

Detection of Outliers as a Possible Method of Accounting Fraud Revelation

Marcela Hradecká

Abstract: *The globalization and the development of enterprising environment allowed the rise of many multinational corporations all over the world. At the same time there appears the phenomenon of sophisticated methods of accounting fraud and tax optimization. Tax avoidance and tax evasion with the aim to reach the highest savings in paying taxes. Financial and tax authorities of individual states have to react to the new situation quickly and to apply the most efficient and effective methods in accounting fraud detection and prevent corporate tax evasion on global scale.*

Both financial and non-financial methods are used in the fraud detection. This article is concerned with the possible use of the outliers method for the accounting fraud detection. This method refers to the extreme values analysis, their detection and demonstrability and draws attention to numerical values that are unusual in some way.

Key words: Fraud · Detection Methods · Outlier · Measures Clustering · Dendrogram

JEL Classification: M2 · M41 · M42

1 Introduction

Accounting undergoes constant development. In reaction to the global changes and to harmonize international accounting the Council of the European Union issued in June 2013 the new Directive 2013/34/EU on the annual financial statements, consolidated financial statements and related reports of certain types of undertakings amending Directive 2006/43/EC and replaced the 4th and 7th EU Directives. The objective of the EU strategy for growth and employment entitled „Europe 2020“ for smart and sustainable growth is to offer a simplified regulatory framework to encourage entrepreneurship and innovation where the SMEs are provided with the right information and supplementary support in markets outside the EU and the bureaucratic load is limited.

Simplification and reduction of administrative load in the area of reporting, accounting and financial information, however, will not prevent creative and sophisticated methods of accounting fraud. Particularly in the area gross operating result of turnover. A really risky area is the services provision often referred to as *transfer services* and the related methods of calculation of the transfer price.

Outliers Detection methods can be classified as one of the possible methods for accounting fraud detection. Extremes and extreme values are part of standard analysis of processes and data. The term “extreme” is usually understood as data which reached an unexpected value or size. This means that they refer to extremely high value or unusual value aberration.

2 Methods

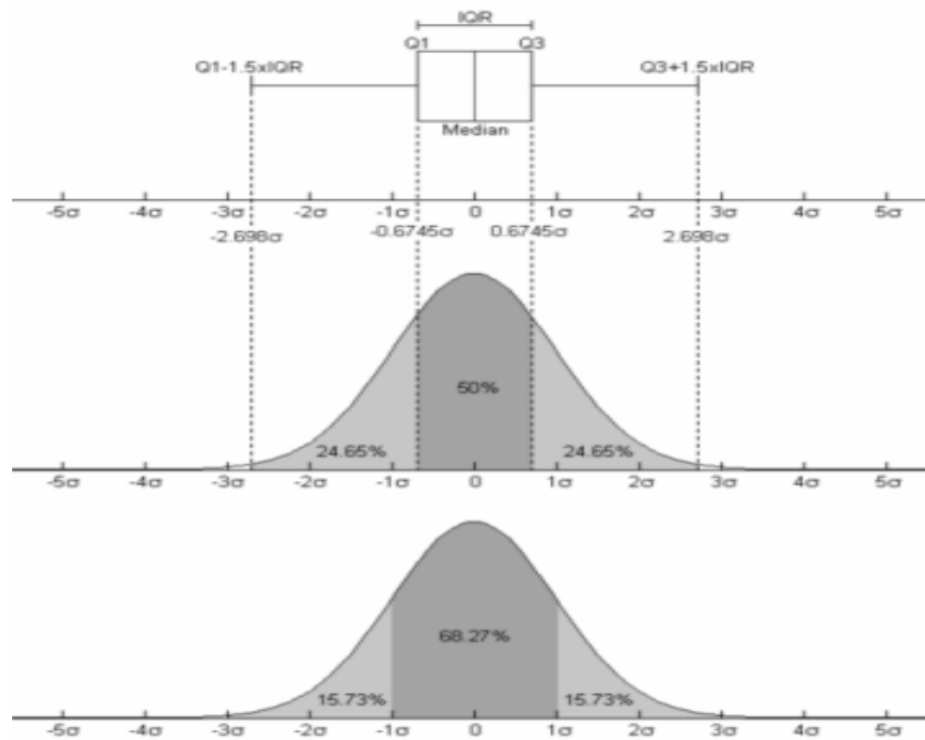
Mathematization of economics and quantitative methods are at present considered a part of research. Outliers Detection methods are divided into one-dimensional and multi-dimensional method. Detection is realized by mathematical analysis of data and can be illustrated by Scatterplot (correlation point diagram). Correlation is the relation between two processes or values. Another possible representation is by Boxplot which is a method for graphically depicting groups of numerical data through their quartiles.

