

Accounting Data Quality under the NACE Classification – Case Slovakia

Ladislav Kabát

School of Public Administration, Slovakia

Abstract.

The management of economic entities on corporate level requires both, timely and content-reliable data. These data are largely taken from the final balance sheets of individual businesses. They are defined by the standard structure, the monitored indicators, as well as by content. Notwithstanding their obligatory content, the submitted data are often incomplete, value-unrealistic, or significantly different from analogous indicators of other undertakings with a similar business focus. These facts weaken the usefulness of such data for an in-depth analysis.

Our paper presents some results of the author's research in the field of statistical data collection and processing, including the data quality checking and data modeling.

The role of the UN System of National Accounts (SNA), the World Trade Organization and the World Custom Organization Harmonized System of Trade Statistics (HS) as well as International Financial Reporting Standards are discussed and clarified.

The main goal of the paper is to present the last results of the author's study on financial corporate ratios reported by individual companies under the NACE classification in Slovakia over several last years. The studied data are selected from the final balance sheets submitted by more than 900 thousand of small and medium enterprises over last years. The sectorial indicators (ratios) and their practical applications in checking the accounting data quality are presented as the core results of this study.

Key words: Data quality, creative accounting, IFRS, financial ratios, tax evasion, tax avoidance, economic criminality

JEL: G10, G34, H32

1 Introduction

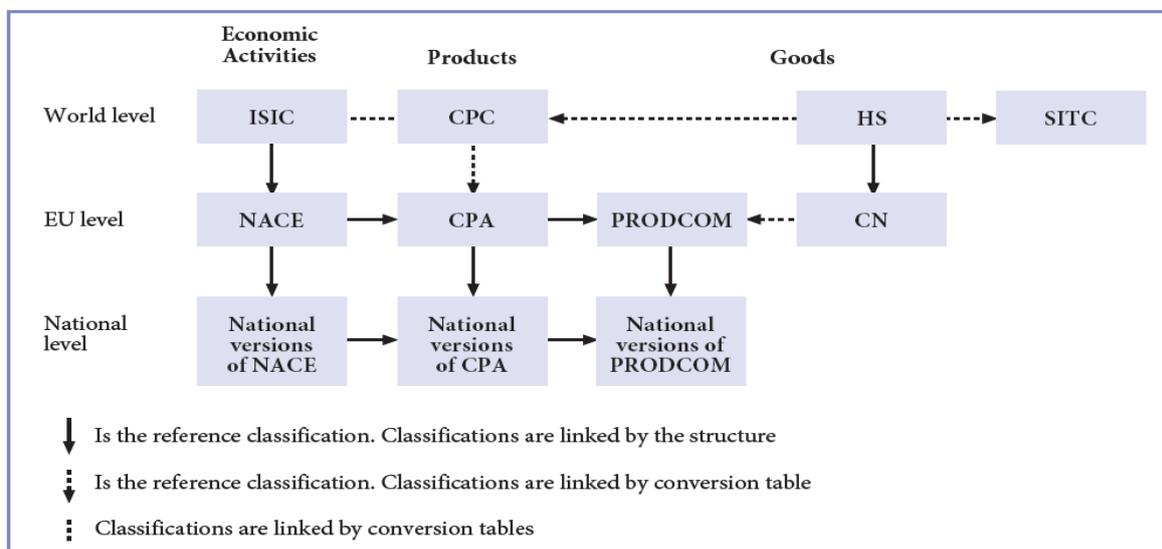
Efficient management in general and, therefore, the management of economic entities on corporate level and ongoing processes requires both, timely and content-reliable data. The data is largely extracted from the final accounts of individual businesses, which are defined by their uniform structure, the scope of the indicators monitored and definition of their content. Notwithstanding the above, data for individual reporting units are often incomplete, value-unrealistic, or significantly different from analogous indicators of other undertakings with a similar business focus.

Problems are even more serious, when national data is compared, or complemented with data generated in foreign economic and/or statistical system. These facts hinder or at least weaken

the usefulness of such data for the more sophisticated analyses, modeling and comparative economic and social studies.

There are various attempts, many projects and active institutions trying to solve these problems and to assist in improving statistical data quality in general. The important role in this effort is played by relevant international institutions.

The attempt to develop efficient, internationally accepted system for statistical data collection, processing and archiving requires enormous energy, financial and material support, as well as full cooperation and support from political bodies in individual countries, on EU level as well as on United Nations bodies' level. The below scheme presents the core levels, components and agents taking part in this endeavor.



Source: www.europeandataportal.eu/

In this structure **economic activities** are classified and reported on world level under the International Standard Industrial Classification (ISIC), on EU level under the standard NACE¹ classification system, while on national levels some versions of NACE are adopted. The NACE is a four-digit classification providing the framework for collecting a large range of statistical data according to economic activity of individual enterprises within the European statistical system (ESS).

Concerning **the products**, the world level adopted the Central Product Classification (CPC), EU level adopted the Classification of Products by Activity (CPA), while on national levels also some versions of the same system are used.

The statistics on the production of manufactured **goods** is provided by PRODCOM system, <https://ec.europa.eu/eurostat/web/prodcom>. The Harmonized Commodity Description and Coding Systems (HS) “allows participating countries to classify traded goods on a common basis for customs purposes”.

¹ The term NACE is derived from the French Nomenclature Statistique des Activités économiques dans la Communauté Européenne).

The Combined Nomenclature (CN) is a tool for classifying goods, set up to meet the requirements both of the Common Customs Tariff (CCT) and of the EU's external trade statistics.

The Standard International Trade Classification (SITC) is maintained by the United Nations and recommended only for analytical purposes. It fulfills special role in following and reporting data on export and import statistics and data for international comparisons of commodities and manufactured goods.

This paper presents some parts of the author's research in the field of statistical data collection, data processing, including data quality checking and data modeling. The roles of the relevant international agents is frequently referred. Namely role of UN System of National Accounts (SNA), World Trade Organization (WTO), World Custom Organization (WCO) Harmonized System (HS), and International Financial Reporting Standards (IFRS) are discussed and clarified.

The main goal of this paper is to present the selected results of the author's study of the sectoral (NACE) selected financial indicators, frequently applied in corporate economics. The studied data are selected from the final balance sheets of more than 900 thousand small and medium enterprises operating in Slovakia. The generalized statistical indicators and their distributions are presented as the core results of the author's study. We believe, they could play an important role in generating, collecting, and processing statistical data. They could serve as an efficient tool in checking and keeping the quality of data on the declared level. This is important not only for the individual firms but also for governments in line with their interest in having the correct tax bases of the individual taxpayers and consequently to maximize tax revenues.

