Ertek G., Sevinç, M., Ulus, F., Köse, Ö., Şahin, G. (2014) "Industrial Benchmarking through Information Visualization and Data Envelopment Analysis: A New Framework", in I. Osman, A.L. Anouze and A. Emrouznejad (eds.) Strategic Performance Management and Measurement Using Data Envelopment Analysis, DOI: 10.4018/978-1-4666-4474-8, ISBN13: 9781466644748, IGI_Global (In Press).

Industrial Benchmarking through Information Visualization and Data Envelopment Analysis: A New Framework

Gürdal Ertek, Mete Sevinç, Firdevs Ulus, Özlem Köse, Güvenç Şahin
Sabanci University, Manufacturing Systems/Industrial Engineering Program
Orhanli, Tuzla, Istanbul, 34956, Turkey

ABSTRACT

We present a benchmarking study on the companies in the Turkish food industry based on their financial data. Our aim is to develop a comprehensive benchmarking framework using Data Envelopment Analysis (DEA) and information visualization. Besides DEA, a traditional tool for financial benchmarking based on financial ratios is also incorporated. The consistency/inconsistency between the two methodologies is investigated using information visualization tools. In addition, k-means clustering, a fundamental method from machine learning, is applied to understand the relationship between k-means clustering and DEA.

Finally, other relevant data, apart from the financial data, is introduced to the analysis through information visualization to discover new insights into DEA results. This study uses information visualization to both explore and reveal the relationships between the different methodologies of financial benchmarking and gain practical insights on the Turkish food industry. The results show that the framework developed is a comprehensive and effective strategy for benchmarking; it can be applied in other industries as well. As a result, our study contributes to the DEA benchmarking literature with a novel methodology that integrates the various benchmarking methods from the fields of operations research, machine learning, and financial analysis.

Keywords: Financial Benchmarking; Data Envelopment Analysis (DEA); Information Visualization; Financial Ratios; K-Means Clustering; Tile Visualization; Graph Visualization; Turkish Food Industry; Istanbul Chamber of Commerce (ISO).

INTRODUCTION

Benchmarking enables companies to see their positions relative to their competitors in order to explore the opportunities to improve their market position. The response to how benchmarking deals with this problem is disguised in its definition: "Benchmarking is the process of continuously measuring and comparing one's business processes against comparable processes in leading organizations to obtain information that will help the organization identify and implement improvements" (Andersen & Jordan, 1998). Various fields of science including computer science, management and operations research, apply alternative methods of benchmarking, such as information visualization (e.g. Self Organizing Maps, SOM), financial ratios (provided that the data is appropriate for financial benchmarking), and analytical methods (e.g. Data Envelopment Analysis, DEA).

This study proposes a framework that integrates a multitude of methods for the purpose of comprehensive benchmarking. As an analytical approach, DEA is employed due to its eligibility among all other methods for its advantage of being a nonparametric technique requiring fewer assumptions (Weill, 2004). Weill (2004) shows that DEA is consistent with standard measures of performance such as the Stochastic Frontier Approach (SFA) and the Distribution-Free Approach (DFA). In addition to its consistency, DEA can be employed to investigate the reasons for a company's inefficiency while showing how much change is needed to achieve efficiency (Galagedera & Silvapulle, 2002).

We combine and compare the DEA results with the insights obtained by other methodologies, including financial ratios, and the results of k-means clustering, as well as other relevant data. Information visualization schemes are extensively and primarily used in a second phase of the analysis in order to bring together the outcomes from all mentioned methods and make them more comprehensive. The steps of the analysis are formalized within an integrated framework. Via this framework, a researcher can combine different benchmarking approaches and data mining techniques to obtain useful insights, as well as to compare the results of the different approaches against each other.

The benchmarking data used in this study is derived from the 2010 Istanbul Chamber of Industry (ISO) List that ranks the top 500 Turkish companies of different industries, and provides financial and other relevant data regarding these companies. In this study, the food industry is selected because it is one of the most developed industries in Turkey and a significant number of companies from this industry are listed in the ISO 500 list.

The complete ISO 500 dataset (for 2000), that includes all the industries, has been subject of earlier research (Ulucan, 2002), and we have selected our inputs and outputs in coherence with this earlier study. While we have not been able to find the efficiency analysis of food industry in Turkey, there have been studies focusing on other industries, such as insurance (Çiftçi, 2004), cement (Karsak and Iscan, 2000), and automative (Bakırcı, 2006). Meanwhile, there exists literature that uses DEA for the benchmarking of food industry or its subsectors in other countries such as agriculture industry in Scotland (Barnes, 2006), food manufacturing plants in the USA (Jayanthi, 1999), productivity growth of Indian food industry (Kumar and Basu, 2008), swine industry in Hawaii (Sharma, Leung and Zaleski, 1997), meat products industry in Greece (Keramidou, Mimis and Pappa, 2011), strawberry greenhouses in Iran (Banaeian, Omid and Ahmadi, 2011), and food industrial companies in Taiwan (Wongchai, Tai and Peng, 2011). An interesting related study by Dadura and Lee (2011) uses DEA for benchmarking the innovativeness of Taiwanese food companies.

Turkish Food Industry Dataset

Sixty-three companies in the Turkish food industry from the 2010 ISO-500 list are included in the study. Each company is presented by its symbol in the Istanbul Stock Exchange (ISE); for the unlisted companies, symbols, which evoke the name of the company, are given. Based on the main food groups, the information regarding the subsector(s) that the companies are involved are appended to the original dataset. These subsectors are organized as Dairy, Flour, Beverage, Sugar, Sea Products, Meat, Poultry, Snack, Oil, Hazelnut, Canned Products, Animal Feed, and other. Data for companies such as Unilever Turkey, a large player in the food industry, was omitted due to missing data. Similarly, Arbel, a leading agricultural producer, was dropped out of calculations due to negative value for one of the inputs.

The financial data used in DEA are the *Equity, Net Asset, Gross Added Revenue, Profit* (from operations) and *Export* values of each company. Apart from the financial ones, some other

useful data such as (number of) *Employees* and *Chamber of Commerce* of each Company, are included in our visual analysis.
These variables can be used for the evaluation of companies in any industry and this is exactly what we intend. The main contribution of the paper is not the analysis of the food industry in Turkey, but the development and the demonstration of a new DEA based analysis framework. Thus, we intended our inputs and outputs to be such that the framework can be applied or adapted to any industry, including food industry, as well as others, such as automobile, machinery, logistics, construction, and many others. The selected inputs and outputs are the drivers and enablers, as well as the basic constructs of the framework. The selected inputs and outputs are among the most basic and popular financial and operational attributes used for benchmarking in any industry, and this is precisely the reason why the Istanbul Chamber of Commerce (ISO) selects and publishes the values for these attributes from among many others. Thus, the selection of the attributes by the Istanbul Chamber of Commerce is a reliable validation for our selection. Furthermore, these attributes are also crucial for the calculation of the financial ratios, which are essential to the multi-perspective nature of our framework.