The detection of extremes requires definition of extreme values. Fundamental for the reviewer is whether the extreme value is in the frequency or data area, The extremes in data set “source matrix” are distinguished

- Extreme values in the frequency area. Frequency of the occurrence of extreme values that excess the frequency of other values of the researched data. The characteristic feature is therefore the frequency of extreme data. Frequency

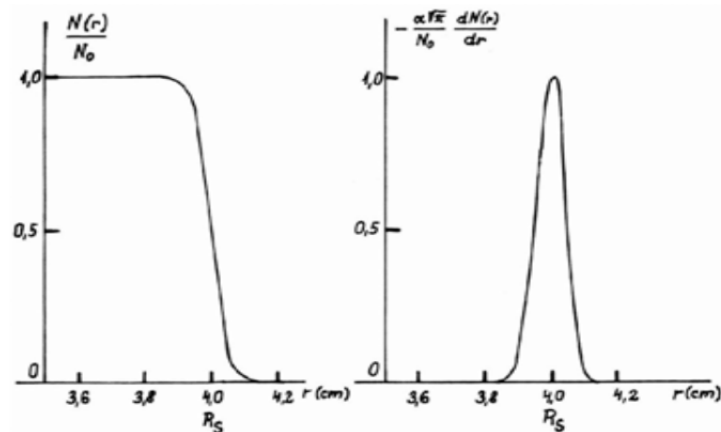
extremes can be visualized by any graphic rendering which shows the frequency or by percent ratio of the individual groups of data in the total data set.

Figure 1 Boxplot with quartiles for the occurrence of outliers in normal distribution



Source: Voženílek, Kaňok, Tuček (2008)

Figure 2 Frequency extremes displayed by graph of probability density



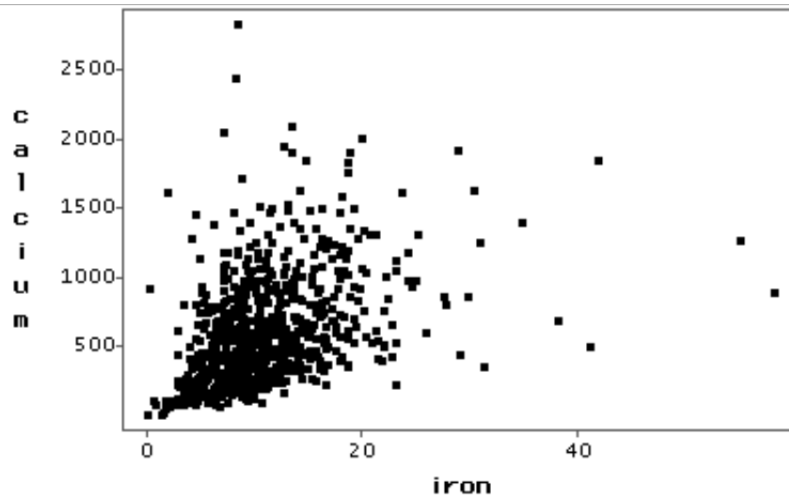
Source: Voženílek, Kaňok, Tuček (2008)

- Extreme values in the data area of the dataset.** Extreme values that exceed or do not reach the characteristic values that are commonly occurring in the researched data. Data extremes can be defined as lying beyond the clustering of the data set. The visualization of the extreme values in the data area can be effectively done by scatter plot.