Our paper presents the group of randomly selected **360** entrepreneurial entities which were statistically analyzed according to the standard methodology. Only companies which presented their reliable final balance sheets were considered for our study. The consequently calculated indicators are presented through statistical parameters for the individual sectors, subsectors and various group of companies as well as for their subgroups. The developed generalized sectoral ratio indicators enable the authorities to check the correctness of partial accounting data provided by individual companies.

2 Statistical Data Availability and Their Processing

Availability and access to reliable statistical data is frequently commented in academic journals, as well as in public media. There are many reasons for that. Results of economic and social development, presented by governments and other official authorities, are evaluated and scrutinized and confronted by various nonpartisan institutions for what the objective data is needed. The data for these purposes should be accessible in real time, in requested content and structure in user friendly and easy understandable form. Another reason for demand for reliable data is narrowly linked with the other scientific studies expecting such data as input data for their own studies (Vlaeminck,1).

Any deficiencies in referred statistical data, namely their incompleteness in content or time coverage, can cause various slowdowns in research, or public service projects. To find the ultimate solution to these problems is not an easy task, as expressed by McCullough (2), who

pointed out that “results published in economic journals are accepted at face value and rarely subjected to the independent verification that is the cornerstone of the scientific method. Most results published in economics journals cannot be subjected to verification, even in principle, because authors typically are not required to make their data and code available for verification.”

In case of the corporate economic entities, these facts can weaken the usefulness and acceptability of such data for an in-depth analysis of the group of companies concerned and for the derivation of universally valid relationships among their performance indicators (Ayodotun,3). It also weakens the possibility of mutual comparing the companies whose databases are inconsistent and reduce the more sophisticated applications of accounting data. This can lead to some question on usefulness and applicability of the economic data in general.

The notion of **extreme values** (Haroon, 4) in the case of economic indicators calls for specific assessment criteria based on deeper knowledge of the economic relations and the context of how they are changed over time and how they are collected under which methodology (Zgurovsky,5). All this can help with correct interpretation and understanding of the extreme values.

A frequent problem that complicates the processing of large files of economic data is *the missing data* (Vroomen,6). This problem occurs very often, which disrupts the representativeness of the relevant database. Theoretically simple solution in estimating the missing values by regression analysis is not always proper. Our approach is based on preliminary analysis of value of explanatory variables through which the missing values can be calculated. The imputed values should be marked as “calculated”, which could be significant information for the future users. Only after data treatment and the creation of representative databases of "reliable data" is possible to continue with their exact statistical analysis or model applications.

Discussion on reliable data covers rather broad spectrum of issues. To keep the discussion more concrete, we narrow it on the data, produced by the corporate accounting system. The subsequent applications of such data can be utilized with the assumption that even the so-called *reliable data* of corporate accounting is not statistically exactly defined. The methodology for their collection, in fact, provides rather a broad option for their "creative construction" and consequently their copying into the archived accounting documents. A possible solution is only through the comparison of these extreme values with identical indicators of other enterprises.

3 Reliable Data and Their Statistical Identification

The concept of *reliable data* refers to characteristics like inputs and outputs of the tracked entities with defined content, the way they are collected, updated, archived and processed. An example, the wide range of documents specifying the principles and content of the accounting data, both at corporate and national levels could be presented.

An essential document defining the obligations of intelligence units at the micro level is known as **International Financial Reporting Standards** (IFRS). The IFRS are developed by

an independent body based in London, the International Financial Reporting Standards Foundation and International Accounting Standards Board (**IFRS,7**). Their implementation within the EU is specified by the EU Commission Directive 2013/34/EU. It stipulates that all business operators must process their consolidated financial reports in accordance with the single system of international standards. They enable the subsequent comparisons of the results of businesses as well as an in-depth analysis of their activities. Also, the search for bottlenecks in production or marketing processes is possible. The identification of the "best" entities allows the creation of so-called league tables based on values of objective indicators (**EU,8**)

The proper classification of enterprises with respect to IFRS standards allows the creation of the NACE groups and the quasi-homogeneous subgroups of enterprises at different levels. This justifies accepting them as individual industry standards. The knowledge of these standards can be utilized by reporting firms for checking their own accounting data prior to their savings into a standardized database of consolidated reports. This creates also a possibility for auditors and auditing institutions to adopt such standards for checking accounting data reported by individual entrepreneurial entities.

At the level of macroeconomic reporting in the international environment, as guidance documents are most often reported **the System of National Accounts (SNA,9)** and **Compilation Guide on Inventories (Eurostat and OECD, 10)**. These internationally respected documents arose on the land of the United Nations Statistical Commission in 1947 and 2017. Since their inception, they have undergone several major updates, the last one being held in 2018. The SNA represents a set of standards on how to collect data on the economic results of individual countries and enterprises. Consequently, these results are mutually comparable and consistent with accepted concepts and definitions, classification and accounting rules. The National Accounts system is one of the pillars of macroeconomic statistics and the basis for economic analysis and the development of economic policy scenarios at national and economic

Several specialized international organizations have developed rules on the collection and processing of information on specific products and services provided by their member States, largely incorporated into the United Nations system. As an example, the World Trade Organization (WTO) should be named. WTO is the only international organization dealing with the global rules of trade. "Its main function is to ensure that the international trade flows are realized smoothly, predictably and freely as possible. By lowering trade barriers through negotiations among member governments, the WTO's system also breaks down other barriers between peoples and trading economies", (WTO, www.wto.org). The special function is fulfilled by an information system that contains statistical indicators on time series on merchandise trade, trade in services market access indicators (bound, applied & preferential tariffs), non-tariff information as well as data retrieval functionalities including metadata. (**User Guide, 11**).

The World Customs Organization (**WCO, 12**) cooperates closely with the WTO and manages the most significant, globally implemented tools for organizing trade and customs data. This platform is called **the Harmonized System (HS) of the trade and tariffs**. The Handbook (**HS Classification, 13**) introduces the Harmonized Commodity Description and Coding

System usually referred to as the Harmonized System or the "HS". This system links the goods or downstream technological components with their statistical and custom specifications. It explains the origins of the Harmonized System, gives a detailed description of it and sets out the provisions for its maintenance. It also contains the text of the International Convention on the Harmonized Commodity Description and Coding System, the Rules of Procedure of the Harmonized System Committee and Sub-Committees.

The product and sector classifications reflect a hierarchical structure, displayed in a tree view. For example, in the Harmonized System classification, HS chapters (2-digit) constitute the first level of the hierarchy, under which HS headings (4-digit) and HS subheadings (6-digit) are listed.

