

THE COMPARISON OF RETAIL UNITS ON THE CZECH REPUBLIC'S MARKET – CLUSTER ANALYSIS AND TIME SERIES ANALYSIS

LUCIE ZOTYKOVÁ

Abstract: This article is focused on the application of cluster analysis and time series analysis in the case of segregation and comparison of large-scale retail units in Czech Republic. The important thing is to find the relationships among retail units of the segment. This particular segment is composed of heterogeneous companies. First it is necessary to classify these retail units into smaller homogeneous clusters according to their similarity. The crucial element here is the comparison of individual clusters (composed of large-scale retail units) with the whole segment in Czech Republic. Thereby we establish which company has a decisive influence on this segment. We will also determine trends and predictions of EBIT, ROA and added value for all clusters.

Keywords: supermarkets, hypermarkets, discount stores, cluster analysis, dendrogram, time series analysis, EBIT, ROA, added value

JEL Classification: M11, C22, C32

1. INTRODUCTION

This paper deals with the application of cluster analysis and time series analysis by processing data on the financial condition of Czech retail stores. Large-scale retail stores in the Czech Republic are presented and compared here. First, we introduced the data source, then we described the methodology used, and finally we presented our results. The main aim of this article is to test the hypothesis of whether the major impact of the economic crisis of 2008 is evident in the performance of time series analysis of individual clusters for the financial ratio indicators: EBIT, ROA and Added Value.

2. LARGE-SCALE RETAIL STORES IN THE CZECH REPUBLIC

A supermarket is a self-service store, which is offering mainly foodstuffs with complementary non-food goods. The primary focus is based on fresh food. The sales area of supermarkets ranges from 400 to 2500 m².

A hypermarket is a large-scale self-service retail store with a wide range of food and non-food products under one roof. Its sales area is more than 2500 m².

A discount store is a variant of retail store which is trying to compete with traditional stores specifically on the lowest price. The self-service food is prevailing, non-food goods is represented to a lesser extent.

3. DATA

The source data were taken from the official site of the Czech judiciary, the web site: justice.cz. This server publishes the data entered in the business register, including the data of large-scale retail units. Database for my research project contains information such as: costs of sales, consumption of production, personal expenses, total assets, equity, registered capital, turnover, production, revenues products services, added value and profit; to each of the retails for each of the year in the period 2008-2011.

This paper compares supermarkets between the years 2008-2011 due to the availability of comparable data.

4. FINANCIAL RATIO INDICATORS

This paper also contains financial ratios, which are derived from database of research for my doctoral thesis: The Apply Statistics Methods With Marketing Research in Supermarket's Customer Satisfaction. The selected ratios were: EBIT, ROA and ADDED VALUE.

EBIT=Earning before Interest and Taxes. It is used in financial analysis to financial ratios. It assesses business performance regardless of the chosen method of financing and taxation. Investors prefer this indicator for its complexity. (Needles, Powers, Crosson, 2007). In Czech accountant document about Profit and loss we find equivalent of EBIT on the 30th row.

ROA=Return on Assets. This indicator shows how profitable are company's assets in generating revenue. It is a useful for comparing competing companies in the same industry. This number mean how profitable a company is before gearing. (Knápková, Pavelková, Šteker, 2012)

$$ROA = \frac{EBIT}{Average\ Total\ Assets} \quad (1)$$

ROA refers to the production power and measures the profit with total assets invested in the business regardless of the financing method. Important is, whether the enterprise can use its capital so effectively.

ADDED VALUE is difference between a particular product's final selling price and the direct and indirect input used in making that particular product. Difference is profit for the firm after all the costs and taxes. This indicator may help investors decide if this business is worth investing, or that there are other and better opportunities. Value added is a higher portion of revenue for integrated companies, e.g., manufacturing companies, and a lower portion of revenue for less integrated companies, e.g., retail companies.