Analysis of extremes – outliers, consequently proves the truth of the claim of extremity of the researched value. Sophisticated multi-dimensional methods are used for such analysis. These can prove that the value is extreme. While searching for the extreme it is necessary to distinguish the set realities. Extreme values are of two types. They are defined as values set by precepts or laws or as values set by statistical methods. During the analysis of a large set of data the values of median and arithmetic mean. An example of a sophisticated method is the cluster analysis. Traditional application of the cluster analysis is the acquisition of detailed description and gaining of insight into the sets of data in

taxonomy. This method aggregates the data of common sets of data on the basis of similarity or dissimilarity and also the distance. The extent of similarity of data is crucial when conducting the cluster analysis. The objective is to divide the set of data into two or more groups, that is clusters based on the similarity of the values of the data. Mutual points of the values of data in multi-dimensional space can also be described by the distance and subsequently graphically visualize their description.

Figure 3 Scatterplot for extreme presentation within data area



Source: Voženilek, Kaňok, Tuček (2008)

Multi-dimensional analysis should be:

- Suitable for resolving the given problem area
- Correctly and precisely calculated according to the set parameters
- Well presented and its conclusions should contribute to the given research

Quantitative results of clustering depend on the existence of clusters in data and on the given criteria of the cluster definition. The result of the analysis is the dendrogram (cluster tree), where there the similar values (points) are in the same cluster while different values will be in another cluster.

In my many years of experience, I have been systematically engaged in identifying problematic items in accounting statements in terms of creative accounting and fraud. Their testing with a focus on corporations and selected directions. In particular payout ratio or fictitious services in performance consumption. The aim of the research is to analyze problematic items of accounting statements in terms of creative accounting and fraud.

Another partial objective is to create a set of synthetic indicators for fraud detection in accounting. The intention is to prevent the reduction of corporate taxes on a global scale. You can use the Cluster Analysis method to detect a large amount of data (corporations) in the data area, allowing multiple dimensions to be tracked within one analysis across selected corporations. The corporation can be assigned to a data set by subject of economic activity. For example, in price controls in the health sector or in comparison to the provision of transfer services in relation to the contribution to the development of the corporation.

The results of the analysis are dependent on the correct input parameters. The analysis carried out in the framework of the research was focused on services in performance consumption in relation to the created value added in the given accounting period and on the created contribution to the development of the corporation. According to the survey, this indicator is a "contribution to corporate development" more important than indicator of the payout ratio. The payout ratio expresses how much of the net profit was paid to shareholders and what portion of the net profit was transferred to the reinvestment of the corporation. The indicator for the development of the corporation analyzes what part of the added value created, adjusted for labor costs, or asset write-offs remains for the development of the corporation after the eventual payment of net profit to the shareholders.

3 Research results

Cluster analysis was used for the detection and identification of outliers. This method is very sensitive on the presence of non-significant outliers that rise over the usual values of data. The research that used the multi-dimensional method of cluster analysis was conducted on a sample of corporations selected on the basis of criteria (character of the economic activity, size of the corporation, employer, and so on) so that the sample contained corporations from different branches. The selected detection method allows the analysis of data from accounting books, managerial reporting and account sheets including closing the book and the profit and loss accounts. The amendment of the Law 563/1991 Coll. on Accounting and the Implementing regulations 500/2002 Coll., led to significant changes in accounting since 1st. January 2016. This refers especially the extent of the entries of the Profit and loss account. The implementation of the Directive 2013/34/EU on the annual financial statements, and consolidated financial statements into the Czech legal framework led to simplification of accounting data with micro and small accounting units (corporations). This simplification, however, has a negative impact as it makes the detection of accounting fraud more difficult. There were changes also in the structure of entries in the other accounting units (corporations) Profit and loss account. Some entries such as “added value” and “margin” were excluded from the list of entries.

The outliers detection method was applied on selected data and entries of Profit and loss account of selected corporations. The objective of the analysis was the detection of outliers in the data area. The calculated values of the accounting indicators were put into the source matrix.