The HS system is valid since 1988 and is managed by the World Customs Organization. The main objective of the above-mentioned instruments is to contribute to maintaining the legislative and statistical clarity of the business environment where an international exchange of goods and services is carried and where the GDP is calculated. In case of the international trade, all statistical information provided by both, the reporting country and the partner country of the business relationship, must comply with the formal requirements of the IFRS, SNA and HS systems.

The gross domestic product as a comparative indicator allows us to monitor and compare the dynamics and results of economic growth in different countries of the world, allocate resources of the international institutions to promote economic growth and the social development of needed countries, but also to address the negative impact of natural disasters on countries and areas that need such assistance. The significance of this indicator, its content and methodological reasoning, as well as reservations against it, are subject to analysis of many publications. As the "politically most important and most controversial number in the world" is analyzed by Fioramonti (14, Fioramonti) who highlights several specific situations where the value of the GDP per capita played a significant role in influencing their economic development.

It can be noted that methodological instruments such as IFRS, SNA, and HS contribute to the creation of highly qualified information sources that ensure the reliability of their subsequent applications. However, the creation of databases that are rated as reliable may not be restricted to applications and respect of the IFRS, SNU and HS instruments. From the very nature of their definition, it is evident that these instruments process information at the level of business operators by the country accounting system. Hence, the monitoring and evaluation of the quality of this information and partial data at this level are also in full competence of the relevant authorities and their active participation in supervising these reports is expected. Such activities are expected namely in supervising data quality and quality assurance at the *primary accounting data at the level of business operators*.

Individual business operators in the accounting presentation of their results pursue *two partially contradictory tasks*. The first one is to keep an eye on the progress and results of economic processes inside the enterprise and on its commercial and technological relations with other relevant entities within the business environment. In the context of the legislative definition of business activities, an important part of the internal accounting data processing is the second task - to deduce the summary indicator which, according to the rules of the competent tax authority, will represent the tax base for this business entity.

4 Base Erosion and Profit Shifting Initiative

Historical experience shows that the two information responsibilities are frequently not fully consistent. While on the side of the government, the uncompromising obligation to respect the legislative definition of the declaring tax persists, on the part of the taxpayers the efforts are still being made to minimize these obligations. Especially the academic sphere, the long term comes with solutions to optimize the tax obligations, which is rather a euphemism for their minimization. In detecting the maneuvering space for adjusting the tax obligations, the legislative gaps in the expression of the content of certain partial constituents that are involved in the design of the economic performance summary indicator are mostly applied. The summary indicator is referred to as the tax base. It is therefore interesting to analyze the factors that are proactive in supporting the fulfillment of the tax obligations (15, ALM, 2018)

Even before we go to the next debate it is necessary to clarify the two basic approaches that appear in finding the area for tax base adjustment. The first approach is **tax avoidance**, which uses legal possibilities to manipulate the tax base by its deferral or decomposition the tax obligations. A different approach is so-called **tax evasion**, which means the conscious not paying taxes, declaration of not full incomes or not ineligible costs and request for their recovery by the tax administrator. The literature presents various methodologies for estimating frequencies of these failures (16, Rizzi, 2017). The different results (17, Hines, 2014) are also reported on their sources and societal seriousness.

To activate the taxpayers to pay their duties, in some cases, the instruments of positive recognition of tax confidence (18, Bazart, 2011) are also used.

The lack of tools to control the **base erosion and profit shifting** (BEPS) leads toward a reduction of the tax base and the transfer of profits to the tax area with low or no tax rates, where no economic activity is carried out. This causes the long-term undermining of tax systems of countries and reduces their efficiency in tax collection. Manipulative shifting of profits into a "more appropriate" tax environment causes problems in countries where real economic activity takes place, where natural and human resources are consumed, while the profits are exported into foreign countries. According to older surveys (OECD, 2013), such operations are hiding 100-240 billion USD per year.

The situation with the reduction of the tax base is reported by the OECD/G20 report (Inclusive, 19). According to this document, the meeting of the top representatives of the countries signing the BEPS declaration shows that its action is being continuously met and the Member States support the fulfilment of the objectives originally defined as measures of program fighting against reducing tax bases.

The theoretical aspects of tax evasion and tax base reduction are examined also in an international context (20, Matsuoka, 2017). 21, Olaoye, 2017) and can act as an impetus for its more rigorous solution in economically developed countries.

The adoption of the action plan to combat the tax base reduction and profit shifting by the G20 leaders in St. Petersburg in 2013 showed high support for efforts to adopt the international tax system, which would reflect the new challenges for the tax system in a

globalized economic environment. Currently, there are 116 countries that adopted fully the BEPS policy in fighting the tax base erosion and profit shifting.

The OECD and EU initiatives and follow-ups have resulted in several concrete actions. Their content and the concretization of future tasks are formulated by the document (OECD, **22**). The document contains 15 detailed measures that should be incorporated into the national tax legislation of all countries that have expressed support for this initiative and have taken part in the development of its legislative framework. Of the fifteen recommended measures, the following are the most important:

1. Resolution of methodological mismatch in the formulation of tax obligations (hybrid mismatch arrangements)
2. Tackling the not efficient tax rules for foreign companies and shifting their profits into tax "haves"
3. Double Taxation Solution
4. Taxation when changing a tax place
5. Abuse of artificially generated debt of the company.

This set of measures is complemented by the agreement of the countries concerned with a timely mutual exchange of information on fraudulent tax activities and attempts by tax operators to abuse their tax systems.

5 Anti-Tax Avoidance Project of the EU and Its ATAD Directive and Its Application in Slovakia

The basic objective of the initiative in question was to create a space for the application of concrete measures in the formulation of the tax principles and their consistent putting into practice at an international level. The European Commission also joined the initiative. A proposal for a directive on the fight against tax evasion (Anti-Tax Avoidance Directive - ATAD) and subsequently endorsed by the European Parliament, (EU) 2016/1164.

Despite the validity of that directive, problems with manipulative processing in data on the calculation of the tax base persist. The last information on progress, and problems with adopting this initiative is presented in OECD progress report inclusive framework on BEPS (OECD, **23**). According this report, the membership of the initiative has increased and by this year achieved 135 countries, which agreed on implementing the recommended measures into their national accounting and tax legislation. Majority of them will come into force by next two years.

In December 2017, the principles of ATAD, elaborated on the terms of Slovakia, were endorsed by the National Council of the Slovak Republic as Law no 344/2017 ECR, in the form of an amendment to Act no 595/2003 on income tax. The objective of this national legislative standard was also to formulate measures to increase protection against aggressive tax policies aimed at reducing the tax base and shifting profits outside the SR. The implementation of the ATAD directive in the SR tax system is one of the tasks focusing on the derivation of sectoral standards of selected ratios (8, Zalai), which is also addressed by the SAF and CRIF.

