	5	6	(5, 1)	4	7
(3, 1)	3	8	(2, 1)	3	8	(10, 9)	6	5	(1, 2)	2	9
(5, 1)	7	4	(2, 3)	10	1	(1, 8)	7	4	(4, 2)	7	4
(1, 2)	6	5	(3, 2)	6	5	(8, 8)	8	3	(1, 9)	8	3
(2, 6)	9	2	(3, 3)	9	2	(1, 9)	10	1	(6, 2)	9	2

Source: Author.

Scheme 9 below shows the number of assigned points for the individual clusters in the years 2015–2018. The replacement of the cluster (1, 1) for the cluster (10, 10) is not considered in the graphs or tables. The replacement would be considered if the cluster (1, 1) was not evaluated as best, as predicted before.



Scheme 9. Number of assigned points for clusters—years 2015–2018 (Assets). Source: Author.

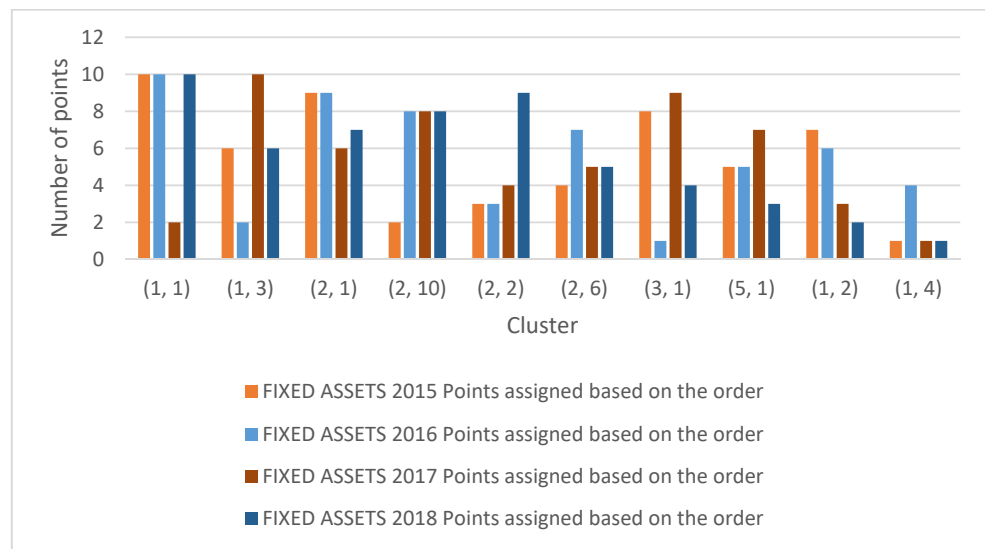
Table 26 shows that in summing up fixed assets of companies in individual clusters, the cluster (1, 1) achieved the highest number of points. In the year 2017, the cluster (10, 10) replaced the cluster (1, 1) again; however, this was not considered in assigning the points.

The total number of points assigned to individual clusters is shown in Scheme 10, where the cluster (1, 1) achieved the highest number of points. In this case, the weakest cluster out of the ten evaluated clusters is the cluster (1, 4).

Table 26. Evaluation of clusters based on their order—years 2015–2018 (Fixed Assets).

Fixed Assets 2015			Fixed Assets 2016			Fixed Assets 2017			Fixed Assets 2018		
Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order
(1, 1)	1	10	(1, 11)	1	10	(1, 8)	9	2	(1, 1)	1	10
(1, 3)	5	6	(1, 2)	9	2	(10, 10)	1	10	(1, 3)	5	6
(2, 1)	2	9	(1, 3)	2	9	(3, 10)	5	6	(3, 1)	4	7
(2, 10)	9	2	(1, 4)	3	8	(8, 10)	3	8	(5, 1)	3	8
(2, 2)	8	3	(1, 5)	8	3	(8, 8)	7	4	(1, 2)	2	9
(2, 6)	7	4	(2, 1)	4	7	(8, 9)	6	5	(4, 2)	6	5
(3, 1)	3	8	(2, 3)	10	1	(9, 10)	2	9	(6, 2)	7	4
(5, 1)	6	5	(3, 1)	6	5	(2, 10)	4	7	(5, 13)	8	3
(1, 2)	4	7	(3, 2)	5	6	(10, 9)	8	3	(1, 5)	9	2
(1, 4)	10	1	(3, 3)	7	4	(9, 7)	10	1	(1, 10)	10	1

Source: Author.



Scheme 10. Number of points assigned to clusters—years 2015–2017 (fixed assets). Source: Author.

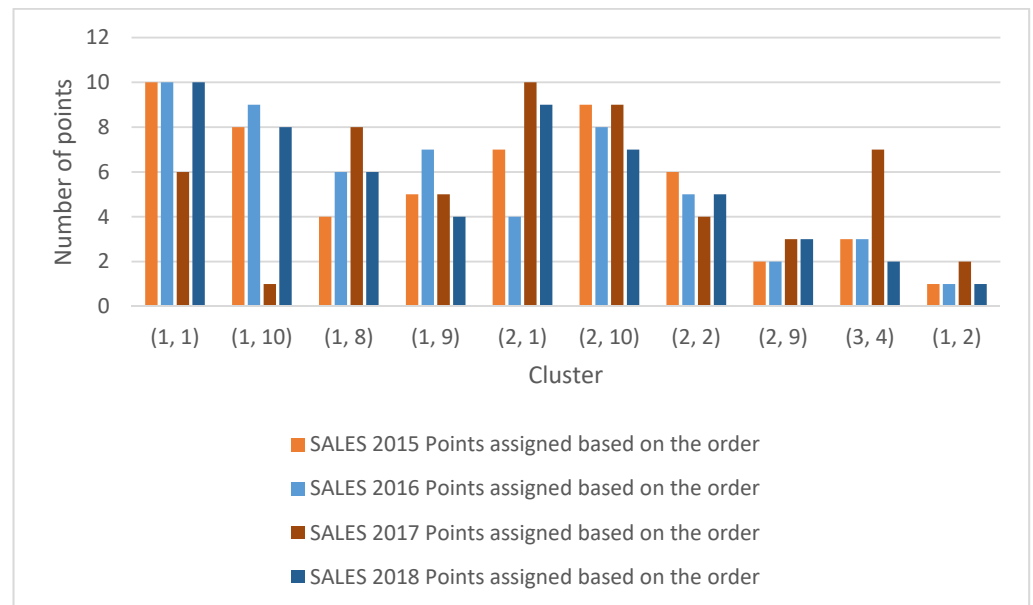
In the case of the sum of individual sales of companies in individual clusters, the highest number of points and the best position was achieved by the cluster (1, 1); in 2017, it was replaced by the cluster (10, 10), but it is still the strongest cluster in the case of analysis Sales variable. For more details, see Table 27.