Input variables for the source matrix:

- Indicator – Services / Added value
- Indicator – Services / Added value adjusted
- Indicator ROS – EBIT (Return on sales - Earnings before interest and taxes) / Performance (Sales from the main activity)
- Indicator PM – EAT (Earnings after Taxes) / Sales derived from the main activity
- Indicator – Contribution to development of corporation (Retained earnings) / Added value adjusted

Table 1 Input variables for the source matrix

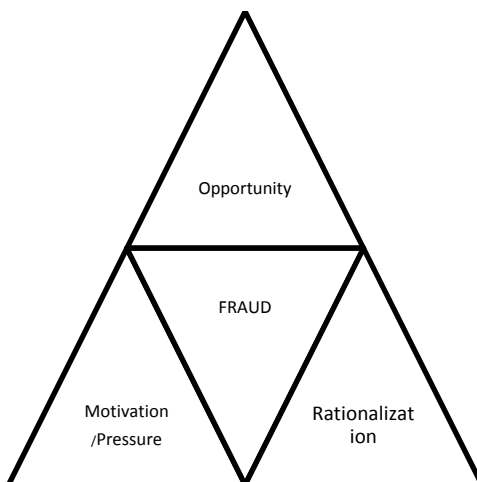
Name of corporation	SE / AV adjusted %	SE / AV %	ROS %	PM %
Corporation 1	214,25	58,35	13,54	11,72
Corporation 2	155,03	57,80	21,61	17,28
Corporation 3	271,84	27,64	0,69	0,69
Corporation 4	381,08	80,43	9,26	7,15
Corporation 5	294,35	63,76	10,76	10,76
Corporation 6	403,95	71,69	8,11	6,55
Corporation 7	171,71	65,54	7,40	5,75
Corporation 8	198,68	90,17	16,47	13,15
Corporation 9	136,57	53,59	7,51	5,91
Corporation. 10	134,60	47,59	26,87	21,59
Corporation 11	546,64	87,20	4,17	4,17
Corporation 12	177,30	91,42	20,59	20,59

Source: Author's processing

It is evident from the source matrix that the indicator of the ratio Services to Added value represents what percentage of the added value do the services represent. The added value is the value which is created by processing, which every provider of services or seller adds by its effort to the value of the purchased or their own semi-products. This indicator is rather unprecise, because the added value in the Profit and loss account does not include the personnel costs (salaries) or amortization amount. The analysis was therefore extended to include the indicator of the ratio between Services and Added value adjusted by the personnel costs (salaries). From the defined ratio it is clear that if the corporations calculated the prices for provided services correctly and in due form the indicator would hardly reach such extreme levels. This partial result of the conducted research confirmed the fact that the calculation of provided services and sold products is not a commonly encountered part of the internal instructions of corporations. The consequence of the

non-existing price calculation of goods and services can be the distorted economic result but also the opportunity for fraud.

Figure 4 Fraud triangle



Source: Author’s graphic processing (<https://www.bccpa.ca>)

The indicator Services to Added value ratio was for the further analysis excluded and replaced by the indicator Contribution to development of corporation / Added value adjusted Contribution to development of corporation (Retained earnings) can be understood as an amount which remains in the corporation after payment of EAT (Earnings after Taxes) to the stock owners or partners.

Table 2 Input variables for the source matrix

Name of corporation	PM – EAT %	ROS %	SE / AV %	Retained earnings / AV adjusted %
Corporation 1	11,72	13,54	58,35	30,44
Corporation 2	17,28	21,61	57,80	23,64
Corporation 3	0,69	0,69	27,64	91,09
Corporation 4	7,15	9,26	80,43	38,82
Corporation 5	10,76	10,76	63,76	16,39
Corporation 6	6,55	8,11	71,69	35,64
Corporation 7	5,75	7,40	65,54	60,38
Corporation 8	13,15	16,47	90,17	24,88
Corporation 9	5,91	7,51	53,59	56,67
Corporation 10	21,59	26,87	47,59	7,63
Corporation 11	4,17	4,17	87,20	38,08
Corporation 12	20,59	20,59	91,42	-17,81

Source: Author’s processing

It is clear from the source matrix that the resulting values of the business performance ratio indicator for adjusted value added represent how much of the created value added is retained in the corporation (corporation) for further development and growth of the corporation. The optimum of this indicator ranges from 38% to 60%.

Presentation of Cluster Analysis Conclusions

Data (the proper variables) in the source matrix were analyzed through the all combination of methods and distances of the cluster analysis. The Nearest Neighbour Method combined with the Euclidean distance based on the minimal distance of the value were used for the detection of outliers in accounting.