The fulfillment of the tasks of the ATAD project requires not only the legislative alignment of existing tax rules with the new requirements of the directive but also to ensure the quality and reliability of the primary data used in the calculation of the value Tax obligations of a taxpayer. It is, therefore, necessary to ensure that, in the primary processing and accounting data circuit, it is also considered sectoral, in line with the structure of the NACE, defined by the standards of the main economic and financial indicators in the field of the corporate economy. By providing well-qualified, industry-class standards to employees of the enterprise who are responsible for collecting and processing primary data, it could be limited to bringing economically incorrect data into the corporate accounting system the information noise in archived data sets.

Based on the experience and results so far, it can be noted that the development of sectoral standards is not numerically difficult, but its implementation has several practical pitfalls. At least the following of them should be mentioned. As the experience shows, many accounting variables are represented by the zero values, which complicates the calculation of the following characteristics expressed as the ratio criterion. The examples are, for example, indicators ROA, ROE, EBITDA, and some others. This founding is confirmed by the long-term experience with the “Average Values” project (24, CRIF).

The results achieved also confirm that the outputs obtained are valuable information to the employees of business practice, as well as students, researchers, educational and managing authorities for comparative analyses of economic and financial the results of the undertakings operating in the same sectoral and legislative group of the NACE.

It can be confirmed that results of the “Average Values” project which so far have been achieved offer the valuable basis for the further developing the sectoral standards, and for developing the “user friendly” presentation to the less experienced users.

6 Data and Methodology

The accounting data, in accordance with their definition, should be a *"non-distorted information mirror"* of economic, financial, production and distribution processes. Its actors are suppliers, manufacturers, distributors, dealers, financial institutions, as well as public authorities. Their mutual interaction generates information that is methodically processed, summarized and presented publicly under the established rules in a broad-based system of accounting documents.

In checking the primary data, it is necessary to analyze and to treat data that differ significantly from the same indicators of undertakings with similar size, sector classification, legislative structure, and business history.

The higher attention deserves the data that are statistically classified as the *"outliers"*. The simple elimination of them according to strict statistical rules is not always the only and rational decision. In some case the identification of “outliers” can be an incentive for further and deeper economic analysis or for efficient checking of the accounting data generating procedures.

Subsequently, after eliminating the extreme values in data files, it is possible to construct the statistically quasi-homogeneous sets of enterprises according their legislative form, size,

NACE group and other relevant attributes and to derive the ratio indicators for individual variables and the specified groups of enterprises. The financial ratios are represented by a set of statistics as **mean, median, modus, standard deviation, and quadrant values, skewness and kurtosis**. As a result, individual businesses can be sorted into performance groups, possibly estimating their distances from the best performing company.

The main goal of the above presented data analysis is to confirm the correctness of the reported data and their qualification for further economic and statistical applications. Namely, we strive for confirmation of their correctness and applicability in deriving the tax base of the individual taxpayers, as it is requested by IFRS documents.

To fulfil the above indicated goals, we continue with calculations of the selected financial ratio indicators (Zalai, 25) as indicated below:

- | | | |
|----|---------|-----------------------------|
| 1. | ROA | - Return on Assets |
| 2. | ROE | - Return on Equity |
| 3. | EBITDAS | - EBITDA in Sales |
| 4. | VAS | - Value Added in Sales |
| 5. | NVAS | - New Value Added in Sales |
| 6. | AT | - Assets Turnover |
| 7. | TVI | - Total Volume of Inventory |
| 8. | ITT | - Inventory Turnover Time |
| 9. | MD | - Maturity of Debts |

In line with the accepted rules, the above ratios are calculated only for enterprises for which:

$$\text{ROA and ROE and VAS and NVAS} > 0 \quad (1)$$

For all financial ratios the following statistics are calculated:

- Descriptive statistics – min, max, mean, median, standard deviation, kurtosis, skew
- Quartile values and number of enterprises in quartile
- Number of enterprises with specific values of ratios (<0, =0)

All enterprises fulfilling condition (1) create the database of the standard entities, which could be used for further studies and possibly for deriving the probability distributions of values of sectorial ratios. In further processing the financial ratios will be analyze for enterprises sorted according

1. The NACE structures
2. Size of enterprises according volume of Assets (small, medium, large)
3. Size of enterprises according annual volume of Sales (small, medium, large)

The ratio indicators provide the useful information on general economic performance of companies. However, the ratio indicators are derived from the other administratively calculated values and lack the clear direct reflection of the core input and output factors on final economic effect of individual companies. Having in mind this, we perform also deeper analysis of assets, sales and inventories, as key production factors. Initial data for our

analysis is extracted from the final balance sheets for 2018 year. The total number of companies under our study was 68 371

7 Results and Discussion

For creating the homogenous groups of enterprises we adopted the EU scheme for splitting these entities into three groups according value of their assets and annual turnover on sales in following, so the assets are considered in categories less than 350 thousand euros, 350 000 to 4 million and over 4 million. The annual turnover is considered as less than 700 thousand euro, 700 000 to 8 million and over 8 million. Distribution of our subgroup of 68 371 companies shows **Table 4**.

Table 4 Distribution of enterprises in absolute values

Distribution of enterprises according Assets and annual Turnover on sales				
	Turnover of sales (1000 euros)			
Assets (1000 euros)	O1(< 700)	O2 (700, 8 000)	O3 (>8 000)	TOTAL
A1 (< 350)	50 159	1 559	323	52 041
A2 (350, 4 000)	5 704	7 342	323	13 369
A3 (> 4000)	339	1 172	1 450	2 961
	56 202	10 073	2 096	68 371

For more detailed study all these companies were divided according the volume of Assets and volume of annual Turnover of sales. The structure of the whole set of companies with their distribution into referred subgroups is shown in **Tables 4**, in absolute values. Table 5 presents these data in relative values.

Table 5 Distribution of enterprises in relative values

Distribution of enterprises according Assets and annual Turnover on sales, %				
	Turnover of sales (1000 euros)			
Assets (1000 euros)	O1(< 700)	O2 (700, 8 000)	O3 (>8 000)	TOTAL
A1 (< 350)	73,4	2,3	0,5	76,1
A2 (350, 4 000)	8,3	10,7	0,5	19,6
A3 (> 4000)	0,5	1,7	2,1	4,3
	82,2	14,7	3,1	100

The data shows that more than 95 % of all companies are classified in category of small and medium size companies.