Table 27. Evaluation of clusters based on their order—years 2015–2018 (Sales).

Sales 2015			Sales 2016			Sales 2017			Sales 2018		
Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order
(1, 1)	1	10	(1, 1)	1	10	(1, 10)	5	6	(1, 1)	1	10
(1, 10)	3	8	(1, 10)	2	9	(1, 7)	10	1	(1, 10)	3	3
(1, 5)	7	4	(1, 2)	5	6	(1, 5)	3	8	(1, 9)	5	5
(1, 9)	6	5	(1, 3)	4	7	(1, 9)	6	5	(2, 2)	7	7
(2, 1)	4	7	(1, 9)	7	4	(10, 10)	1	10	(3, 1)	2	2
(2, 10)	2	9	(2, 1)	3	8	(10, 9)	2	9	(1, 2)	4	4
(2, 2)	5	6	(2, 3)	6	5	(2, 10)	7	4	(1, 3)	6	6
(2, 9)	9	2	(2, 4)	9	2	(2, 6)	8	3	(3, 4)	8	8
(3, 4)	8	3	(2, 2)	8	3	(9, 10)	4	7	(3, 10)	9	9
(1, 2)	10	1	(1, 8)	10	1	(6, 10)	9	2	(6, 1)	10	10

Source: Author.

Graphical representation of points assigned to individual clusters in the case of the “Sales” variable is seen in Scheme 11.



Scheme 11. Number of points assigned to clusters—years 2015–2018 (Sales). Source: Author.

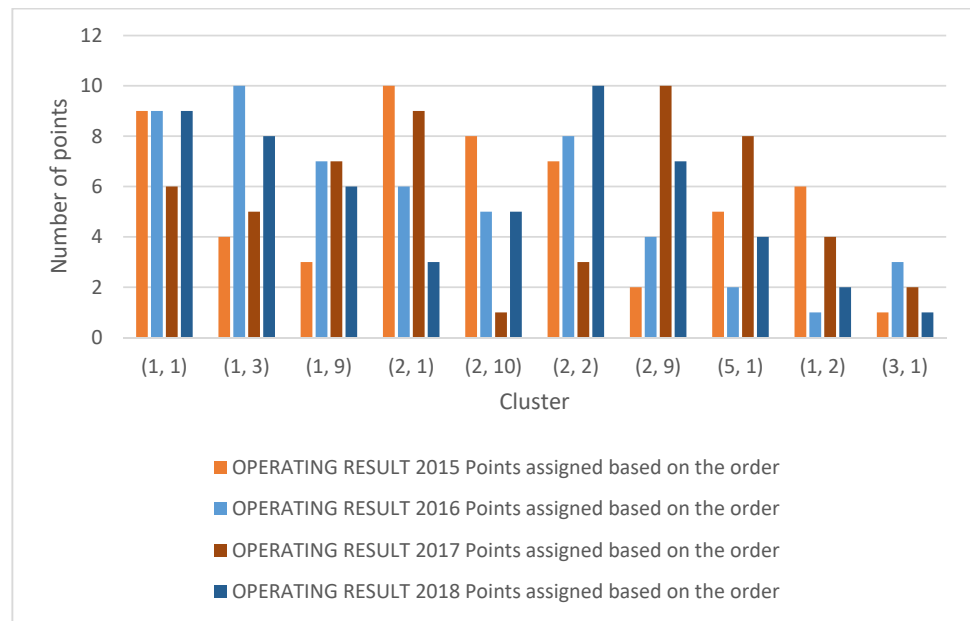
The variable “Operating result” (hereinafter also referred to as “OR”) shows notable changes in all monitored years in terms of the strongest clusters (see Table 28). The sum of OR of the companies in the cluster (2, 1) was the highest of all ten clusters and was thus assigned the highest number of points in the year 2015. In 2016, the cluster (1, 2) was assigned 10 points out of 10 due to its position as the strongest cluster. In 2017, there was another change in the first position, where the cluster (9, 10) was assigned the highest possible number of points (10). In 2018, the highest number of points was assigned to the cluster (1, 1).

Table 28. Evaluation of clusters based on their order—years 2015–2018 (Operating result).

Operating Results 2015			Operating Results 2016			Operating Results 2017			Operating Results 2018		
Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order	Clusters	Sum of Points Based on the Order	Points Assigned Based on the Order
(1, 1)	2	9	(1, 10)	2	9	(1, 10)	5	6	(1, 1)	2	9
(1, 3)	7	4	(1, 2)	1	10	(1, 5)	6	5	(1, 10)	3	8
(1, 9)	8	3	(1, 3)	4	7	(1, 9)	4	7	(1, 9)	5	6
(2, 1)	1	10	(1, 5)	5	6	(10, 10)	2	9	(2, 9)	8	3
(2, 10)	3	8	(1, 9)	6	5	(3, 10)	10	1	(5, 1)	6	5
(2, 2)	4	7	(2, 1)	3	8	(8, 9)	8	3	(1, 2)	1	10
(2, 9)	9	2	(2, 2)	7	4	(9, 10)	1	10	(3, 1)	4	7
(5, 1)	6	5	(2, 3)	9	2	(2, 10)	3	8	(4, 2)	7	4
(1, 2)	5	6	(2, 4)	10	1	(10, 9)	7	4	(5, 3)	9	2
(3, 1)	10	1	(3, 1)	8	3	(2, 6)	9	2	(2, 2)	10	1

Source: Author.