From among the selected attributes, *Employees, Equity* and *Net Assets* have been selected as inputs since these are operational and financial resources a company uses in conducting its business. On the other hand, *Gross Added Revenue* and *Profit* have been selected as outputs, since these are the financial performance measures for assessing the performance of a company.

METHODOLOGY

Data Envelopment Analysis (DEA)

DEA is an approach used to measure efficiency (or productivity) of an activity or an entity, denoted as a Decision Making Unit (DMU), (Cooper et al., 2006). Typically, a multitude of inputs affect a multitude of outputs for a DMU. It is a challenge to assess how much weight an individual input has with respect to the determination of outputs. DEA does not make any assumptions to choose these weights; instead, the weights are directly derived from the data. The DEA model itself chooses the "best" weights, which have the most adequate properties for the DMU benchmarking. Thus, these weights are not given, but they are defined as variables. Since the weights of the inputs and outputs are determined from the data, there is no need for the units to be congruent. In other words, the units can be of different types (e.g. number of workers, number of machines, dollars). The DMUs with the highest rank of efficiency levels together determine the "efficient frontier" while the efficiencies of other (inefficient) DMUs are calculated according to this efficient frontier. The efficient DMUs that reside on the efficient frontier and have an efficiency score of 1 are either same as or superior to the inefficient DMUs (that have efficiency less than 1) in all productivity dimensions; however, they are inferior to each other in at least one dimension. Each inefficient DMU has its *reference set* which consists of the efficient DMUs that are used to calculate the efficiency of inefficient DMU.

The objective of the input-oriented model is to minimize the input by concurrently keeping the output at least at the current level. The objective of the output-oriented DEA models is to maximize the output level by making use of at most the same amount of inputs. In our analysis, output-oriented BCC model has been applied.

The BCC model by Banker, Charnes & Cooper (1984) is a well-known DEA model; it has almost all of the characteristics of the earlier CCR model. However, unlike the CCR Model, the BCC Model enables the problem to be variable-returns-to-scale. This property is assured through a convexity constraint, which makes the production frontier to be spanned by the DMU a convex hull. As a result, we employ the BCC Model as it can accommodate variable-returns-to-scale without restriction to constant-returns-to-scale; when all the inputs increase by a constant, the output levels do not change.

The input and output determination is one of the challenging steps in performing a DEA study. The dataset for this case study includes financial data, and unfortunately, a standard scheme for the input and output selection does not exist in the financial benchmarking literature. Therefore our choice of input and output measures are based on the idea that achieving more of an output with less of an input is always preferred in DEA. The inputs are selected as *Employees*, *Equity* and *Net Assets*; outputs are selected as *Gross Added Revenue* and *Profit* (from operations)

In our study, DEA computations have been carried out using the Smart DEA Solver software (Akcay et al., 2012), which generates the DEA results in a tabular structure, making it easy to analyze the results using readily available data mining and information visualization software.

Financial Ratios

Can we determine a company's success by analyzing its financial data only? Although one possible way is to examine the growing revenue or positive feedback from customers, one needs more concrete ways to both measure performance and make comparisons within a set of companies. One of these methods is benchmarking companies through the comparison of *financial ratios*. Financial ratios are ratios calculated based on financial data extracted from the balance sheet and/or the income statement (Bragg, 2006). Financial ratios are commonly used in comparing companies with respect to the standardized industry level. They are also used in evaluating a company in terms of its past and current performances.

In general, financial ratios can be categorized in five groups:

- Group I measures how efficient a company uses its assets to make profit. Some ratio examples for Group 1 are Gross Margin, Efficiency Ratio, and Operating Leverage.
- Group II assesses the utilization of assets to generate profits. Some examples of ratios in Group 2 are Return on Net Assets, and Return on Equity.
- Group III determines the company's leverage level. Debt to equity ratio and the interest coverage ratio fit are included in this group.
- Group IV measures the liquidity level of the company. A popular ratio used for this purpose is the Current Ratio.
- Group V is used for stock valuation; examples include price/earnings ratio, and price/sales ratio.

It is easy to find several case studies where financial ratios are used as means of benchmarking. A study on Motorola Company (Collier, H.W., 2004) uses financial ratio analysis to study the operating characteristics of Motorola by comparing the company to two other. Another example considers 20 different financial ratios in evaluating the Critical Access Hospitals (CAH) in Baltimore, Maryland (Pink et al., 2006), using health care information system (Flex Monitoring Team, 2005). The authors discover that CAHs without long-term care make more profit, have fewer liabilities, are more liquid and have higher beneficial use of beds when compared with the long-term care CAHs.

Our study uses the *turnover ratio* which is the ratio of sales revenue to net assets. The justification for this preference is two-fold. First of all, the turnover ratio helps to make logical comparisons among companies in terms of their efficiencies. It is the measure of the overall sales for each dollar of asset that the company has. If the turnover ratio of the company is high, it shows that they manufacture fast and put their products out on the market quickly. Secondly, the turnover ratio data is widely available, or easily computable from widely available sales revenue and net assets data.

A review on the use of financial ratios for measuring efficiency yields several results. The first group of studies uses both ratio analysis and DEA by comparing and/or combining them. Biener and Eling (2011) use DEA in order to evaluate the performance of micro-insurance programs, and also financial techniques such as bootstrapping and truncated regression analysis, and they compare the two results. Feroz et al. (2003) claim that DEA can provide additional benefits for the analyst besides ratio analysis. Hollingsworth and Smith (2003) explain why sometimes DEA results are not sufficient while ratio analysis might yield better results, and they also discuss why an analyst has to use BCC model in DEA when taking financial ratios into the model as inputs. The second group of papers uses financial ratios as inputs and/or outputs in their DEA models. Yeh (1996) also constructs a DEA model using financial ratios as inputs and evaluates the performances of Taiwanese banks. Wen, Lim and Huang (2003) evaluate a DEA model to measure the efficiency of e-commerce business models, again using financial ratios as outputs in the DEA model.

K-Means Clustering

One of the most popular data mining techniques, *k*-means clustering, is an algorithm to cluster entities into *k* partitions based on their attributes (Han and Kamber, 2006). This technique starts with a set of entities and their attributes. It is initiated by defining a specified number of clusters randomly or by using some heuristic data to find the *centroid* (mean point) of each cluster. Then, the algorithm builds new clusters in which all entities belong to the nearest centroid. The "distance" here is the aggregate distance between the cluster centroid and the actual attribute values of entities. Subsequently, the algorithm finds the centroids of the new clusters and repeats itself until the total number of moves of all data objects from a cluster to another is zero. At the end of an epoch, which consists of *N* iterations (where *N* is the number of elements in the dataset), the algorithm may end up with fewer clusters than the number specified at the beginning.