7 Results and Discussion

The proposed methodology in calculation the financial ratios was verified on the subgroup of 360 randomly selected enterprises of the NACE 01110. Among them are only small and medium size companies, as shown in Table 6 where small companies represent 54,7 % and medium size companies over 45 % of companies.

Table 6 Distribution of enterprises of NACE 01110

Distribution of enterprises according Assets and annual Turnover on sales				
	Turnover on sales (1000 euros)			
Assets	O1	O2	O3	TOTAL
A1	100	0	0	100
A2	94	104	0	198
A3	3	59	0	62
	197	163	0	360

Distribution of enterprises according Assets and annual Turnover on sales, %				
	Turnover on sales (1000 euros)			
Assets	O1	O2	O3	TOTAL
A1	27,8	0,0	0,0	27,8
A2	26,1	28,9	0,0	55,0
A3	0,8	16,4	0,0	17,2
	54,7	45,3	0,0	100,0

In following parts we present the detailed descriptive statistics for selected financial ratios as defined under the methodology section. In the first step we present the values of financial ratios for all 360 companies, **Table 7**.

In **the first section** of this table there are **the descriptive statistics for financial ratios**. These results could be used for the further in-depth data analysis. Specifically, they could be used for comparative analysis with the accounting data of the companies of the same NACE and size group. Such comparison could help to avoid the uncorrected data from the balance sheets documents.

The second section of the table presents **the quartile structure of the data**, which enables to classify the companies into relatively homogenous groups of companies according the values of their financial ratios. The smallest and largest values of the individual ratios could be identified and checked.

The last part of the table presents **information on statistically and economically incorrect or dubious data** or data classified as outliers. Mainly, they identify the data missing for calculation the individual ratios. Such finding could help in checking quality of accounting data before they are transferred and saved as official data. These findings gave impetus for deeper study of the respective original accounting data.

The data presented in **Table 7** and all following tables could be used for comparison an individual company with its economic, technological and/or marketing peers.

Table 7 Overview of the selected financial ratios for 360 companies of the NACE 01110

Overview of selected financial ratios for 360 companies of NACE 01110									
Selected ratios									
	ROA	ROE	EBITDA/S	NVHvS	NVAinS	AT	TZ	IT	MD
Measures of position									
MINIMI	0,04	0,01	0,32	-168,58	-157,22	0,81	0,03	0,00	0,00
MAXIM	74,49	1286,12	484,02	99,26	118,43	99,72	1522,28	1200,00	922,21
Mean	8,48	21,78	33,42	-15,18	15,59	55,66	23,00	115,29	69,15
STDEV	12,16	71,89	35,90	30,82	30,38	27,91	136,01	122,11	106,02
MEDIA	3,96	10,16	26,41	-12,30	18,87	57,50	4,13	91,39	37,28
KURTOS	10,06	268,69	75,02	6,09	6,57	-1,05	90,46	22,92	17,60
SKEWN	2,90	15,45	6,98	-1,43	-1,34	-0,31	9,39	3,46	3,56
Quartile values									
0-Minim	0,04	0,01	0,32	-168,58	-157,22	0,81	0,03	0,00	0,00
1	1,40	3,24	16,01	-23,24	5,07	32,25	1,94	38,77	6,42
2-Media	3,96	10,16	26,41	-12,30	18,87	57,50	4,13	91,39	37,28
3	10,87	24,16	40,09	-0,35	30,20	79,66	8,83	158,47	80,60
4-Maxim	74,49	1286,12	484,02	99,26	118,43	99,72	1522,28	1200,00	922,21
Number of companies within quartile									
1	90	90	90	90	90	90	90	90	90
2	90	90	90	90	90	90	90	90	90
3	90	90	90	90	90	90	90	90	90
4	90	90	90	90	90	90	90	90	90
Sum	360	360	360	360	360	360	360	360	360
Share of companies within quartile in %									
1	25,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0
2	25,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0
3	25,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0
4	25,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0
Sum	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
The number of companies with specific values									
< 0	0	0	0	273	72	0	0	0	0
. = 0	0	0	0	0	1	0	0	46	48
Mean+-	322	353	337	298	300	213	354	324	330
Mean+-	343	358	353	335	338	360	355	349	341
Mean+-	350	358	356	351	353	360	356	355	350
Outliers	0	0	0	8	6	0	0	0	0
Outliers	10	2	4	1	1	0	4	5	10
The ratios with specific values in %									
< 0	0,00	0,00	0,00	75,83	20,00	0,00	0,00	0,00	0,00
. = 0	0,00	0,00	0,00	0,00	0,28	0,00	0,00	12,78	13,33
Mean+-	89,44	98,06	93,61	82,78	83,33	59,17	98,33	90,00	91,67
Mean+-	95,28	99,44	98,06	93,06	93,89	100,00	98,61	96,94	94,72
Mean+-	97,22	99,44	98,89	97,50	98,06	100,00	98,89	98,61	97,22
Outliers	0,00	0,00	0,00	2,22	1,67	0,00	0,00	0,00	0,00
Outliers	2,78	0,56	1,11	0,28	0,28	0,00	1,11	1,39	2,78

Source: Author

The following **tables 8, 9 and 10** presents the values of financial ratios for three subgroups of companies, namely for subgroup A1-O1 with 100 companies, subgroup A2-O1 with 94 companies and subgroup A2-O2 with 104 companies. We do not calculate the financial ratios for subgroups with very small number of companies.