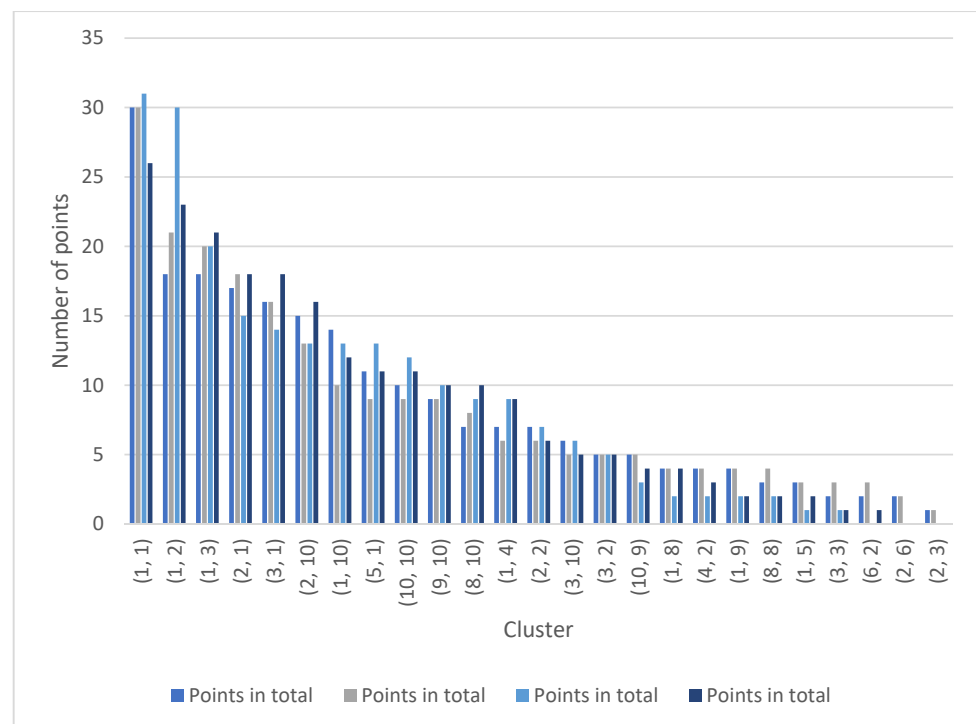
Graphical representation of the number of points assigned in the individual years for individual clusters is shown in Scheme 12.



Scheme 12. Number of assigned points for individual clusters—years 2015–2018 (Operating result). Source: Author.

4.5. The Strongest Cluster of the Overall Ranking

Before compiling the overall ranking of all clusters that appeared in the top ten list by all variables in all years, it is necessary to graphically represent the number of points the individual clusters were assigned in the sum of all years for each variable separately (see Scheme 13). It follows from the graph that the cluster (1, 1) is the best candidate for being the strongest cluster in the overall ranking of all clusters for all years and all variables. The cluster (1, 2), which is often among the three strongest clusters in the previous analyses, should also be mentioned. The data are based on the table presented in Table 29.



Scheme 13. Number of points assigned to clusters by individual variables. Source: Author.

Table 29. Table with the data on the number of points for individual clusters.

Assets		Fixed Assets		Sales		Operating Results	
Cluster	Points in Total	Cluster	Points in Total	Cluster	Points in Total	Cluster	Points in Total
(1, 1)	30	(1, 1)	30	(1, 10)	31	(1, 2)	26
(1, 2)	18	(1, 3)	21	(1, 1)	30	(1, 10)	23
(1, 3)	18	(3, 1)	20	(1, 9)	20	(1, 9)	21
(2, 1)	17	(1, 2)	18	(2, 1)	15	(2, 1)	18
(3, 1)	16	(2, 1)	16	(1, 2)	14	(1, 1)	18
(2, 10)	15	(5, 1)	13	(2, 2)	13	(2, 10)	16
(1, 10)	14	(10, 10)	10	(1, 5)	13	(2, 2)	12
(5, 1)	11	(1, 4)	9	(2, 10)	13	(3, 1)	11
(10, 10)	10	(2, 10)	9	(1, 3)	12	(1, 3)	11
(9, 10)	9	(9, 10)	9	(10, 10)	10	(5, 1)	10
(8, 10)	7	(8, 10)	8	(10, 9)	9	(9, 10)	10
(1, 4)	7	(3, 2)	6	(3, 1)	9	(10, 10)	9
(2, 2)	7	(3, 10)	6	(9, 10)	7	(1, 5)	6
(3, 10)	6	(1, 5)	5	(3, 4)	6	(1, 8)	5
(3, 2)	5	(4, 2)	5	(2, 3)	5	(2, 9)	5
(10, 9)	5	(8, 9)	5	(2, 6)	3	(4, 2)	4
(1, 5)	4	(6, 2)	4	(3, 10)	2	(10, 9)	4
(4, 2)	4	(3, 3)	4	(2, 9)	2	(8, 9)	3
(1, 9)	4	(2, 6)	4	(2, 4)	2	(5, 3)	2
(8, 8)	3	(8, 8)	4	(6, 10)	2	(2, 6)	2
(1, 5)	3	(2, 2)	3	(6, 1)	1	(2, 3)	2
(3, 3)	2	(5, 3)	3	(1, 7)	1	(2, 4)	1
(2, 6)	2	(10, 9)	3			(3, 10)	1
(2, 6)	2	(1, 8)	2				
(2, 3)	1	(9, 7)	1				
		(1, 10)	1				
		(2, 3)	1				

Source: Author.