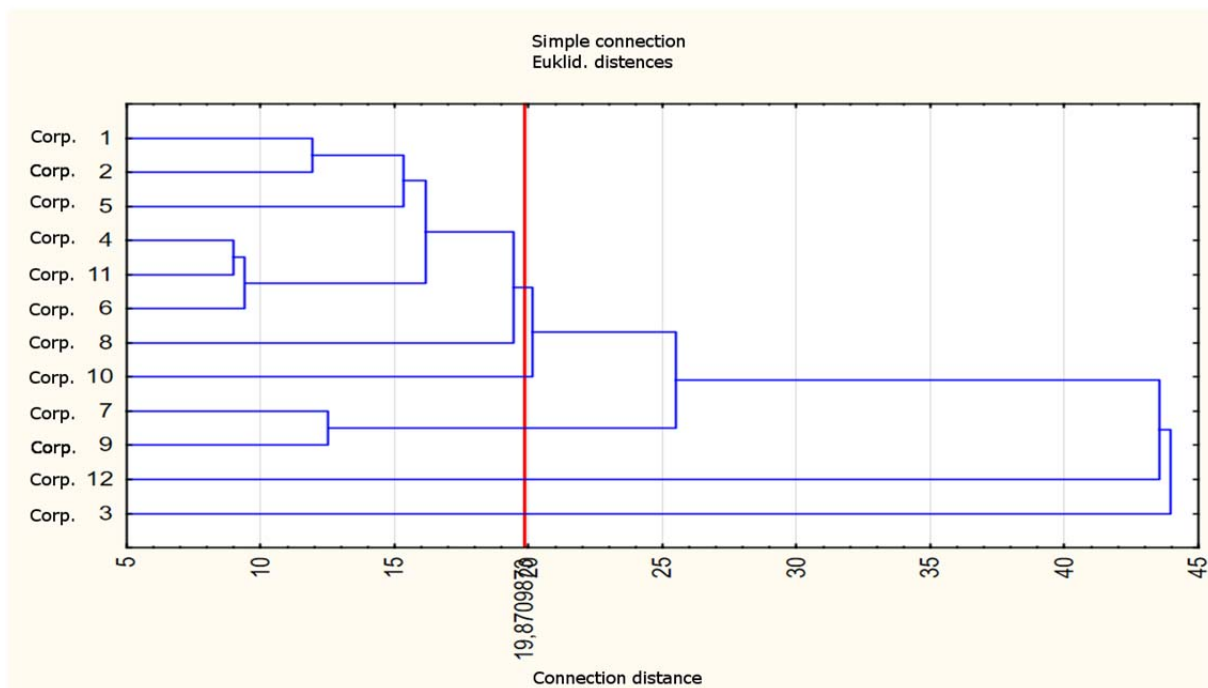
The output of the multi-dimensional method “Cluster analysis” are the separate „Dendrograms“ from selected combinations of methods and distances.

Table 3 Resultant Distances Matrix

Distances connection	Merger schedule											
	Obj. N.1	Obj. N.2	Obj. N.3	Obj. N.4	Obj. N.5	Obj. N.6	Obj. N.7	Obj. N.8	Obj. N.9	Obj. N.10	Obj. N.11	Obj. N.12
27,8099	Corp. 9	Corp. 10										
40,5752	Corp. 4	Corp. 6										
40,98258	Corp. 2	Corp. 12										
45,27614	Corp. 1	Corp. 7										
46,19888	Corp. 1	Corp. 7	Corp. 9	Corp. 10								
49,8458	Corp. 1	Corp. 7	Corp. 9	Corp. 10	Corp. 8							
63,60174	Corp. 1	Corp. 7	Corp. 9	Corp. 10	Corp. 8	Corp. 2	Corp. 12					
105,5718	Corp. 1	Corp. 7	Corp. 9	Corp. 10	Corp. 8	Corp. 2	Corp. 12	Corp. 5				
162,673	Corp. 1	Corp. 7	Corp. 9	Corp. 10	Corp. 8	Corp. 2	Corp. 12	Corp. 5	Corp. 4	Corp. 6		
437,2879	Corp. 1	Corp. 7	Corp. 9	Corp. 10	Corp. 8	Corp. 2	Corp. 12	Corp. 5	Corp. 4	Corp. 6	Corp. 11	
3960,066	Corp. 1	Corp. 7	Corp. 9	Corp. 10	Corp. 8	Corp. 2	Corp. 12	Corp. 5	Corp. 4	Corp. 6	Corp. 11	Corp.3