Table 8 Overview of the selected financial ratios for subgroup of 100 companies

Overview of selected financial ratios for group of 100 companies of NACE - 01110 - a1 - o1									
Selected ratios									
	ROA	ROE	EBITDA/S	NVHvS	NVAinS	AT	TZ	IT	MD
Measures of position									
MINIMI	0,070	0,160	0,320	-157,220	-157,220	0,810	0,046	0,000	0,000
MAXIM	74,490	161,370	484,020	67,160	118,430	98,420	1522,280	1200,000	922,212
Mean	13,625	25,842	40,964	-19,895	3,512	54,312	27,441	81,673	66,552
STDEV	17,248	27,928	55,915	38,799	41,576	29,720	155,988	149,475	137,504
MEDIA	5,915	16,865	29,595	-12,350	9,405	56,375	3,079	42,120	4,878
KURTOS	3,428	5,797	42,273	2,272	3,310	-1,180	87,558	31,644	16,589
SKEWN	1,948	2,120	5,851	-1,166	-0,899	-0,332	9,147	4,772	3,626
Quartile	Quartile values								
0-Minin	0,070	0,160	0,320	-157,220	-157,220	0,810	0,046	0,000	0,000
1	2,310	7,103	16,700	-35,028	-13,218	27,793	1,325	0,000	0,000
2-Media	5,915	16,865	29,595	-12,350	9,405	56,375	3,079	42,120	4,878
3	18,065	32,078	46,370	1,595	26,253	79,443	6,160	98,424	74,349
4-Maxim	74,490	161,370	484,020	67,160	118,430	98,420	1522,280	1200,000	922,212
Quartile	Number of companies within quartile								
1	25	25	25	25	25	25	25	34	37
2	25	25	25	25	25	25	25	16	13
3	25	25	25	25	25	25	25	25	25
4	25	25	25	25	25	25	25	25	25
Sum	100	100	100	100	100	100	100	100	100
Quartile	Share of companies within quartile								
1	25,00	25,00	25,00	25,00	25,00	25,00	25,00	34,00	37,00
2	25,00	25,00	25,00	25,00	25,00	25,00	25,00	16,00	13,00
3	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00
4	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00
Sum	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Cluster	Number of companies with specific values								
< 0	0	0	0	73	39	0	0	0	0
. = 0	0	0	0	0	0	0	0	34	37
Mean+-	87	87	96	76	82	62	97	91	90
Mean+-	93	94	98	94	93	100	99	97	95
Mean+-	98	98	98	99	98	100	99	99	97
Outliers	0	0	0	1	2	0	0	0	0
Outliers	2	2	2	0	0	0	1	1	3
Klaster	Share of companies with specific values								
< 0	0,00	0,00	0,00	73,00	39,00	0,00	0,00	0,00	0,00
. = 0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	34,00	37,00
Mean+-	87,00	87,00	96,00	76,00	82,00	62,00	97,00	91,00	90,00
Mean+-	93,00	94,00	98,00	94,00	93,00	100,00	99,00	97,00	95,00
Mean+-	98,00	98,00	98,00	99,00	98,00	100,00	99,00	99,00	97,00
Outliers	0,00	0,00	0,00	1,00	2,00	0,00	0,00	0,00	0,00
Outliers	2,00	2,00	2,00	0,00	0,00	0,00	1,00	1,00	3,00

Source: Author

Table 9 Overview of the selected financial ratios for subgroup of 94 companies

Overview of the selected financial ratios for group of 94 companies of NACE - 01110 - a2 - o1									
Selected ratios									
	ROA	ROE	EBITDA/S	NVHvS	NVAinS	AT	TZ	IT	MD
Measures of position									
MINIMI	0,04	0,07	1,12	-168,58	-114,89	1,11	0,07	0,00	0,00
MAXIMI	68,64	1286,12	211,03	99,26	99,26	99,72	1420,84	508,72	673,56
Mean	6,89	32,76	39,25	-16,62	17,95	58,92	34,81	124,06	82,32
STDEV	9,62	135,21	32,41	38,57	32,82	27,94	182,32	114,97	121,70
MEDIA	3,40	10,26	31,54	-13,15	19,98	63,72	4,90	86,20	30,87
KURTOS	17,99	81,75	8,60	4,54	2,98	-0,74	47,72	1,81	7,62
SKEWN	3,48	8,83	2,44	-0,92	-0,54	-0,50	6,91	1,39	2,58
Quartile Quartile values									
0-Minim	0,04	0,07	1,12	-168,58	-114,89	1,11	0,07	0,00	0,00
1	1,40	3,82	19,50	-30,87	1,98	37,48	2,35	40,52	7,61
2-Media	3,40	10,26	31,54	-13,15	19,98	63,72	4,90	86,20	30,87
3	8,42	21,55	48,01	0,23	29,63	80,23	12,35	172,67	99,77
4-Maxim	68,64	1286,12	211,03	99,26	99,26	99,72	1420,84	508,72	673,56
Quartile Number of companies within quartile									
1	24	24	24	24	24	24	24	24	24
2	23	23	23	23	23	23	23	23	23
3	23	23	23	23	23	23	23	23	23
4	24	24	24	24	24	24	24	24	24
Sum	94	94	94	94	94	94	94	94	94
Quartile Share of companies within quartile									
1	25,53	25,53	25,53	25,53	25,53	25,53	25,53	25,53	25,53
2	24,47	24,47	24,47	24,47	24,47	24,47	24,47	24,47	24,47
3	24,47	24,47	24,47	24,47	24,47	24,47	24,47	24,47	24,47
4	25,53	25,53	25,53	25,53	25,53	25,53	25,53	25,53	25,53
Sum	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Cluster Number of companies with specific values									
< 0	0	0	0	70	22	0	0	0	0
. = 0	0	0	0	0	0	0	0	9	9
Mean+-	80	92	80	77	74	58	92	68	84
Mean+-	92	93	90	87	86	90	92	89	89
Mean+-	92	93	92	91	93	94	92	91	92
Outliers	0	0	0	2	1	0	0	0	0
Outliers	2	1	2	1	0	0	2	3	2
Klaster Share of companies with specific values									
< 0	0,00	0,00	0,00	74,47	23,40	0,00	0,00	0,00	0,00
. = 0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	9,57	9,57
Mean+-	85,11	97,87	85,11	81,91	78,72	61,70	97,87	72,34	89,36
Mean+-	97,87	98,94	95,74	92,55	91,49	95,74	97,87	94,68	94,68
Mean+-	97,87	98,94	97,87	96,81	98,94	100,00	97,87	96,81	97,87
Outliers	0,00	0,00	0,00	2,13	1,06	0,00	0,00	0,00	0,00
Outliers	2,13	1,06	2,13	1,06	0,00	0,00	2,13	3,19	2,13

Source: Author

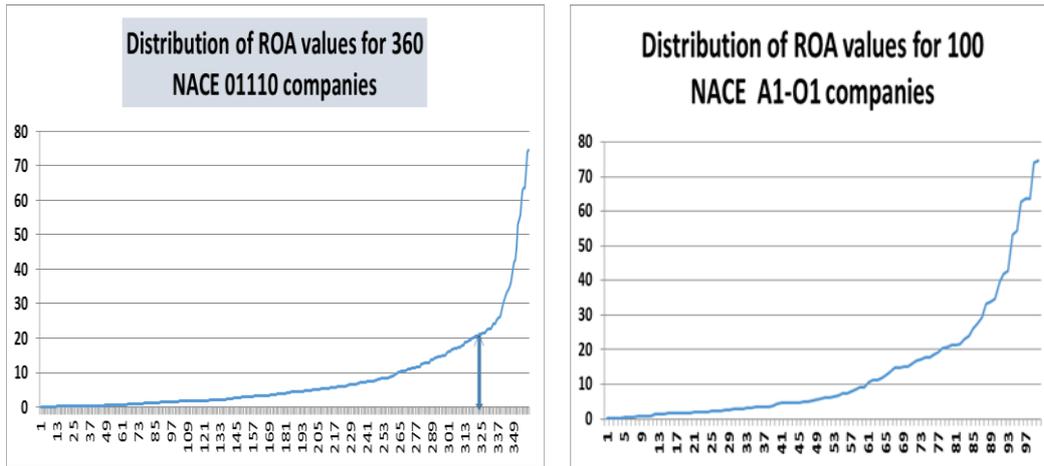
Table 9 Overview of the selected financial ratios for subgroup of 94 companies