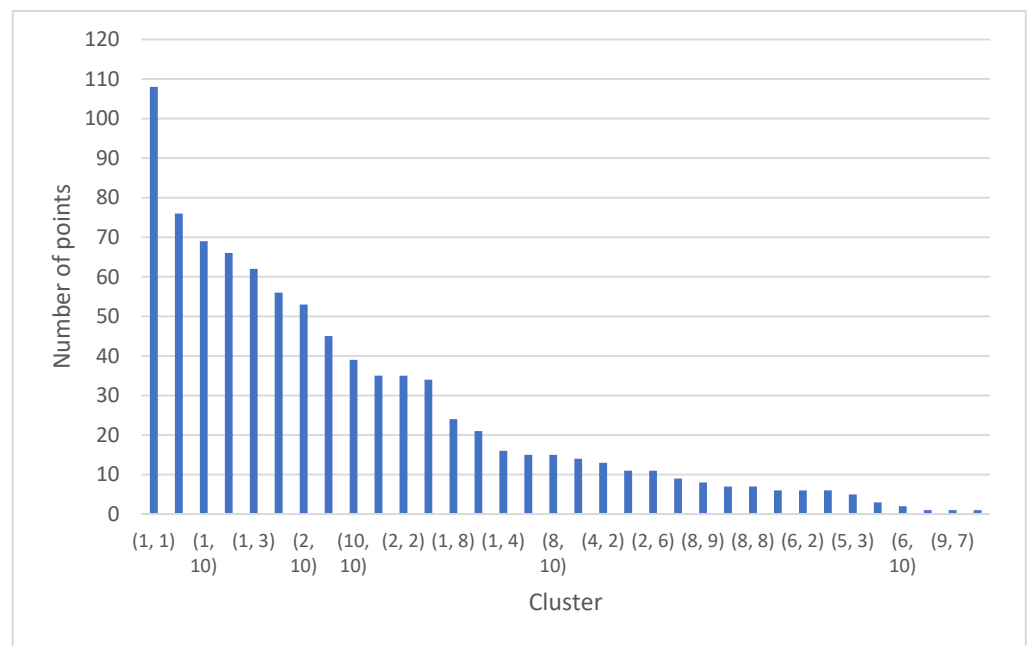
4.6. Ranking of Strongest Clusters

After compiling the ranking, the overall table was, due to its extent, put in the appendices part of this text under the title Appendix B: Overall ranking of clusters. Table 30 presents the order of the first ten strongest clusters. As predicted in the previous chapter, the first position is occupied by the cluster (1, 1) with a total number of 108 points. Scheme 14 is a graphical representation of clusters by the number of points in total.

Table 30. Overall ranking of clusters (first ten clusters).

Clusters	Sum of Points
(1, 1)	108
(1, 2)	76
(1, 10)	69
(2, 1)	66
(1, 3)	62
(3, 1)	56
(2, 10)	53
(1, 9)	45
(10, 10)	39
(9, 10)	35

Source: Author.



Scheme 14. Number of points assigned to clusters in total. Source: Author.

4.6.1. Analysis of Companies in Cluster (1, 1)

The strongest cluster based on the variables considered the business value generators turned out to be the cluster (1, 1). Table 31 shows companies from the cluster (1, 1) in individual years. It follows from the table that in 2015, the cluster (1, 1) contained five companies; in 2016, it was four companies; five companies in the year 2017, and three companies in the year 2018. Here, it shall be noted that in grouping the companies in the clusters, the situation in the year 2017 is considered, when, as already mentioned, the cluster (10, 10) replaced the cluster (1, 1).

Table 31. Representation of companies in cluster (1, 1) in the years 2015–2018.

2015	2016	2017	2018
70,994,234	70,994,234	70,994,234	70,994,234
25,663,135	25,663,135	5886	25,663,135
5886	5886	47,114,983	5886
47,114,983	28,196,678	60,193,531	
28,196,678		28,196,678	

Source: Author.

Table 31 also indicates that only two companies appeared in all years of the monitored period, specifically the company with the registration number 70,994,234 and the company with the registration number 5886. Those two companies are subject to subsequent analysis.

4.6.2. Analysis of Selected Companies from Cluster (1, 1) and Their Participation in the Creation of the Selected Components

An interesting fact is how the selected companies from the cluster (1, 1) participated in the creation of the selected components—assets, fixed assets, sales, and operating result—in individual years of the monitored period. This is presented in the table in Figure 4, where the column “Assets” shows the sum of assets of a given company for a given year, while the column “Total assets” shows the sum of assets of all companies created in the same year. The same system is used for all components in the table. The data were taken from the dataset, which was presented to the Kohonen network for training; based on this dataset, a map of clusters was created as an output of the network.

Company reg. number	Assets	Assets in total	Share	Fixed assets	Fixed assets in total	Share	Sales	Sales in total	Share	Operating profit	Operating profit in total	Share
2015												
70994234	72341103	413449913	17.50%	66897651	280639204	23.84%	6649848	324506112	2.05%	-951874	21019952	-4.53%
5886	80691116	413449913	19.52%	71264356	280639204	25.39%	17654721	324506112	5.44%	1971978	21019952	9.38%
2016												
70994234	72746546	395272967	18.40%	66982500	268999373	24.90%	6653888	304570767.8	2.18%	-2465847	15728229	-15.68%
5886	79078485	395272967	20.01%	72006551	268999373	26.77%	17580169	304570767.8	5.77%	1569062	15728229	9.98%
2017												
70994234	67492454	423054095	15.95%	62673153	281846380	22.24%	6899311	338280118.8	2.04%	-1295243	18716954	-6.92%
5886	81223107	423054095	19.20%	72983728	281846380	25.89%	18594271	338280118.8	5.50%	2432683	18716954	13.00%
2018												
70994234	65647392	414892884	15.82%	60587503	287500737	21.07%	6480726	298666892.4	2.17%	-199987	17869706	-1.12%
5886	82669816	414892884	19.93%	74087947	287500737	25.77%	19368813	298666892.4	6.49%	2190990	17869706	12.26%

Figure 4. Participation of selected companies on the creation of individual components. Source: Author.