The Comparison of Retail Units on the Czech Republic's Market – Cluster Analysis and Time Series Analysis

Table 1 Large-scale retail units in Czech Republic

Large-scale retail unit	Characteristics
Albert (hypermarkets and supermarkets)	AHOLD Czech Republic, a. s., is a Czech subsidiary of multinational company Ahold headquartered in Amsterdam. Founded in 1991 with first supermarket name Mana.
Billa (supermarket)	BILLA, spol. s r.o. the Czech legal entity, established 1990 in order to build a retail network - a wide range of food and goods common needs.
Globus (hypermarket)	German retail chain, which entered the Czech Republic's market in 1996.
Hruška (supermarket)	Company Maloobchodní síť HRUŠKA, spol. s r.o. is Czech chain - engaged in the sale of food, meat, sausages, drugstores, consumer goods, tobacco, fruits and vegetables.
JIP (discount)	Czech family company JIP východočeská, a.s. founded in 1994, and in 1996 started to build its own network of retail discount stores throughout the Czech Republic.
Kaufland (hypermarket)	A German food chain, which operates hundreds of stores in the Czech Republic, Germany, Poland, Croatia, Romania and the Slovak Republic. In 1998 opened the first store in the Czech Republic.
Lidl (discount)	Lidl Česká republika v.o.s. was founded in 2000 and its core business is discount sales of food and other goods. It is part of the consolidated German business groups Schwarz.
Norma (discount)	The NORMA, k.s. Company deals with discount sales, especially food products.
Penny (discount)	Penny Market s.r.o. operating in the Czech Republic market since 1997, is a subsidiary of the international chain REWE.
Spar (hypermarkets and supermarkets)	SPAR Česká obchodní společnost s. r. o. belongs to the group Austria SPAR. Customers also find non-food goods in addition to food goods in this chain.
Tesco (hypermarkets and supermarkets)	Tesco stores ČR a.s. is part of a group Tesco PLC. They started business in the Czech Republic in 1996.
Žabka (supermarket)	Tesco Franchise Stores ČR a.s. operates retail stores under the brand Žabka, it is also part of a group Tesco PLC, which is one of the largest retail chains in the world.

Source: servers of individual large-scale retail units and Justice.cz

5. METHODOLOGY

Cluster analysis (CA)

The important thing in this paper is the realization of links between large-scale retail units including supermarkets, hypermarkets and discounts. First it is necessary to classify these retails according to their similarity into smaller homogeneous cluster. A cluster analysis in SPSS (Statistical Package for the Social Sciences) will help us with this. As the source database we will use data representative of the Inputs and Outputs of retail companies, as mentioned above.

Table 2 Categories of data

Inputs	Outputs
costs of sales	turnover
consumption of production	production
personal expenses	revenues products services
total assets	added value
equity	profit
registered capital	

Source: Own selection.

The grouping of cluster organizations is based on cluster analysis (CA) that is multivariate statistical method whose primary purpose is to group objects based on the characteristics they possess (Hair, Black et al., 2009). Cluster analysis is a major technique for classifying a large number of information to subgroups, called clusters. It allows to identify similarity of objects and combine them to homogenous groups. Method of cluster analysis requires that transformed data have zero mean and unit variance. The resulting cluster of objects should exhibit high internal (within-cluster) homogeneity and high external (between-cluster) heterogeneity (Hair, Black, et al., 2009).

The most popular procedures of CA are represented by the hierarchical methods and non-hierarchical methods. The hierarchical cluster analysis (agglomerative or divisive) we chose, uses the dissimilarities such as distances between objects during the formation of the clusters. We chose the Euclidean distance (Řezánková, Húsek, Snášel, 2009):

$$D_{ES}(x_i, x_j) = \sum_{l=1}^m (x_{il} - x_{jl})^2 \quad (2)$$

where x_{il} is value of l -th observation on the i -th element, x_{jl} is value of l -th observation on j -th element. Finding similarities of objects (firms) we obtain the matrix of distance (Proximity Matrix). The greater the value of the distance, the smaller degree of similarity between objects.

The last step of cluster analysis is the graphical representation of distance at which clusters are combined through the special figure, called dendrogram. Finally, the extracted clusters are interpreted with using the profile of clusters based on the mean value of the standardized variable.

6. TIME SERIES ANALYSIS

The principle of time-series analysis is the prediction of future developments based on the assumption that history always repeats itself. Generally, the graphs are used for time series illustration. These graphs show the trend of the individual indicators in the time period and the prediction of future development. In this paper, we primarily conducted cluster analysis of the each of Inputs and Outputs for each year.

Retails were divided (by CA) into three homogeneous groups for which we will examine time series derived indicators. All three groups underwent changes in composition, as well as in values in the time period 2008-2012. Each of the smaller clusters are shown in chapter with results. We will try to verify the hypothesis that in line with the trend of the global crisis, our three indicators: EBIT, ROA, ADDED VALUE will be declining to retails. Consequently, the prognosis of the future development will be a downward trend.