Overview of the selected financial ratios for group of 104 companies of NACE - 01110 - a2 - o2									
Selected ratios									
	ROA	ROE	EBITDA/S	NVHvS	NVAinS	AT	TZ	IT	MD
Measures of position									
MINIMUM	0,07	0,01	3,15	-95,74	-51,20	1,29	0,03	0,00	0,00
MAXIMUM	55,68	113,38	96,12	26,59	47,88	99,16	39,96	330,80	432,65
Mean	7,53	16,57	22,92	-11,46	18,96	55,35	6,50	117,76	54,44
STDEV	9,45	20,72	13,69	15,46	14,80	24,39	7,11	73,53	57,88
MEDIA	5,02	10,19	19,19	-11,49	19,55	56,18	4,18	106,34	43,11
KURTOSIS	8,40	7,75	7,10	7,89	4,34	-0,93	9,99	-0,21	17,72
SKEWNESS	2,55	2,58	1,96	-1,54	-1,32	-0,15	2,89	0,52	3,35
Quartile	Quartile values								
0-Minimum	0,07	0,01	3,15	-95,74	-51,20	1,29	0,03	0,00	0,00
1	1,07	3,89	14,16	-17,99	12,12	35,21	2,59	59,66	20,15
2-Median	5,02	10,19	19,19	-11,49	19,55	56,18	4,18	106,34	43,11
3	9,06	22,40	28,97	-2,50	27,91	76,40	8,13	164,31	67,01
4-Maximum	55,68	113,38	96,12	26,59	47,88	99,16	39,96	330,80	432,65
Quartile	Number of companies within quartile								
1	26	26	26	26	26	26	26	26	26
2	26	26	26	26	26	26	26	26	26
3	26	26	26	26	26	26	26	26	26
4	26	26	26	26	26	26	26	26	26
Sum	104	104	104	104	104	104	104	104	104
Quartile	Share of companies within quartile								
1	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00
2	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00
3	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00
4	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00	25,00
Sum	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Cluster	Number of companies with specific values								
< 0	0	0	0	84	9	0	0	0	0
. = 0	0	0	0	0	0	0	0	3	2
Mean+-	91	95	79	80	77	61	95	70	95
Mean+-	100	97	101	100	100	103	100	100	100
Mean+-	102	101	103	103	103	104	101	104	102
Outliers	0	0	0	1	1	0	0	0	0
Outliers	2	3	1	0	0	0	3	0	2
Klaster	Share of companies with specific values								
< 0	0,00	0,00	0,00	80,77	8,65	0,00	0,00	0,00	0,00
. = 0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2,88	1,92
Mean+-	87,50	91,35	75,96	76,92	74,04	58,65	91,35	67,31	91,35
Mean+-	96,15	93,27	97,12	96,15	96,15	99,04	96,15	96,15	96,15
Mean+-	98,08	97,12	99,04	99,04	99,04	100,00	97,12	100,00	98,08
Outliers	0,00	0,00	0,00	0,96	0,96	0,00	0,00	0,00	0,00
Outliers	1,92	2,88	0,96	0,00	0,00	0,00	2,88	0,00	1,92

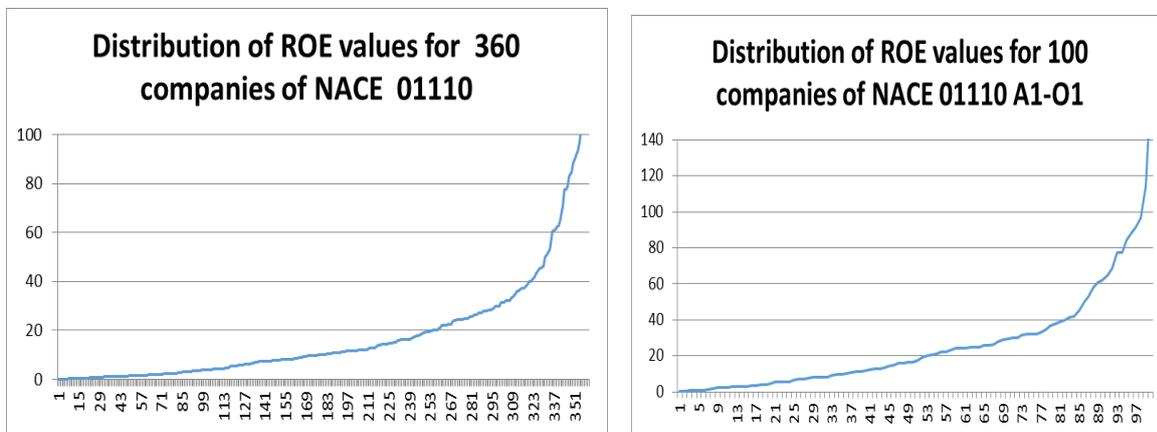
Source: Author

The obtained information on values of financial ratios could be graphically presented and interpreted. Fig. 1 presents distribution of ROA values for 360 companies of the NACE 01110. According this graph, roughly 325 companies (cca 90% of all presented companies) report ROA value less than 20. This finding could be used in further comparative analyses.