In this study, the attribute set is selected based on our input-output measures for the DEA study since we aim to analyze the relationship between these two approaches. The algorithm is initially applied with 20 clusters while it finally ends up with five.

In the related literature, there are several studies that combine DEA with clustering methodologies. Sharma and Jin (2011) cluster the DEA results in their paper using self organizing maps. Yang (2010) clusters DEA results in order to perform cross-system evaluation of 982 branches of a Canadian bank. Zhou, Huang and Hsu (2008) integrate the DEA results as inputs into *k-means* clustering to evaluate the efficiencies of hotel services in 31 regions of China. Samoilenko and Osei-Bryson (2008) perform a clustering of DEA results and also utilize decision trees in order to visualize the resulting clusters.

Data and Information Visualization

Data visualization is an indispensible method of data analysis, which is highly useful for detecting outliers, finding patterns that may not be seen by some standard methods of data mining, and accordingly, asserting new hypotheses. The biggest advantage of data visualization is that it enables even the novice analysts to interact with and learn from data due to the advanced cognitive capabilities of the visual brain. One of the well-known references in data visualization literature is Chambers et al. (1983), which includes plenty of earlier statistics-related visualization techniques, such as quantile plots, histograms, box plots, symmetry plots, and scatter plots. Nevertheless, most of these techniques could not pass beyond supporting statistical analysis; there was not effective usage of colors and patterns, and it was very difficult to visualize larger datasets in most cases.

In 1990's, with the developments in the field of computer science, the graphical methods of 1980's have elevated into a new level, being referred to as information visualization. The idea in information visualization is to identify undiscovered patterns, come up with actionable insight and constitute maximum benefit out of data, particularly when the data is large-scale and complex. The new era in information visualization is built on ideas from data mining, statistics, and computer graphics, and it compensates for the shortcomings of graphical methods of 1980's. Well-known references on information visualization include Spence (2001), Soukup & Davidson (2002), Keim (2002), and Hoffman and Grinstein (2002).

This study employs three types of visualization tools. The tool known as "colored scatter plot", developed in the 1980's, is also referred to as "starfield display" in the literature. Even though this tool is very simple, it can facilitate comprehension of the data to a maximum extent. We use the Miner3D software (Miner3D) to perform the starfield display. In addition to the starfield display, "the tile visualization" is applied to the dataset, which makes it possible to group the data according to some key characteristics and obtain further insights about it. The Visokio Omniscope software is used for the tile visualization (Visokio Omniscope) tool. Finally, a "graph structure" is implemented to identify relationships between the DMUs, according to the DEA results. Via the implementation of this visualization tool, some additional analyses, which the study is unable to carry out by using the other two tools, are performed. The yEd - Graph Editor (yWorks), a software drawing graphs by using several layouts, is employed to obtain the graphs.

PROPOSED FRAMEWORK

The proposed framework is illustrated in Figure 1. The database that contains the complete data is queried to select the data for DEA and clustering. Clustering, financial ratio analysis and DEA are independently conducted, and the results are obtained in tabular form. Then different visualizations (VIS 1a, 1b, 2, 3, and 4) are constructed using the various combinations of the results and other data in the database in order to derive deeper insight about the benchmarked DMUs and the domain.

Earlier studies in literature have extensively used information visualization for the analysis of DEA results. Ulus et al. (2006) analyze the transportation companies traded in NYSE (New York Stock Exchange) by visualizing the efficiency scores against various attributes of the companies, implementing VIS1a in Figure 1. Ertek et al. (2007) analyze the apparel retail industry in Turkey, introducing VIS1b in Figure 1 in addition to VIS1a. Ertek et al. (2012) introduce the graph visualization of the reference sets, corresponding to VIS4 in Figure 1.

Osther studies that employ visualization for understanding DEA results are reviewed in Ertek et al. (2012).

To perform the complete analyses, we have followed the instructions as outlined in the framework, and the resulting visualizations are categorized into four main types. First three types of visualizations are displaying the relationship between the efficiency scores of DEA and

(i) other relevant data taken from the company database (VIS1a and VIS1b),
(ii) the turnover ratios of each company (VIS2), and
(iii) the cluster indices by *k-means* clustering (VIS3).

Finally, in the fourth type of visualization, reference sets of each DMU, which are found as a result of DEA analysis, are combined with company related data and interpreted accordingly (VIS4). We now present the results based on four different visualization methods that rely on the results of our three benchmarking techniques in the subsequent sections.