The trend reflects a general trend effect for a long period and it can be described by trendy features. In this paper we used a polynomial trendline. This is suitable to use with fluctuating data for estimated representation of the direction and prediction. The degree of the polynomial trend can be determined by the number of variations in the data or the number of curvature (maxima and minima) in the curve. Therefore, a second-degree polynomial trendline has usually one top, third-degree two tops, etc. (Cipra,

1986) The level of conformity data with given model is represented by the value of R^2 .

7. RESULTS

The specific results of cluster analysis for individual years are shown as dendrograms in Appendix 1. As shown in these figures, segmentation of retail stores can be categorized in three clusters in all cases. The first cluster is composed of (in 2008) Hruška, Norma, Žabka, JIP. In 2009, 2010 and 2011 it also includes Spar. The second cluster consists already of Billa, Penny, Globus and Lidl. The third cluster then consist from 2008 to 2011 of Albert, Kaufland and Tesco. Based on these dendrograms, the summaries and the averages of indicators for each of sub-cluster in each of years were calculated. These derived indicators have been analyzed by time series analysis.

Figure 1 shows a growing tendency of the financial ratio indicator – EBIT in all three clusters, and forecasts for future years is equated with an increasing trend. Regarding indicators for the entire segment, we can see the compliance with some clusters. As for cluster 3, which forms the most important supermarkets in the Czech Republic, we see a deeper downturn in 2009, which reflects the impact of the crisis on this segment, even on the EBIT for the entire segment. Positive estimates for future years is a bit distorted by the deep slump, but we can expect another increase in profits in this segment.

Figure 2 shows a little a small diversion in the prediction – ROA indicator for the first and third cluster have an increasing trend, while the second cluster has downward tendency. The second cluster is the average retail unit in

this segment, and therefore could significantly affect the final segment. However, if we look on presented trend and prediction to future years for the whole segment, we will be able to see that the main effect is caused by members of the third cluster.

We see in sub-clusters 1 and 2 decreasing tendency and in cluster 3 increasing trend regarding the Added value indicator. Then we can finally verify the theory of the biggest influencers segment – members of the third cluster here. If we look at is the tendency of added value for the entire segment again we see a growing trend, supported by high levels of $R^2=0,9874$. The segment is therefore expected to further increase in EBIT, ROA, and indicator of added value.

8. CONCLUSION

In this work we focus on the processing of data about the financial condition of Czech retail units using cluster analysis and time series analysis. Large-scale retail stores operating on the Czech republic's market are presented and compared. First we presented the data source, then we showed the methodology of the article, and finally we published our results. The most important thing in this paper is to test the hypothesis of whether the impact of the economic crisis is seen in the analysis of time series of individual clusters in the case of EBIT, ROA and Added value. We provide also a prediction of the financial indicators in the future. The results in this paper confirm the impact of the crisis on this segment especially in 2009 and later improvement. It is shown that people need food whether or not it is a crisis.

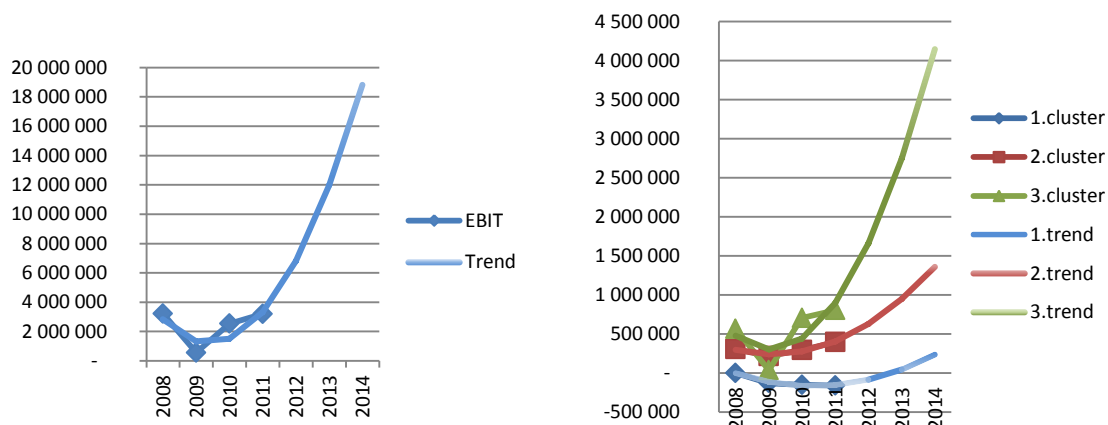


Figure 1 Time series sub-clusters and whole segment with polynomial trendline EBIT

Source: Justice.cz 2014. Own elaboration.