The table shows that the company with the registration number 70994234 participated in the creation of the total assets as follows: 17.5% in 2015, 18.4 in 2016, and 15.95% in 2017, and in 2018, it decreased to its lowest level of 14.82%. The company with the registration number 5886 showed a slightly higher share on the creation of the total assets: in 2015, the company participated in the creation of the total assets of all companies in the dataset by 19.52%. It was 20.01% in 2016, 19.20% in the year 2017, and 19.93% in 2018. In the case of fixed assets, the participation of both companies was higher: in 2015, the share of the company with the registration number 70994234 on the total fixed assets was 23.84%, while the share of the company with the registration number 5886 was 25.93%. In 2016, the share of the company with the registration number 70994234 increased to 24.90% and to 26.77% in the case of the company with the registration number 5886. In the year 2017, there was a decrease to 22.24% and 25.89%. The lowest share was recorded in the year 2018, specifically 21.07% and 25.77%. As for sales, in the years 2015–2018, the share of the company with the registration number 70994234 ranged between 2.04% and 2.18%, while in the case of the company with the registration number, the share was slightly higher, ranging between 5.44% and 6.49%.

The problem is the operating result, where the company with the registration number 70994234 shows a negative operating result. This is because this company is controlled by the state, where financing is different than in the case of private companies. Companies managed by the state provide strategic services and functions for the state and their operation is financed from the state budget, and they are not intended for generating profit but ensure various functions for all citizens; their functioning is thus logically different, and their principle is completely different from private companies. They are state organizations whose task is to ensure the transport infrastructure, not to generate profit. It may seem that this company should not be included in the other companies in the cluster, but it should be considered that it exists and creates the environment this contribution deals with. The negative operating result will thus not be commented on in terms of the percentage share. However, the company with the registration number 5886 showed positive operating results for all years of the monitored period, and its share in the years 2015–2018 ranged between 9.38% and 13%.

4.7. Financial Analysis of Selected Companies

4.7.1. Company with the Registration Number 70994234—Správa Železnic, Státní Podnik

Table 32 presents a financial analysis of the company in the form of ratios. The first three components in the table need not be commented on since their negative values are explained in the previous chapter (see Negative operating result). As for cash position ratio, it indicates to what extent the company can pay its short-term liabilities immediately, i.e.,

with cash, money in bank accounts, cheques, or short-term securities. The recommended value of cash position ratio ranges between 0.2 and 0.5 and is calculated by dividing financial assets by short-term liabilities + current bank loans. The company with the registration number 70994234, as seen in Table 32, shows the value of cash position ratio of 0.11 in 2015, which is not even the lower limit of the recommended value. Lower values of cash position ratio indicate a reduced ability to pay short-term liabilities; too high values indicate inefficient management. In the case of this company, the cash position ratio increased to 0.45, 0.54, and 0.56 in the years 2016–2018, which indicates inefficient management. Quick ratio, sometimes also called second-degree liquidity, provides information on how many times the company can pay its short-time liabilities if financial assets and short-term liabilities are converted into cash. The recommended value ranges between 0.7 and 1.2. As seen in Table 32, quick ratio in all years of the monitored period corresponds to the recommended range. As for current liquidity, its values range between 0.73 and 1.09, where the recommended range is 1.5–2. This ratio indicates how many times the company would be able to pay its short-term liabilities if all current assets are converted into cash.

Table 32. Results of ratios for the company Správa železnic, státní podnik.

	2015	2016	2017	2018
Return on assets	−1.46	−2.89	−1.62	−0.81
Return on equity	−1.94	−3.78	−2.01	−0.99
Return on sales	−15.88	−31.55	−15.84	−8.17
Current ratio	0.73	0.86	1.05	1.09
Quick ratio	0.63	0.8	0.93	0.99
Cash position ratio	0.11	0.48	0.54	0.56
Return on assets	3959.81	3990.52	3560.83	3687.19
Accounts receivable turnover	184.38	103.08	87.29	106.94
Payables turnover ratio	264.59	362.26	238.77	258.63
Inventory turnover	39.14	22.57	28.34	25.93
Debt ratio	22.82	23.43	19.23	18.08
Debt-to-equity	30.33	30.69	23.88	22.15
Financial leverage	1.33	1.31	1.24	1.23
EBITDA	3,111,929,000	1,813,509,000	2,939,349,000	3,642,561,000
Liability turnover ratio	496.94	666.49	502.5	540.72
Debtor days ratio	199.79	113.14	92.96	108.76
Debt coverage ratio	24.75	23.67	19.46	18.38
Financial debt coverage ratio	18.22	12.79	24.2	27.92
Fixed asset coverage ratio	0.93	0.94	0.96	0.98
Bank indebtedness	7.49	4.47	1.75	0.9
Financial ratio	0.75	0.76	0.81	0.82
Contribution margin	−76.75	−85.86	−85.34	−117.38
Contribution margin break—even point		146.80	166.71	213.53
Wages/Sales	135.08	138.78	141.96	162.55
Wages/Added value	−95.06	−116.23	−171.06	−118.28
Cost efficiency ratio	−3.4	−3.93	−2.22	−0.96
Return on investment	4.16	3	4.66	5.05

Source: Author.

The items including the word “turnover” indicate how long the assets remain in their initial form before being converted into sales or cash. In the case of assets, this period should be as short as possible, while in the case of liabilities, it is the other way around. Financial leverage shows the ratio of debt capital to the total assets of a given company. The higher the financial leverage (smaller share of equity on the total resources), the greater the effect of the financial leverage is on return on equity.

4.7.2. Company with the Registration Number 5886—Dopravní Podnik hl. Města Prahy

Table 33 shows the financial analysis using ratios for the years 2015–2018. Comparison of the tables for both companies shows that the results of the company 5886 are slightly better. Return on assets is significantly shorter than in the case of the company 70994234, while financial leverage is comparable with the company Správa železnic, státní podnik. EBITDA is the so-called earnings before Interest, Taxes, Depreciation, and Amortization; it shows the operating performance of a company. Table 33 shows that its value is significantly higher than in the case of the company 7099423. On the contrary, liquidity shows worse values. Current ratio does not achieve even the lower limit of the recommended value in any year of the monitored period. Quick ratio shows better values with the values being in the recommended range. Cash position ratio exceeds the limits of the recommended range for all years, which indicates inefficient management.