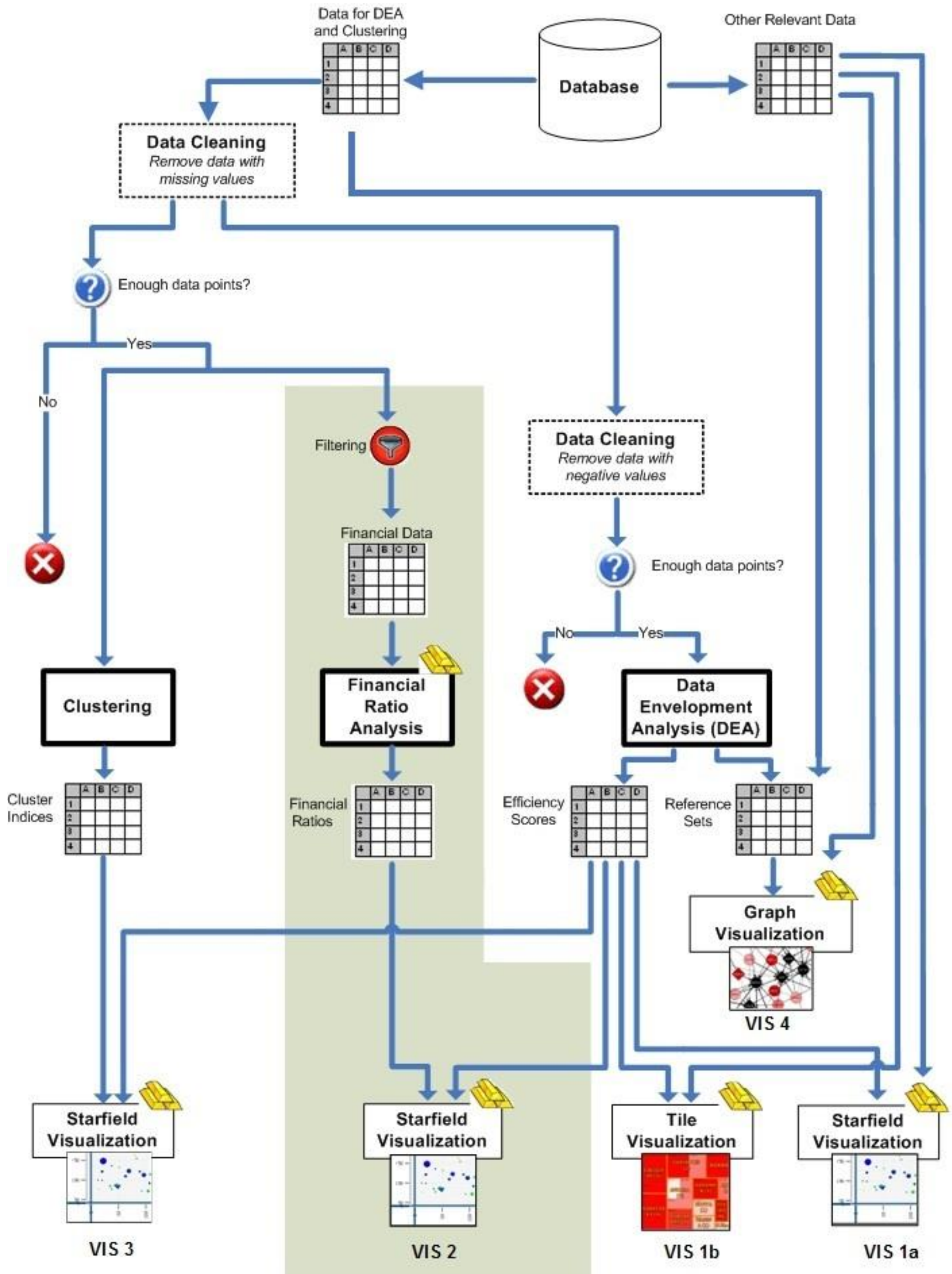

Figure 1. The proposed analysis framework.

DEA Results vs. Relevant Data from ISO 500 (VIS1)

With the DEA analysis, it is usually expected that the efficiency scores should increase as the output values increase and input values decrease. However, since DEA does not take only one output or one input value into consideration, this expectation does not have to hold all the time. In order to see the relationship between the input/output values and the efficiency scores, we use starfield visualization. Figure 2 shows the relationship between the sales revenue and the efficiency values.

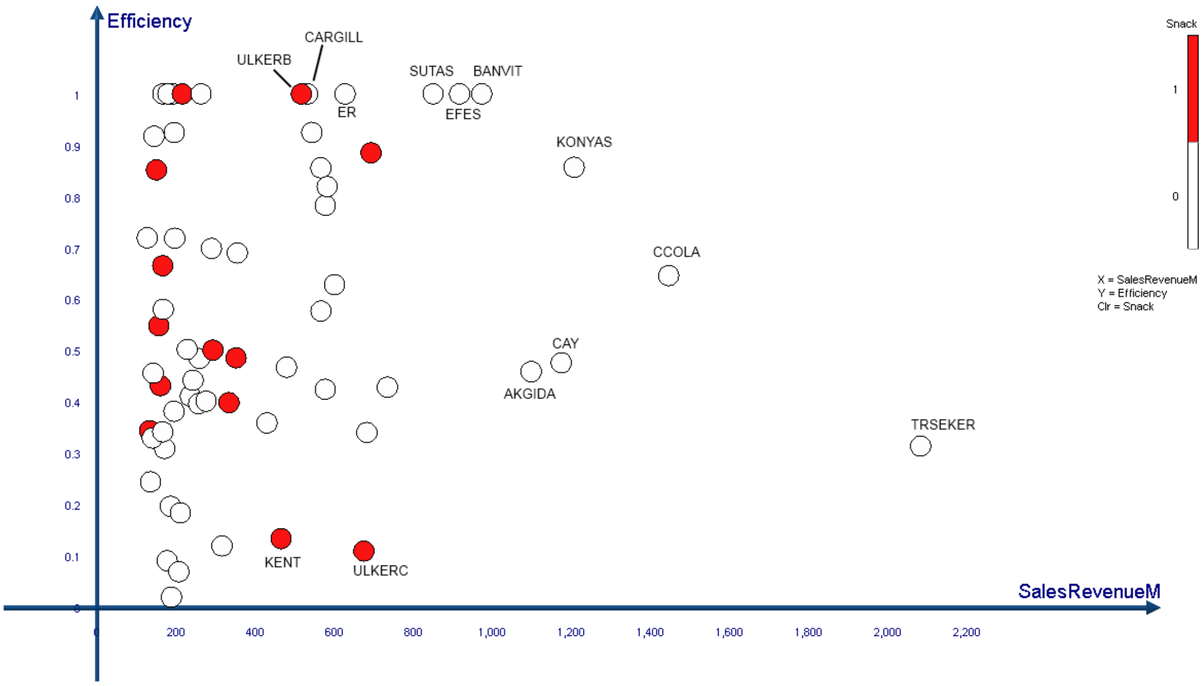

Figure 2. Efficiency vs. SalesRevenueM; red colored circles indicate the SNACK subsector (VIS1a).

In Figure 2 (VIS1a), the X-axis denotes to Sales Revenue value in millions, whereas the Y-axis denotes the Efficiency scores. The circles colored in red are denoting the companies in the Snack subsector. Figure 2 shows that the DMUs having a sales revenue figure larger than 1000 million TL are all inefficient. Companies OZYLMZ, KEREVIT, PROGIDA, NATURA, KESKINK form a cluster on the upper left side of the graph having low Sales Revenue (lower than 400), and the only efficient company is NATURA. As one can see in the above figure, the companies NATURA and ULKERB in Snack industry are located on the top area and are efficient with notably lower Sales Revenue. Another insight reveals that company ETI has the highest Sales Revenue with 692 million TL, and it is not efficient, indeed. This shows that having a higher sales revenue does not necessarily guarantee efficiency. On the contrary, one can observe that those companies with relatively higher sales revenues are generally inefficient in the Snack subsector. KENT and ULKERC are represented as two red circles on the lower side of the chart. These two companies have high sales revenues of 464 million and 674 million, respectively, but they are the two least efficient companies among the Snack companies from the ISO 500 list.