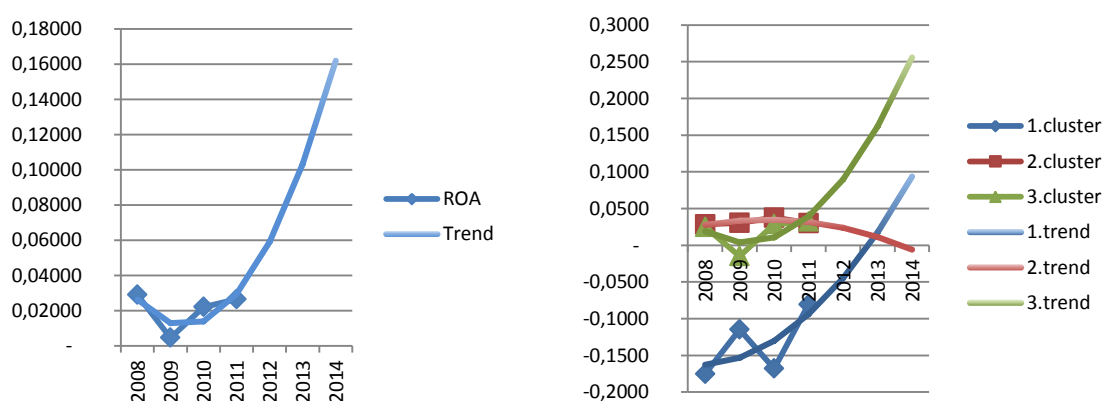


Figure 2 Time series sub-clusters and whole segment with polynomial trendline ROA

Source: Justice.cz 2014. Own elaboration.

The Comparison of Retail Units on the Czech Republic's Market – Cluster Analysis and Time Series Analysis

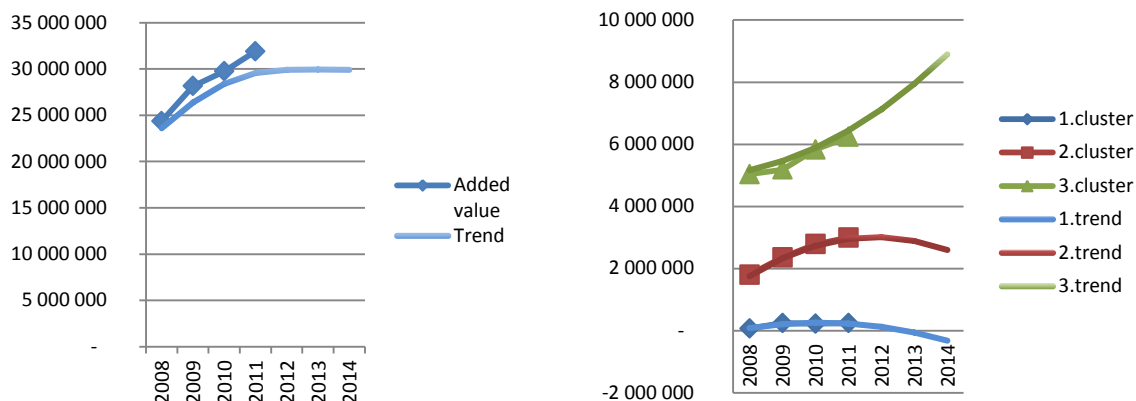


Figure 3: Time series sub-clusters and whole segment with polynomial trendline ADDED VALUE

Source: Justice.cz 2014. Own elaboration.