Source: Author’s processing using software Statistica

Table 4 Resultant dendrogram



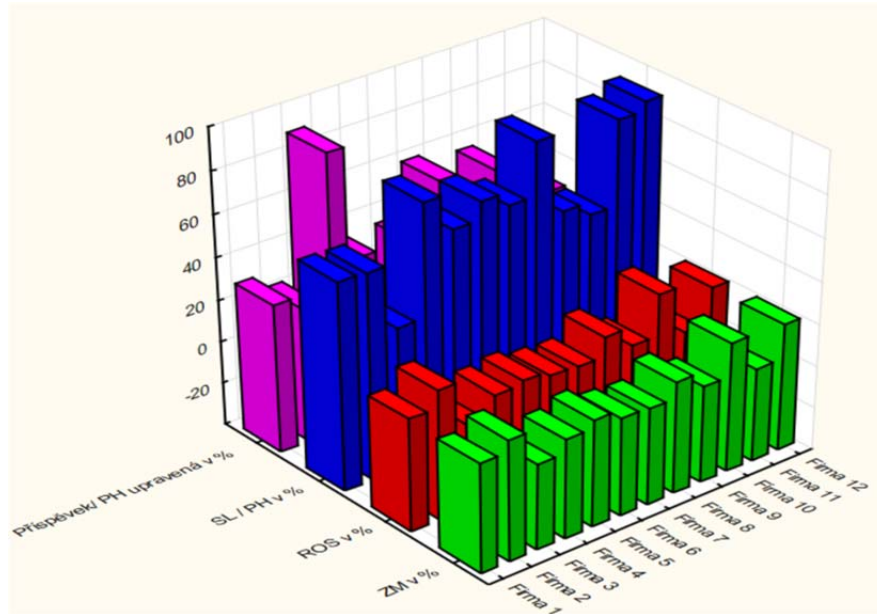
Source: Author’s processing using software Statistica

Resultant dendrogram and the matrix of clusters describes the submission of corporations into clusters according to distance of connectors and presents the detected outliers.

- Indicator – Services / Added value, Ukazatel – Služby / Přidaná hodnota (SL/PH)
- Indicator ROS – EBIT (Return on sales - Earnings before interest and taxes) / Performance (Sales from the main activity, Ukazatel ROS – EBIT (hospodářský výsledek před zdaněním) / Výkony (tržby z hlavní činnosti)
- Indicator PM – EAT (Earnings after Taxes) / Sales derived from the main activity, Ukazatel ZM – EAT (čistý zisk po zdanění) / Tržby z hlavní činnosti
- Indicator – Contribution to development of corporation (Retained earnings) / Added value adjusted, Ukazatel – Příspěvek na rozvoj podniku / Přidaná hodnota PH upravená

The conducted analysis shows that from the selected sample of data and from the pre-set ordinary criteria it is necessary to concentrate on corporations with low values of indicators ROS and PM. And to research in detail the internal and external effects on low cost effectiveness and the profit margin under 10%. In the case of the indicator Services to Added value ratio it is necessary to concentrate on corporations with high value of this indicator. These values indicate wrongly set or missing calculations of provided services or goods and also rather high values of transfer services.