Table 33. Results of ratio for the company Dopravní podnik hl. města Prahy.

	2015	2016	2017	2018
Return on assets	1.56	2.32	2.34	1.85
Return on equity	2.05	2.91	2.89	2.27
Return on sales	7.11	10.44	10.22	7.89
Current ratio	0.81	0.71	0.92	1.04
Quick ratio	0.77	0.67	0.85	0.97
Cash position ratio	0.66	0.56	0.65	0.78
Return on assets	1663.57	1641.74	1589.94	1553.53
Accounts receivable turnover	25.68	19.22	31.89	22.98
Payables turnover ratio	230.8	201.58	172.53	151.23
Inventory turnover	8.16	9.44	11.36	10.6
Debt ratio	23.43	19.27	18.12	17.46
Debt-to-equity	30.93	24.18	22.37	21.42
Financial leverage	1.32	1.25	1.23	1.23
EBITDA	4,970,710,000	4,883,945,000	5,803,489,000	5,598,106,000
Liability turnover ratio	356.36	305.09	278.1	263.04
Debtor days ratio	27.72	21.37	36.16	28.86
Debt coverage ratio	24.27	20.28	19.01	18.52
Financial debt coverage ratio	22.51	33.79	35.82	34.21
Fixed asset coverage ratio	0.95	0.94	0.98	0.99
Bank indebtedness	1.28	0	0	0
Financial ratio	0.76	0.8	0.81	0.81
Contribution margin	48.29	46.81	47.45	46.61
Contribution margin break-even point	38.75	60.45	65.18	63.90
Wages/Sales	36.05	37.6	37.79	38.51
Wages/Added value	54.99	63.63	61.02	62.32
Cost efficiency ratio	6.96	8.03	8.74	6.5
Return on investment	5.27	6.51	6.49	5.97

Source: Author.

5. Discussion

Within this contribution, one research question was formulated on whether there are only a few companies that largely affect the growth in the transport sector. The research question was answered in the research section of this contribution. The dataset including the accounting data of all companies was presented to the Kohonen network for training; the result was a map of clusters graphically representing the presence and number of companies in individual clusters. Within the research, the so-called business value generators, on the basis of which it was determined to what extent the individual clusters participated in their creation. A ranking of the strongest clusters was compiled for individual years and individual variables, where the first position was mostly occupied by the cluster (1, 1), which was imaginarily replaced by the cluster (10, 10) in the year 2017. All clusters were assigned points based on their position in the individual tables; subsequently,

a ranking of the strongest clusters for all variables in all years of the monitored period was compiled. The strongest cluster turned out to be the cluster (1, 1) containing 3–5 companies between 2015 and 2018. The second strongest cluster was the cluster (2, 1). As the strongest cluster in terms of the largest volume of assets, fixed assets, sales, and operating result in the years 2015–2018, it contained 3–5 companies, thus providing the answer to the research question: Are there only a few companies that have a great effect on the growth of the transport sector? The answer to this research question is thus yes. The transport sector is to a large extent influenced only by 3–5 companies, which are interconnected with the transport sector in such a way that if the companies included in the strongest cluster (1, 1) make a profit, stagnate, or even go bankrupt, this will have a significant influence on the transport sector of the CR. The companies that appeared in the cluster (1, 1) in all four years of the monitored period are Správa železnic, státní podnik and Dopravní podnik hl. města Prahy. These companies differ from each other in terms of their financing and management; however, both were identified as leading companies in the transport sector of the Czech Republic in terms of generating business value. The share in the total assets of all companies in the transport sector included in the original dataset ranges between 15.82% and 20.01%. This means that specifically Dopravní podnik hl. města Prahy, with its share of 20.01% makes up more than a fifth of the total assets of all companies. In the case of fixed assets, this company achieves the highest share in the year 2016 (26.77%). In the case of operating result, its share is highest in the year 2017, when it achieves 13%. A similar application of Kohonen networks was discussed by (Stehel et al. 2019), whose aim was to analyze companies operating in the agriculture sector of the Czech Republic using the Kohonen network and subsequent prediction of their further development. They created a dataset, which contained, as in the case of the research concerning the transport sector, data from financial statements of 4201 companies operating in the agriculture sector of the Czech Republic in the year 2016. The dataset was generated based on Bisnode's Albertina database. The dataset was subsequently subject to cluster analysis using the Kohonen network. The advantages of using the Kohonen network are also mentioned by (Du Jardin and Séverin 2012), whose results show that out-of-sample error achieved by the map remains more stable over time than the error achieved by traditional methods used for proposing failure models (discriminant analysis, logistic regression, Cox model, and neural networks).

6. Conclusions

The transport sector is very important for the Czech Republic, as it is interconnected with other sectors and represents a great contribution to the economy of the Czech Republic. The transport sector includes mainly companies providing services. The goal of the research was to analyze the financial situation of the companies operating in the transport sector and the whole transport sector of the Czech Republic using artificial intelligence, specifically Kohonen networks. The contribution pointed to the importance of the transport sector and explained how cluster analysis and specifically an artificial intelligence method—Kohonen network—can be used for creating clusters based on similar characteristics; it also presented the clustering system and the importance of clusters as such. The methodology chapter was quite extensive because of the necessity to explain the methods and procedures used in the research and to specify the methods used for identifying the strongest cluster in the transport sector based on raw data, since the neural networks and working with them is a complex and exacting discipline. The methodology also included the information on the dataset used for the training and on the modification of the dataset so that it could be presented to the network. Both companies included in the cluster (1, 1) in all years of the monitored period were finally subject to financial analysis and subsequently described. The performance of a thorough analysis of the economic development of the transport sector using Kohonen network and subsequent prediction of its development enabled meeting the objective of the research. The added value of the research and its benefits can be seen in its complexity and details, which provided the information on what companies most affect

the transport sector and how. This approach can be considered as being of general validity; it is thus applicable for other sectors or groups of companies. It also provides analysts with the opportunity to monitor the situation in these companies and in the case of fall or sudden growth, to focus on specific companies that are most likely to have given rise to the situation, since they had undergone the changes that affected the whole transport sector. The limitations of the research consist in the lack of international reach; however, this can be a subject of further investigation. With a slight modification of the choice of variables (business value generators), this model would be applicable for any group of companies in the same or a different sector. A possible improvement of the research can be the application of an optimized Kohonen network, as done by (Harchli et al. 2014), who used the generation of a heuristic method prior to the phase of network learning. The main objective was to find the initial parameters of the map, which means finding centroids in the most homogeneous areas of the dataset. Subsequently, the Kohonen learning phase was started and the obtained clustering was evaluated using the map. These two phases were repeated until reaching a desired number of iterations. The result shows that the proposed method can provide better clustering results than a traditional topological map. In further research on the application of the Kohonen map to certain sectors or groups of companies, the optimization of the model could thus provide better results.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this work do not come from public sources.