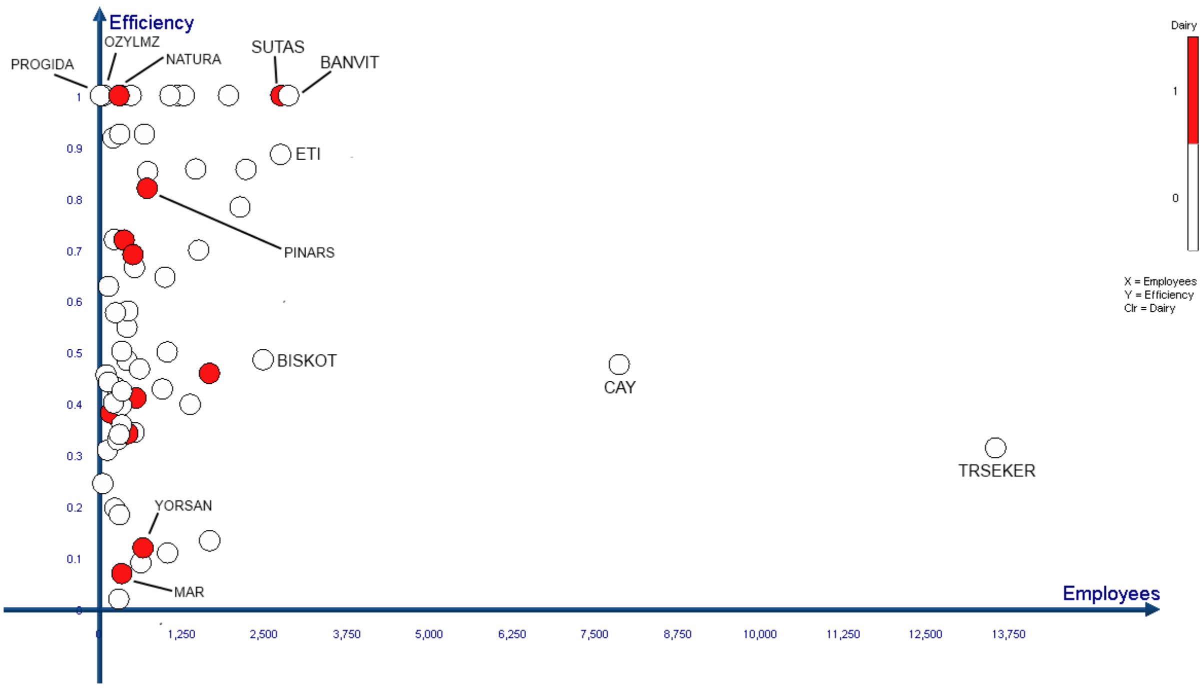

Figure 3. Efficiency vs. Employees; red colored circles indicate the DAIRY subsector (VIS1a).

In Figure 3 (VIS1a), the X-axis denotes to the number of Employees in a company, whereas the Y-axis denotes to the Efficiency scores. The circles colored in red denote the companies in Dairy industry. The two efficient companies are SUTAS and NATURA. We can note that the number of employees of a Dairy company, is not an indicator of its efficiency or inefficiency. While SUTAS has the highest number of employees among all Dairy companies, and NATURA has low employee value of 322, both of them are efficient companies. The two companies with more than 3750 employees are inefficient; CAY and TRSEKER are both state-owned companies. Following CAY and TRSEKER, the two companies with the highest number of employees are BANVIT and SUTAS, and they are both efficient. In Figure 3, we can also remark that SUTAS is the most efficient company with the largest number of employees in the Dairy subsector.

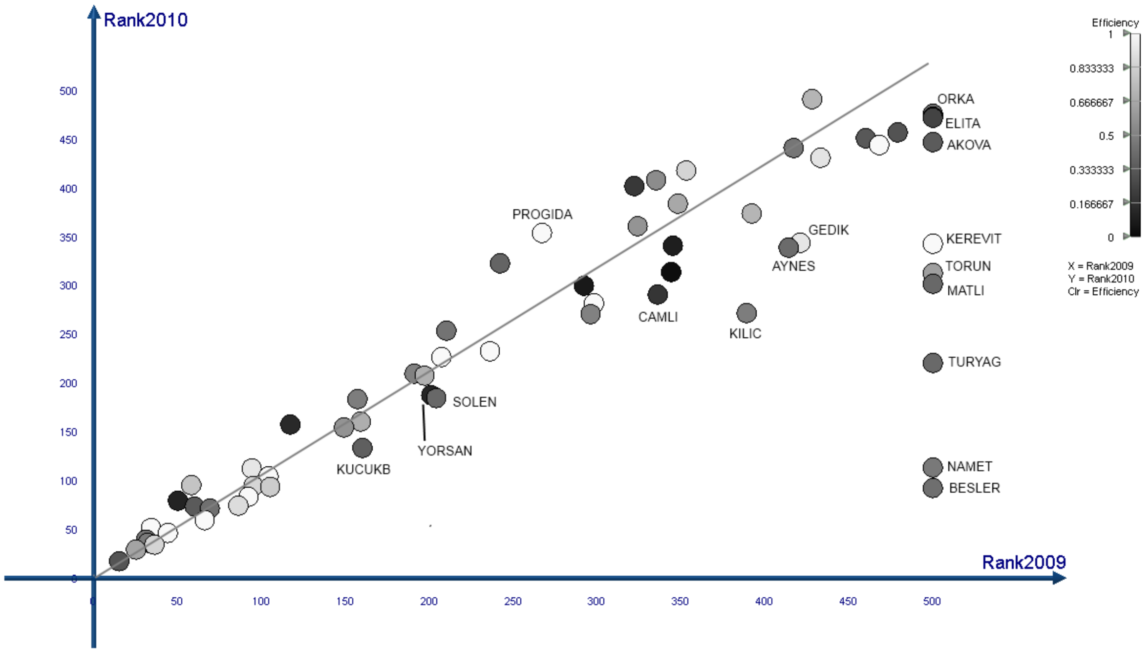

Figure 4. Rank2010 vs. Rank2009; lighter colors denote higher efficiency values.

In Figure 4 (VIS1a), the X-axis denotes the rank of the company in the ISO 500 list in 2009, whereas the Y-axis denotes the rank of the company in the ISO 500 list in 2010. The shades of grey indicate efficiency values; lighter coloring denotes higher efficiency while the white circles represent the most efficient companies. ORKA, ELITA, AKOVA, KEREVIT, TORUN, MATLI, TURYAG, NAMET, BESLER, lined up on the right hand side of the chart were not in the list in 2009, and they made it to the list in 2010. The only efficient company among those is KEREVIT, represented in the graph with white.

Figure 5 (VIS1b) illustrates the tile visualization of ISO 500 Food Companies, grouped according to Provincial Chamber of Industry. Each company is represented by a rectangle, and the size of the rectangle denotes the net asset value of the company. This visualization enables the analyst to focus on the patterns within a city to compare patterns of different cities. For example, comparing the companies in Istanbul Chamber of Commerce to others, we can observe that the total net assets is very large in Istanbul, and there is only a single company, ULKERC, that has very low efficiency score. Within Istanbul, EFES is efficient, while having a great amount of net assets. Ülker Company Group, the largest food group in Turkey, operates in Snack industry with two brands, ULKERC and ULKERB, which produce chocolate and biscuits respectively, and the efficiency gap between the two brands is excessive.