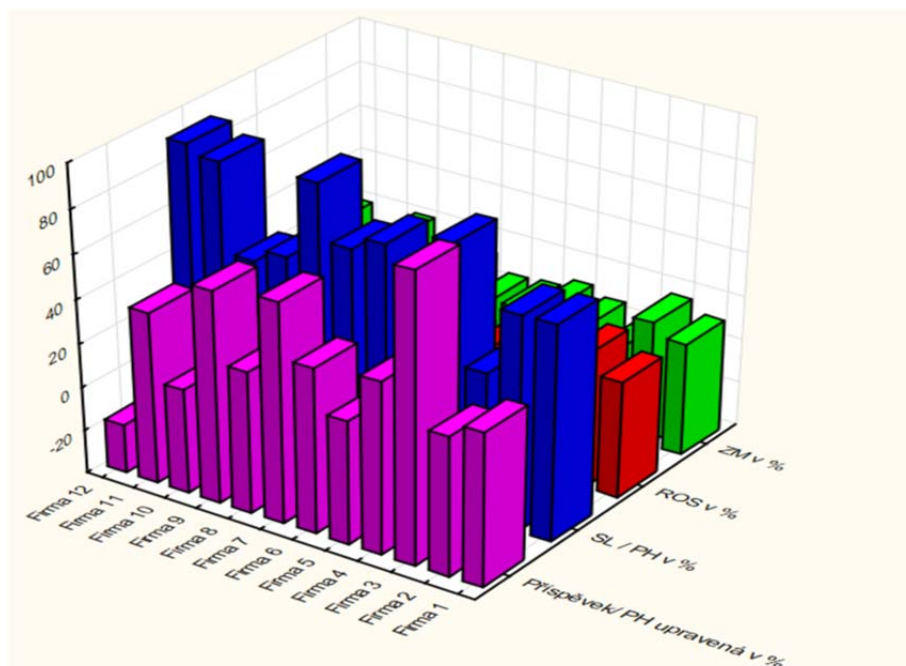
Figure 1 3D Sequence graph



Source: Author’s processing using software Statistica

The last researched new indicator is the ratio between the Contribution to development of corporation (Retained earnings) and the Added value adjusted. While considering the results of the analysis it is necessary to concentrate on corporations with low values. These values indicate wrongly set or missing calculations of provided services or goods. Another criterion of this indicator are the legal norms of the Czech Republic. Principally the Law 90/2012 Coll. on commercial companies and cooperatives (Law on commercial corporations), §40, section 1. Business Corporation shall not pay out net profit if that could lead to its bankruptcy. This means that if stock owners or partners of the corporation prefer the payment of dividends and shares on profit to retaining profits for re-investment the statutory bodies of the corporations have to take into account that some part of the created added value (added value adjusted) have to remain in the corporation for its development.