Acknowledgments: Many thanks to reviewers for improving the quality of the paper.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Table A1. Selected financial ratios (NACE classification), transport and storage.

Ratio	Unit of Measurement	Year									
		2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Number of active companies		41,785	42,094	41,873	41,232	40,064	38,944	38,610	38,159	39,016	39,791
Number of employed persons in total—in natural persons	persons	305,801	292,676	284,021	279,027	267,444	262,886	264,532	271,389	281,680	287,996
From the above: Average registered number of employees—in natural persons	persons	261,887	248,932	240,301	234,335	225,906	220,100	221,430	227,885	236,524	242,654
Average registered number of employees—equivalent	persons	256,886	245,192	236,696	231,002	222,648	217,010	218,578	225,023	233,348	239,954
Average gross monthly wage per natural person	CZK	22,233	22,332	22,351	22,480	22,686	22,657	23,172	23,871	24,963	26,656
Average gross monthly wage per 1 equivalent person	CZK	22,666	22,672	22,692	22,804	23,018	22,980	23,475	24,175	25,302	26,956
Net turnover (Total revenues)	million CZK	631,182	539,302	580,318	591,781	608,224	619,828	649,497	634,861	647,091	679,345
including Sales of products and services and sales of goods	million CZK	541,165	470,076	512,957	528,654	543,574	547,580	575,238	572,396	584,773	615,837
including Sales of products and services, Sales of goods	million CZK	455,460	399,637	438,805	449,196	449,240	447,626	476,725	499,679	516,451	547,284
Other operating income	million CZK	85,705	70,440	74,152	79,459	94,334	99,954	98,513	72,717	68,322	68,553
Income including trade margin	million CZK	73,739	57,600	52,002	54,057	56,341	61,924	67,724	56,389	52,768	53,165
Costs in total	million CZK	461,202	404,484	443,254	454,118	454,000	452,271	481,334	504,236	521,183	552,565
including Consumption of material and energy, costs of services	million CZK	606,610	525,103	551,443	567,717	586,484	600,027	624,401	604,586	617,987	648,669
including Cost of goods sold	million CZK	332,138	284,139	308,370	321,987	324,396	323,289	342,302	350,662	362,646	385,880
Other operational costs	million CZK	81,813	66,878	70,798	75,670	91,027	96,658	95,262	69,485	65,025	64,899
Personnel costs	million CZK	45,756	34,619	38,149	37,427	44,507	46,209	56,728	42,652	36,932	34,525
including Wages without UN	million CZK	98,286	93,806	91,816	89,730	86,727	85,400	87,520	92,766	100,849	110,595
Trade margin	million CZK	69,870	66,709	64,452	63,213	61,499	59,842	61,572	65,279	70,851	77,618
Share of trade margin on sales of goods	%	3892	3561	3354	3788	3308	3295	3251	3232	3297	3653
Added value	million CZK	4.5	5.1	4.5	4.8	3.5	3.3	3.3	4.4	4.8	5.3
Earnings for current accounting year	million CZK	129,064	120,346	134,884	132,130	129,605	128,981	139,032	153,574	158,537	166,685
Net assets	million CZK	24,572	14,199	28,875	24,064	21,740	19,801	25,096	30,274	29,104	30,677
	million CZK	554,067	551,091	559,267	540,713	537,802	548,815	559,505	572,898	592,724	605,721

Table A1. Cont.

Ratio	Unit of Measurement	Year										
		2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	
including	Fixed intangible assets (net)	million CZK	3478	3685	3576	5911	5927	5716	5477	5296	5714	5631
	Fixed tangible assets (net)	million CZK	344,678	348,973	346,560	328,366	334,975	326,886	331,924	335,636	338,762	340,331
	Fixed financial assets (net)	million CZK	29,079	28,499	29,195	27,802	22,559	28,845	31,274	33,026	33,118	34,488
	Inventories without advances provided (net)	million CZK	12,140	11,005	9893	10,398	10,369	9896	9971	9581	10,723	10,544
Liabilities	receivables (net)	million CZK	108,115	104,888	114,827	115,860	107,675	115,193	109,395	113,498	127,337	134,200
including	equity	million CZK	554,067	551,091	559,267	540,713	537,802	548,815	559,505	572,898	592,724	605,721
	liabilities	million CZK	342,575	343,616	352,957	317,118	308,863	310,668	284,611	293,220	312,711	323,313
Acquisition of fixed tangible and intangible assets	million CZK	195,036	192,106	191,197	209,726	216,818	226,346	262,906	266,519	266,689	268,377	
		million CZK	59,782	42,473	45,149	56,050	47,716	50,857	46,651	51,919	49,219	42,720

Source: CSO.

Appendix B

Table A2. Overall ranking of clusters.

Clusters	Number of Points in Total
(1, 1)	108
(1, 2)	76
(1, 10)	69
(2, 1)	66
(1, 3)	62
(3, 1)	56
(2, 10)	53
(1, 9)	45
(10, 10)	39
(9, 10)	35
(2, 2)	35
(5, 1)	34
(1, 8)	24
(10, 9)	21
(1, 4)	16
(3, 10)	15
(8, 10)	15
(1, 5)	14
(4, 2)	13
(3, 2)	11
(2, 6)	11
(2, 3)	9
(8, 9)	8
(2, 9)	7
(8, 8)	7
(3, 4)	6
(6, 2)	6
(3, 3)	6
(5, 3)	5
(2, 4)	3
(6, 10)	2
(6, 1)	1
(9, 7)	1
(1, 7)	1

Source: Author.