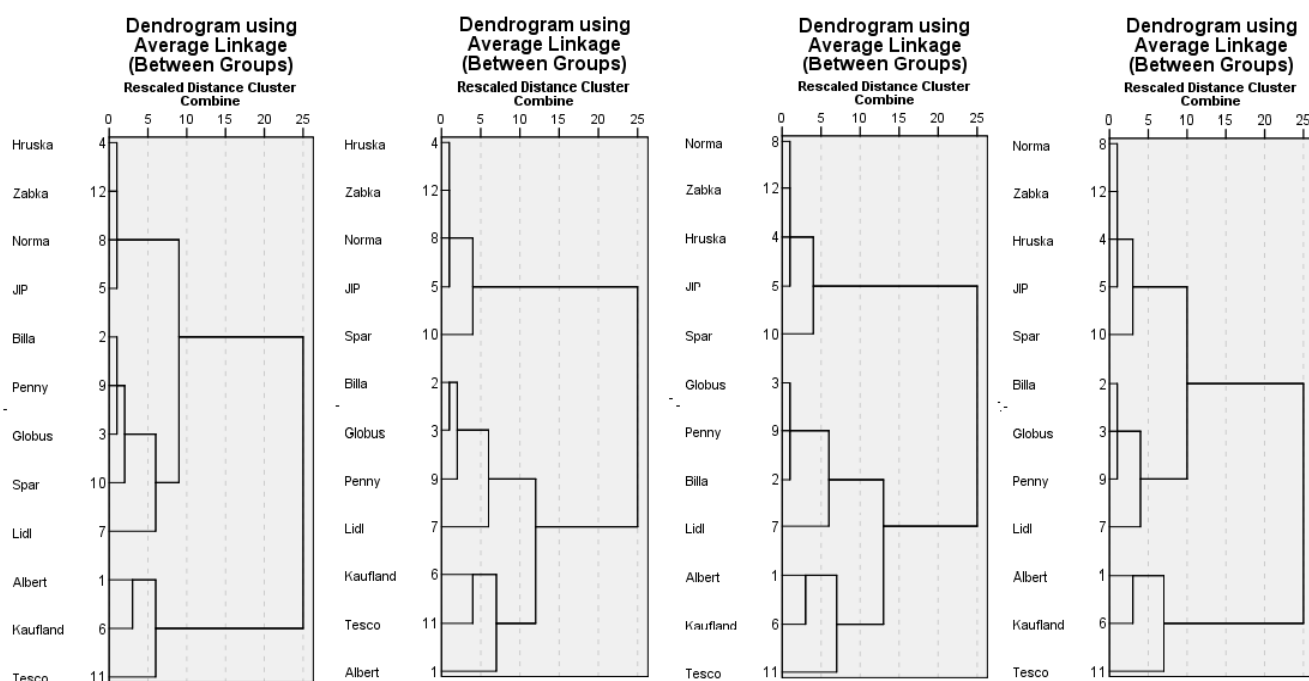
REFERENCES

- [1] BAZARNIK, J. et al. *Handel detailizny: funkcjonowanie i kierunki rozwoju*. Kraków: Oficyna Ekonomiczna, 2004. 260 s. ISBN 83-89355-20-5.
- [2] BOX, G.E.P. *Analiza szeregów czasowych : prognozowanie i sterowanie*. Warszawa: Państwowe Wydaw. Naukowe, 1983. 574 s. wyd. 1. ISBN 83-01-00360-X.
- [3] BROCKWELL, P. J., DAVIS, R. A. *Time series : theory and methods*. New York: Springer, 2006. XVI, 577 s. ISBN 978-1-4419-0319-8.
- [4] CIPRA, T. (1986) *Analýza časových řad s aplikacemi v ekonomii*. 1. vyd. Praha: SNTL/ALFA. 248 s. ISBN 99-00-00157-X.
- [5] HAIR, J. F., W. C. Black et al. (2009). *Multivariate Data Analysis*. (7th ed.). Prentice Hall.
- [6] KNÁPKOVÁ, A., PAVELKOVÁ, D., ŠTEKER, K. (2012) *Finanční analýza: Komplexní průvodce s příklady*. 2.vyd. Praha: GRADA. 240 s. ISBN: 978-80-247-4456-8.
- [7] NEEDLES, B.E., POWERS, M., CROSSON, S.V. (2007) *Principles of accounting*. 10th ed. Boston: Houghton Mifflin Co. p. 1296. ISBN 0-618-73661-1.
- [8] NOWOROL, Cz. *Analiza skupień w badaniach empirycznych: rozmyte modele hierarchiczne*. Warszawa : Państwowe Wydaw. Naukowe, 1989. 147 s. wyd. 1; ISBN 83-01-10009-5.
- [9] ŘEZÁNKOVÁ, H., Húsek, D., Snášel, V. (2009) *Shluková analýza dat*. 2. vyd. Praha: Professional Publishing. 220 s. ISBN 978-80-86946-81-8.
- [10] SULLIVAN, M., ADCOCK, D. *Marketing w handlu detailiznym*. Kraków: Oficyna Ekonomiczna, 2003. 444 s. ISBN 83-89355-09-4

Ing. Lucie ZOTYKOVÁ

Silesian University in Opava, School of Business Administration in Karvina

Univerzitní nám. 1934/3, 733 40 Karviná, Czech Republic



Appendix 1: cluster analysis 2008-2011

Source: Justice.cz 2014. SPSS: Own elaboration.