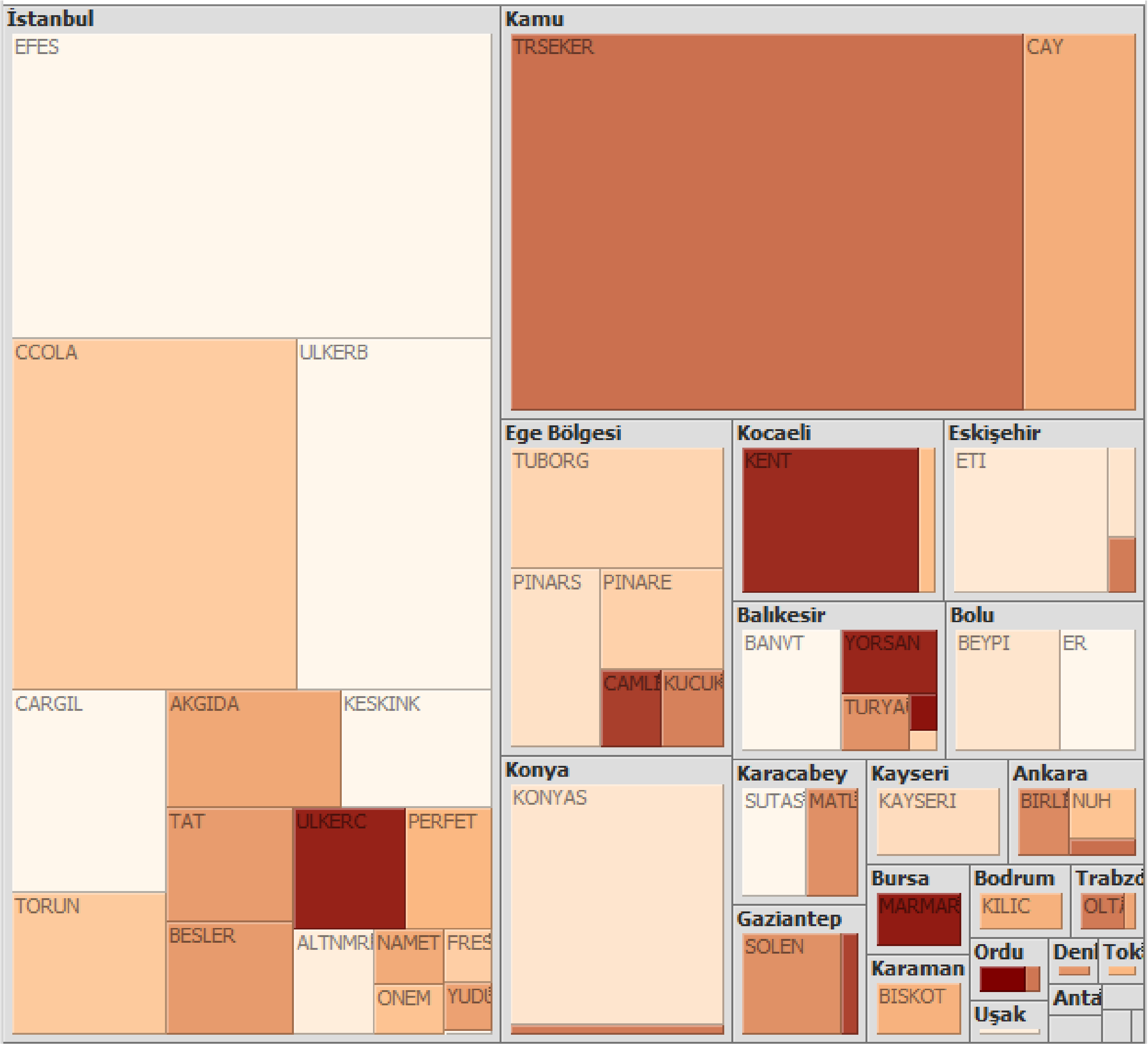

Figure 5. Grouping by Chamber, color denotes efficiency whereas size denotes net assets

Analysis of DEA Results vs. Financial Ratios

In Figure 6 (VIS2), the X-axis denotes the turnover ratio, while the Y-axis denotes the Efficiency Score. The circles colored in red are in Hazelnut (and its byproducts) subsector. Turnover ratio is obtained by dividing the Net Assets of a company by its Sales Revenue. Turnover ratio is a success indicator as to how efficient a business is in terms of generating sales revenue by using its net assets. We can observe that the two outlier companies PROGIDA and OZYLMZ are not only efficient in the Hazelnut subsector , but are also the two efficient companies with the highest turnover ratios (87% and %34 respectively) within the entire dataset. Financial ratios and DEA results produced consistent results for these two companies. However, this may not always hold true, and considering the input and output relations multi-dimensionally using DEA might yield significantly different results. For instance, the two companies with negative turnover ratios on the right side of the chart are KEREVIT and NATURA. This particular outcome shows that DEA results are not by themselves sufficient for the analysis of efficiency, and financial ratios should also be considered. Yet, a multi-dimensional benchmarking method such as DEA is necessary, in addition to ratio analysis. For example, even though SEKERP has higher turnover ratio, it's not very efficient according to DEA results.

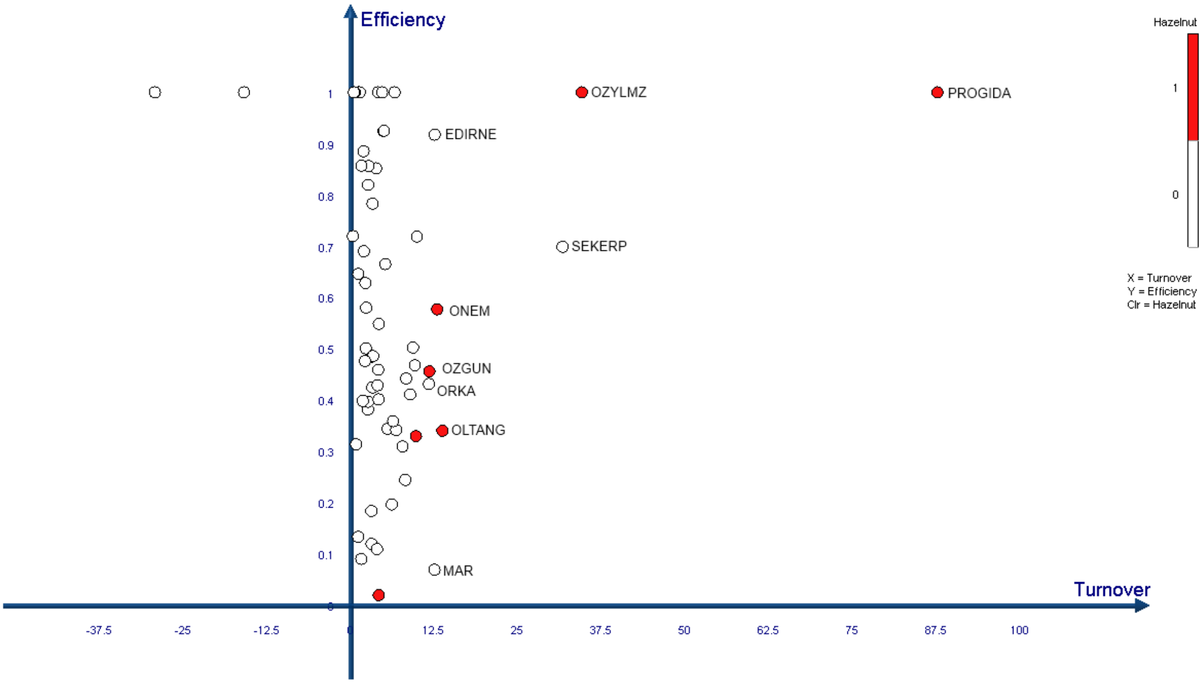

Figure 6. Efficiency vs. Turnover; red color denotes companies in the hazelnut subsector (VIS 2).