References

- Abdelkafi, Inès, Manel Zribi, and Rochdi Feki. 2018. New Classification of Developed and Emerging Countries Based on the Effects of Subprime Crises: Kohonen Map Method. *Journal of the Knowledge Economy* 9: 908–27. [\[CrossRef\]](#)
- Awrejcewicz, Jan, Yu B. Lind, Aigul R. Kabirova, Liana F. Nurislamova, Irek M. Gabaidullin, and Rinat A. Mulyukov. 2011. Prediction of drilling fluids loss during oil and gas wells construction. Paper presented at the 11th Conference on Dynamical Systems—Theory and Applications, Łódź, Poland, December 5–8; pp. 163–70.
- Bakoben, Maha, Tony Bellotti, and Niall Adams. 2017. Identification of credit risk based on cluster analysis of account behaviours. *Journal of the Operational Research Society* 71: 775–83. [\[CrossRef\]](#)
- Cahyana, Bambang Eka, Umar Nimran, Hamidah Nayati Utami, and Mohammad Iqbal. 2020. Hybrid cluster analysis of customer segmentation of sea transportation users. *Journal of Economics, Finance and Administrative Science* 25: 321–37. [\[CrossRef\]](#)
- Cho, Sungbin, Jinhwa Kim, and Jae Kwon Bae. 2009. An integrative model with subject weight based on neural network learning for bankruptcy prediction. *Expert Systems with Applications* 36: 403–10. [\[CrossRef\]](#)
- Du Jardin, Philippe, and Eric Séverin. 2012. Forecasting financial failure using a kohonen map: A comparative study to improve model stability over time. *European Journal of Operational Research* 221: 378–96. [\[CrossRef\]](#)
- Feranecova, Adela, Martina Sabolova, and Peter Remias. 2016. Cluster analysis of automotive industry companies. *Actual Problems of Economics* 182: 95–102.
- Harchli, Fidae, ES-Safi Abdelatif, and Ettaouil Mohamed. 2014. Novel method to optimize the architecture of kohonen's topological maps and clustering. Paper presented at the 2nd IEEE International Conference on Logistics Operations Management, GOL 2014, Rabat, Morocco, June 5–7; pp. 8–14. [\[CrossRef\]](#)

- Kohonen, Teuvo. 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics* 43: 59–69. [CrossRef]
- Kohonen, Teuvo. 1986. Self-organization, memorization, and associative recall of sensory information by brain-like adaptive networks. *International Journal of Quantum Chemistry* 30: 209–21. [CrossRef]
- Lakhawat, Piyush, and Arun K. Somani. 2017. Big Data Cluster Analysis: A Study of Existing Techniques and Future Directions. In *Big Data Analytics: Tools and Technology for Effective Planning*, 1st ed. Edited by Arun K. Somani and Ganesh Chandra Deka. New York: Chapman and Hall/CRC, p. 413.
- Liashenko, Olena, Tetyana Kravets, and Tetiana Verhai. 2018. Estimation of the Production Potential of Ukraine's Regions Using Kohonen Neural Network. Paper presented at the CEUR Workshop Proceedings, 2105, ICTERI 2018, Kyiv, Ukraine, May 14–17; pp. 19–34.
- Li, Hui, Hong Lu-Yao, Zhou Qing, and Yu Hai-Jie. 2015. The assisted prediction modelling frame with hybridisation and ensemble for business risk forecasting and an implementation. *International Journal of Systems Science* 46: 2072–86. [CrossRef]
- MPO. 2016. Metodika Výpočtu—Analytické Materiály a Statistiky [Online]. Available online: <https://www.mpo.cz/assets/cz/rozcestnik/analyticke-materialy-a-statistiky/2016/11/metodika-vypoctu.pdf> (accessed on 21 June 2021).
- Rayala, Venkat, and Satyanarayan Reddy Kalli. 2021. Big data clustering using improvised fuzzy C-means clustering. *Revue d'Intelligence Artificielle* 34: 701–8. [CrossRef]
- Rodrigue, Jean-Paul, Claude Comtois, and Brian Slack. 2013. *The Geography of Transport Systems*, 3rd ed. New York: Routledge, p. 432.
- Rowland, Zuzana, and Jaromír Vrbka. 2016. Using artificial neural networks for prediction of key indicators of a company in global world. In *Globalization and Its Socio-Economic Consequences, 16th International Scientific Conference Proceedings*. Žilina: EDP Sciences, pp. 1896–903.
- Stehel, Vojtěch, Jakub Horák, and Marek Vochozka. 2019. Prediction of institutional sector development and analysis of enterprises active in agriculture. *E a M: Ekonomie a Management* 22: 103–18. [CrossRef]
- Svobodová, Hana, Antonín Věžník, and Eduard Hofmann. 2013. *Vybrané kapitoly ze socioekonomické geografie České republiky*, 1st ed. Brno: Masarykova Univerzita, p. 163.
- Tyukhova, Elena, and Dmitry Sizykh. 2019. The cluster analysis method as an instrument for selection of securities in the construction of an investment portfolio. Paper presented at the 2019 12th International Conference "Management of Large-Scale System Development", MLSD 2019, Moscow, Russia, October 1–3; p. 621.
- Vahalík, Bohdan, and Michaela Staníčková. 2016. Key factors of foreign trade competitiveness: Comparison of the EU and BRICS by factor and cluster analysis. *Society and Economy* 38: 295–317. [CrossRef]
- Vochozka Marek. 2020. *Metody Komplexního Hodnocení Podniku*, 2nd ed. Prague: Grada Publishing.
- Vochozka, Marek, and Veronika Machová. 2017. Enterprise value generators in the building industry. *SHS Web of Conferences* 39: 01029. [CrossRef]