Graf 2 3D Sequence graph

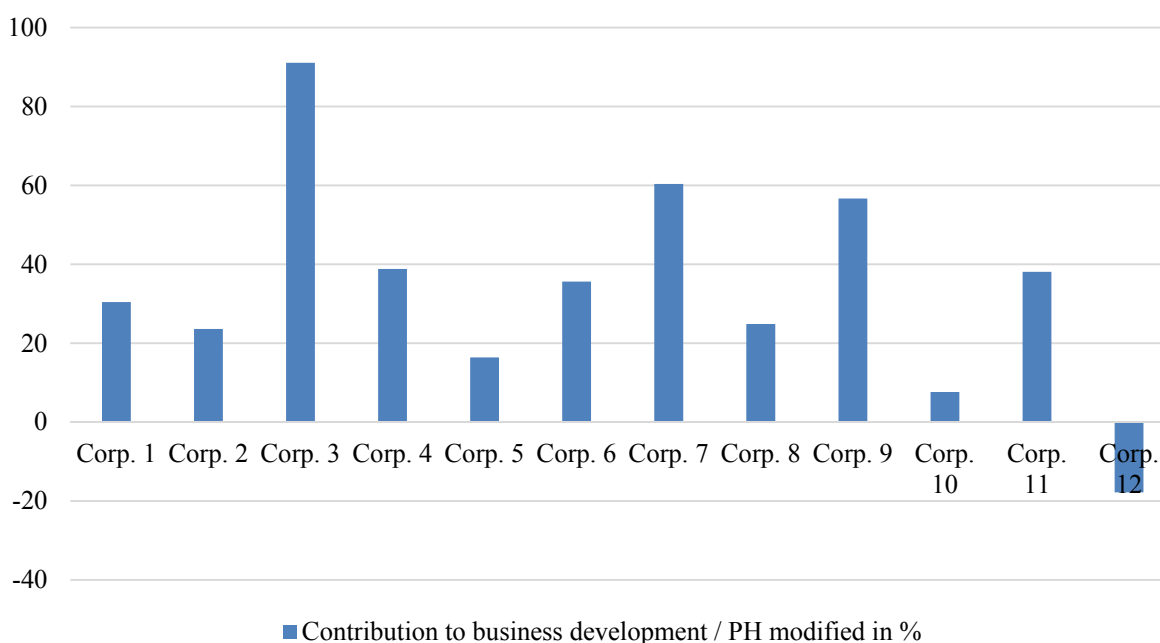


Source: Author’s processing using software Statistica

- Indicator – Services / Added value, Ukazatel – Služby / Přidaná hodnota (SL/PH)
- Indicator ROS – EBIT (Return on sales - Earnings before interest and taxes) / Performance (Sales from the main activity, Ukazatel ROS – EBIT (hospodářský výsledek před zdaněním) / Výkony (tržby z hlavní činnosti)
- Indicator PM – EAT (Earnings after Taxes) / Sales derived from the main activity, Ukazatel ZM – EAT (čistý zisk po zdanění) / Tržby z hlavní činnosti
- Indicator – Contribution to development of corporation (Retained earnings) / Added value adjusted, Ukazatel – Příspěvek na rozvoj podniku / Přidaná hodnota PH upravená

Provided the indicator of the ratio between Contribution to development of corporation (Retained earnings) and the Added value adjusted was researched only separately it could be graphically presented also in the following way.

Graph 3 Indicator of the ratio between Contribution to development of corporation (Retained earnings) and the Added value adjusted



Source: Author's processing

The conclusive research question: „Why should or could the Tax and Financial Authorities use Outliners Taxonomy for detection of accounting fraud? “

Financial Authority processes every year data of many corporations. Outliners Taxonomy allows the analyses of large data files and effectively detect fraud in accounting as well as tax evasions. Multi-dimensional analyses can be used also for detection of wrongly set transfer prices. Principally for setting the usual market price and also the interest on loans to connected persons or otherwise connected parties with the aim of tax avoidance or evasion.

4 Conclusions

The amendment of accounting introduced by the new Directive 2013/34/EU on the annual financial statements brought no effectivity into comparative accounting. Detection of accounting fraud would profit from separate statements assigned by subject matter. The entry of services in the profit and loss account should be reported separately in the attachment to the closing of books as calculation services that can be assigned to the calculation of prices of services or goods that create the added value (added value adjusted) and the other operational services. These other operational services are usually part of transfer services for international tax optimization. Financial Authority suggests that the tax recognition of cost interests should be at max. 20% of EBITDA (Earnings Before Interest, Taxes, Depreciation and Amortization). For the area of services could be proposed and recommended independent taxation of operational services which cannot be assigned to the calculation. Provided there was a link between the tax recognition of costs only according to material and time assignability into the calculation of the price of services and goods, Financial Authority would improve its effectivity and the tax payers could prepare better for the possible tax control.

References

- Ben-Gal, I. (2005). Outlier detection. In Maimon O. and Rockach L. (Eds.) *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*. Kluwer Academic Publishers. ISBN 0-387-24435-2.
- Jarkovský, J., & Littnerová, S. (2011). *Vícerozměrné statistické metody* [online]. Evropský sociální fond. Evropská unie. Available from: <http://www.iba.muni.cz/esf/res/file/bimat-prednasky/vicerozmerne-statisticke-metody/VSM-01.pdf>
- Meloun, M., & Militký, J. (2004). Přednosti analýzy shluků ve vícerozměrné statistické analýze, In *Sborník přednášek z konference Zajištění kvality analytických výsledků* (pp. 29-46), Medlov, 22.-24. 3. 2004. ISBN 80-86380-22-X. Available from: <https://meloun.upce.cz/docs/publication/152.pdf>
- Meloun, M., Militký, J., & Hill, M. (2005). *Počítačová analýza vícerozměrných dat v příkladech*. Praha.: Academica 449 p. ISBN 80-200-1335-0.
- Voženílek, V., Kaňok, J., & Tuček, P. (2008). Extreme detection, provableness and visualization of statistical data files. *Kartografické listy*, 16 p.
- Vokac (2016). *Kmlinux* [online]. Available from: <https://kmlinux.fjfi.cvut.cz>
- TIBCO (2017). *Statsoft* [online]. Available from: www.statsoft.cz
- Zákon č. 563/1991 Sb. o účetnictví
- Směrnice č.2013/34/EU