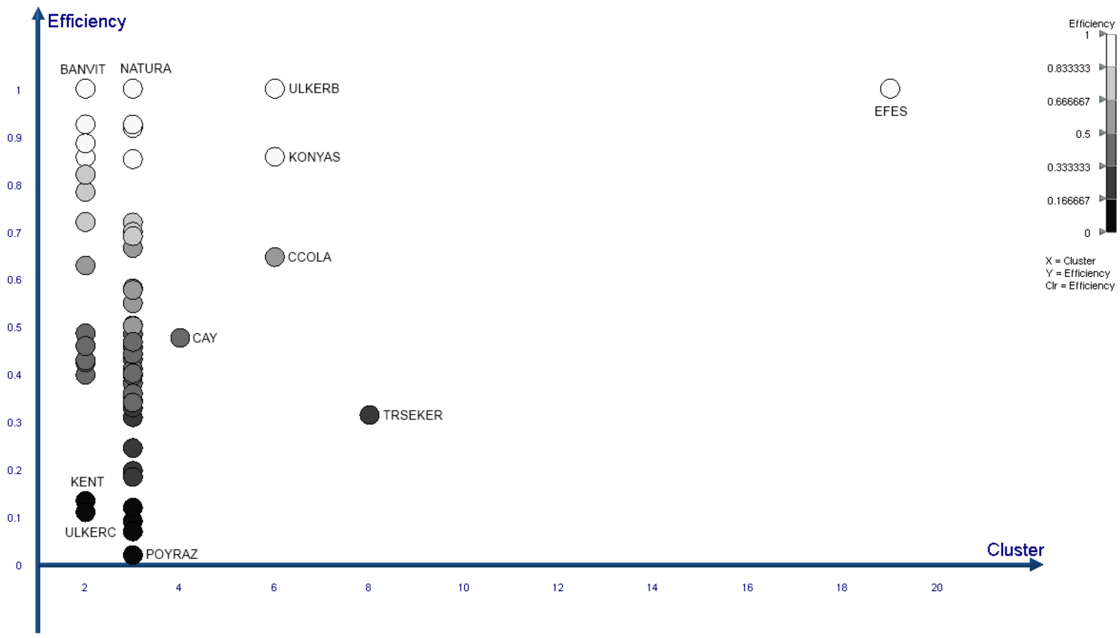

Figure 7. Efficiency vs. ClusterID; lighter colors denote higher efficiency values.

Analysis of DEA Results vs. k-Means Clustering Results

In Figure 7 (VIS3), the X-axis denotes the ClusterID, while the Y-axis and the coloring denotes Efficiency scores. This chart portrays the results of DEA together with the results of k-means clustering. *K-means* clustering was applied with Miner3D software using 20 initial clusters, and running the algorithm for 200 iterations. The variables used in clustering are the five inputs and outputs used in DEA model. The clustering algorithm ended with five clusters, three of which (clusters 4, 8, 19) consist of only a single company. Clustering considers the magnitude of the variable values, ignoring the input-output relationships. Therefore, it is expected that clustering results may not be consistent with DEA results. This expectation is observed in our case study; clusters 2 and 3, which contain many companies, have a high level of variability in efficiency values within the clusters. For example, in cluster 2, where 15 food industry companies with similar magnitudes are clustered, one can observe very inefficient companies such as KENT and ULKERC, besides a very efficient company such as BANVT. Similarly, cluster 3 contains the very inefficient POYRAZ, as well as the efficient NATURA. Figure 8 clearly shows that DEA analysis brings a much better understanding upon the clustering results. Meanwhile, clustering helps us interpret the results of DEA even better. For example, a known property (and also a typical shortfall) of DEA is that it might declare outlier observations as efficient, due to their extreme values in one or two dimensions. In Figure 8, EFES, an efficient company, is seen as the single member of cluster 19, which suggests that the efficiency of EFES might be due to its being an outlier in magnitude in comparison to the other companies. Similarly, CAY and TRSEKER are obviously outlier companies, which appear as the single members of two different clusters. These two companies are inefficient, suggesting that their low scores might be due to the extremely high values of some of its inputs.

Analysis of Reference Sets According to the DEA Results

Figure 8 (VIS4) presents the visualization of the reference sets for the Snack companies within the food industry; each node represents a company, and the darker colors denote larger number of employees in the company. The arcs emanating from node i to node set J denote that J is the reference set for company i. The thickness of an arc denotes the weight for that reference. Efficient companies are those that have arcs emanating and terminating at themselves. One can observe two groups of companies in Figure 8. The first group has ETI, NATURA and ULKERB as the role model, and the second group takes TAM and ORKA as the role model. NATURA exhibits an efficient business model that should be followed by other companies, as it manages to be efficient with relatively low number of employees. The nodes with the same inbound and outbound arc signifies that these companies are efficient.

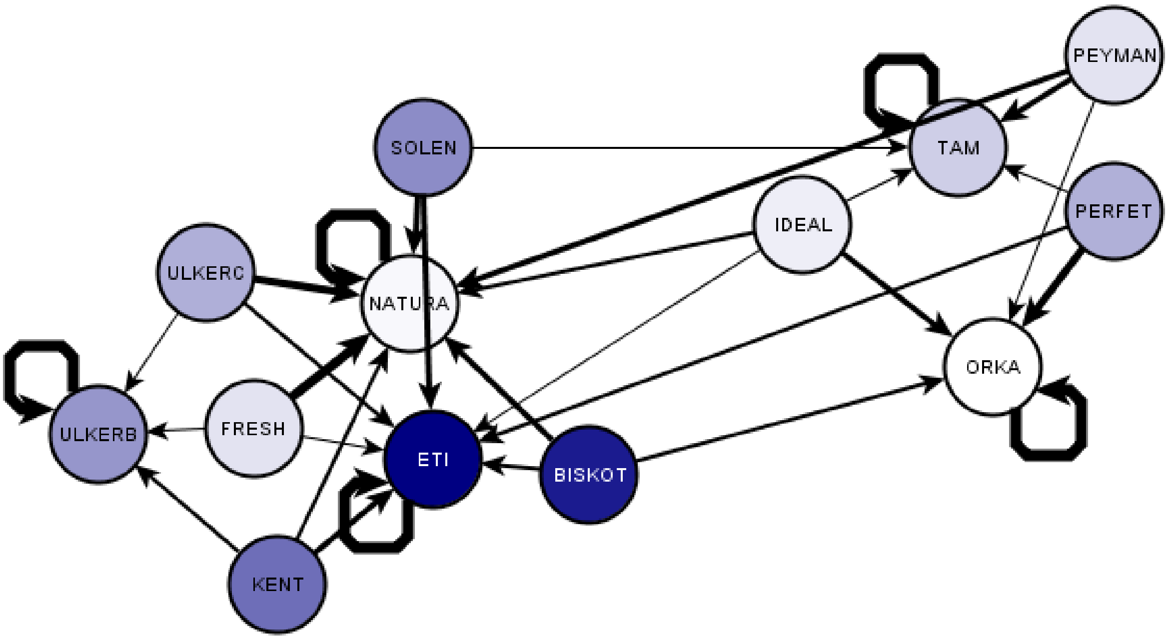

Figure 8. Visualization of Reference Sets; colors denote magnitude of the number of employees.