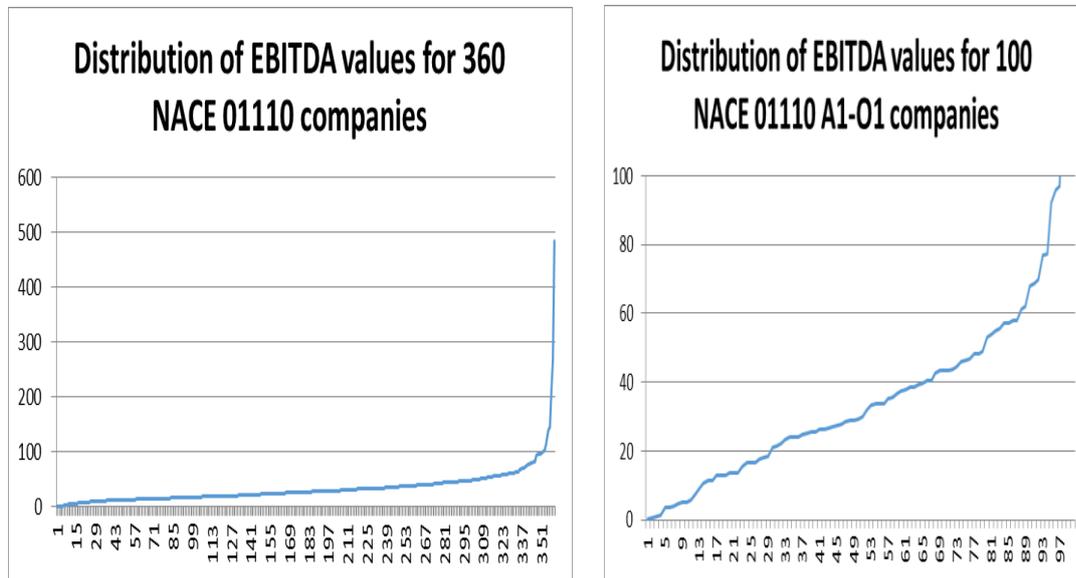
**Fig. 1 Distribution of ROA values for 360 NACE 01110 companies
and 100 NACE A1-O1 companies**



**Fig. 2 Distribution of ROE values for 360 companies of NACE 01110
and 100 NACE 01110 A1-O1**



**Fig. 3 Distribution of EBITDA values for 360 NACE 01110 companies
and 100 companies of NACE 01110 A1-O1**



Conclusion

The aim of this paper was to present the current theoretical views and opinions on the quality of statistical data, their collection, checking their correctness, processing and archiving into publicly accessible data sets. Particular attention is devoted to the primary data generated by entrepreneurial entities classified under various NACE groups. Data analysis is based on study of the final balance sheets of more than 900 thousand entrepreneurial entities. Based on our findings, the quality of this data is analyzed and numerical tools are offered for checking their quality and consistency with the relevant available data of the other entrepreneurial units belonging to the same NACE subgroupings.

The purpose of such approach is to present the outputs which are reliable in presenting economic situation of the studied companies and enabling their mutual comparison. This could be useful both, for individual company as well as for all its potential technological and marketing partners.

To underline significance of the data collection processes, the core methodological documents issued by United Nations agencies as well as by some international organizations are referred.

Statistical data quality is presented not only as the microeconomic problem, related to the mutual relation among the various members of the market environment. The data quality is discussed also as a requirement of the OECD/G20 initiative, expressed in the Base Erosion and Profit Shifting document (BEPS) based on the International Finance Reporting Standards (IFRS). The paper points the harmfulness of violations of these requirements by disrupting the stability and efficiency of the tax and derivative system. The real threat to this situation in Slovakia is documented by statistical data on the development of economic crime and particularly crime in tax cuts.

The developed sectorial indicators could be used as an efficient tool in checking the initial accounting data on correctness, before they are adopted for modelling and calculations of various important economic indicators as GDP and/or tax base.

Literature

¹ Vlaeminck, S.: Research Data Management in Economic Journals, ZBW – German National Library of Economics / Leibniz Information Centre for Economics,

<http://openeconomics.net/resources/data-policies-of-economic-journals/>

² McCullough/McGeary/Harrison: “Lessons from the JMCB Archive”, 2006

³ Oby Ayodotun, J. at all: Contribution of small and medium enterprises to economic development: Evidence from a transiting economy, Data in Brief, Volume 18, June 2018, Pages 835-839 <https://doi.org/10.1016/j.dib.2018.03.126>

⁴ Haroon, M, Barakat, El-Sayed Nigm, Osama M. Khaled: Statistical Techniques for Modelling Extreme Value Data and Related Applications, Cambridge Scholars Publishing, 2019, ISBN (13) 978-1-5275-3207-6

⁵ Zgurovsky, M.: Methodology of the Extreme Value Theory Application for Data Analysis, Research Bulletin of the National Technical University of Ukraine, 2016

⁶ Vroomen MacNeis, J. at all: Multiple imputation strategies for zero-inflated cost data in economic evaluations: which method works best? Eur J Health Econ. 2016; 17(8): 939–950.

⁷ IFRS: IFRS_ www.ifrs.org

⁸ European Commission: Financial Reporting, https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/financial-reporting_en, 2013

⁹ UNDESA: Handbook on Supply, Use and Input-Output Table with the Extension and Applications, 2018

¹⁰ UN: Compilation Guide on Inventories (Eurostat and OECD), 2017

¹¹ WTO: User Guide, www.wto.org

¹² WCO: WCO in Brief, <http://www.wcoomd.org/en.aspx>

¹³ WCO: The HS Classification Handbook, ISBN 9782874921360, 2013

¹⁴ Fioramonti, L.: Gross Domestic Problem: The Politics Behind the World's Most Powerful Number, Amazon, 2017

¹⁵ ALM, J.: WHAT MOTIVATES TAX COMPLIANCE? Journal of Economic Surveys, Vol 33, 2018

¹⁶ Rizzi, D.: Tax Evasion Indices and Profiles, Public Finance Review, Vol. 45, 2017

¹⁷ Hines, R., J.: How Serious Is the Problem of Base Erosion and Profit Shifting? University of Michigan Law School, <https://repository.law.umich.edu/articles/1379>

¹⁸ Bazart, C., Pickhardt, Michael: Fighting Income Tax Evasion with Positive Rewards, Public Finance Review, Vol. 39, 2011

¹⁹ OECD/G20: Inclusive Framework on BEPS: Progress Report July 2017-June 2018

²⁰ Matsuoka, A.: What made base erosion and profit Shifting project possible?: Identifying factors for building momentum for reform of international taxation, Journal of Financial Crime, Vol. 25, 2017

²¹<http://www.apexjournal.org>, ISSN 2315-8727

²² OECD: OECD Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations 2017, OECD Publishing, ISBN 978-92-64-26512

²³ OECD/G20: OECD/G20 Inclusive Framework on BEPS, Progress Report July 2019-2020,
<https://www.oecd.org/tax/beps/oecd-g20-inclusive-framework-on-beps-progress-report-july-2019-july-2020.htm>

²⁴ CRIF: Mean values, CRIF-Slovak Credit Bureau, ISBN 978-80-971109-4-9, 2016

25 Zalai, K.: Finančno-ekonomická analýza podniku, SPRINT 9. vyd., 9788089710225