Conclusions and Future Work

In this chapter, we carried out a financial benchmarking of food companies in Turkey using DEA, and compare and combine the DEA results with the results of two other popular benchmarking/analysis methods, namely financial ratio analysis and *k*-means clustering. Three different information visualization techniques (starfield visualization, tile visualization and graph visualization) were implemented for visual comparisons and analysis.

The study confirms the earlier results in literature in the sense that the results of various methodologies are not necessarily consistent with each other. Therefore, in order to perform comprehensive and fair benchmarking, one needs to employ multiple methods of benchmarking, and interpret the results of these methods together to get the correct results. However, it is difficult to choose the methodologies to employ and interpret them in an acceptable way. The integrated benchmarking framework presented in this chapter serves to this purpose. To the best of our knowledge, such a comprehensive integrated framework does not exist in the existing literature. Therefore, this study makes a methodological contribution to the benchmarking literature, besides presenting a benchmark of the Turkish food industry.

This work can be followed with a crucial research area in which one can carry out formal statistical tests to prove or disprove the hypotheses that are claimed. Moreover, choosing more appropriate financial data, one can use a larger set of financial ratios to compare against the DEA results in order to obtain deeper insights and to compare the different methodologies integrated here.

References

Akcay, A.E., Ertek, G., & Buyukozkan G., (2012). Analyzing the solutions of DEA through information visualization and data mining techniques: SmartDEA framework. *Expert Systems with Applications, 39*, 7763–7775.

Andersen, B., & Jordan, P. (1998). Setting Up A Performance Benchmarking Network. *Journal of Production Planning & Control: The Management of Operations, 9*(1), 13-19.

Bakırcı, F. (2006). Sektörel Bazda Bir Etkinlik Ölçümü: VZA ile bir analiz. *İktisadi ve İdari Bilimler Dergisi, 20*(2).

Banker, R., D., Charnes, A. & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Journal of Management Science, 30,* 1078–1092.

Barnes, A.P. (2006). Does multi-functionality affect technical efficiency? A non-parametric analysis of the Scottish Dairy Industry. *Journal of Environmental Management, 80,* 287 – 294.

Banaeian, N., Omid, M. & Ahmadi, H. (2011) Application of Data Envelopment Analysis to Evaluate Efficiency of Commercial Greenhouse Strawberry. *Research Journal of Applied Sciences, Engineering and Technology 3*(3), 185-193.

Biener, C., & Eling, M. (2011). The Performance of Microinsurance Programs: A Data Envelopment Analysis. *The journal of Risk and Insurance, 78*(1), 83-115.

Bragg, S.M. (2006). Financial Analysis: *A Controller's Guide,*Wiley, 2nd Edition.

Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). *Graphical methods for data analysis*. Pacific Grove, CA: Wadsworth & Brooks/Cole Publishing Company.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429-444

Cooper, W. W., Seiford, L. M., & Tone, K. (2006). *Introduction to data envelopment analysis and its uses*. New York, NY: Springer.

Çiftçi, H. (2004). Türk Sigorta Sektörünün Sorunları; DEA Analizi ile Türk Sigorta Şirketlerinin Etkinlik Düzeylerinin Belirlenmesi. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 13*(1), 121-149.

Dadura, A. M & Lee, T. R. (2011) Measuring the Innovation ability of Taiwan's food industry using DEA. *Innovation: The European Journal of Social Science Research 24*(1)-2, 151-172.

Ertek, G., Can, M. A., & Ulus, F. (2007). *Benchmarking the Turkish apparel retail industry through data envelopment analysis (DEA) and data visualization*. EUROMA 2007, Ankara, Turkey.

Ertek, G., Tunç, M.M., Kurtaraner, E., & Kebude, D. (2012). *Insights into the Efficiencies of On-Shore Wind Turbines: A Data-Centric Analysis*, Accepted for INISTA 2012 Conference. July 2-4, 2012, Trabzon, Turkey (indexed in IEEE Electronic Library).

Feroz, E.H, & Kim, S. (2003). Financial Statement Analysis: A Data Envelopment Analysis Approach. *Journal of the Operational Research Society, 54*, 48-58.

Galagedera, D. U. A., & Silvapulle, P. (2002). Australian mutual fund performance appraisal using data envelopment analysis. *Journal of Managerial Finance*, *28*(9), 60–73.

Han J. & Kamber M. (2006) *Data mining: concepts and techniques*, Morgan Kaufmann Publishers.

Hollingsworth, B., & Smith, P. (2003). Use of Ratios in Data Envelopment Analysis. *Applied Economics Letters, 10*, 733-735.

ISE. Istanbul Stock Exchange. http://www.ise.org.

ISO (2011). Istanbul Sanayi Odası Dergisi, August 2011, issue:545, ISSN 1307-6418. Istanbul Chamber of Industry, Istanbul, Turkey.

Jayanthi S., Kocha B. & Sinha, K.K. (1999). Competitive Analysis of Manufacturing Plants: An application to the US processed food industry. *European Journal of Operational Research, 118*, 217-234.

Karsak, E.E., & İşcan, F. (2000). Çimento Sektöründe Göreli Faaliyet Performanslarının Ağırlık Kısıtlamaları ve Çapraz Etkinlik Kullanılarak Veri Zarflama Analizi ile Değerlendirilmesi.. *Endüstri Mühendisliği Dergisi, 11*(3), 2-10.

Keramidou, I., Mimis A. & Pappa, E. (2011) Estimating Technical and Scale Efficiency of Meat Products Industry: The Greek Case. *Journal of Applied Sciences 11*(6)*,* 971-979.

Kumar M. & Basu P., (2008). Perspectives of productivity growth in Indian food industry: a data envelopment analysis. *International Journal of Productivity and Performance Management, 57*(7), 503 – 522.

Miner3D. http://www.miner3d.com/index.php.

Omniscrope. http://www.visokio.com/omniscope/.

Pink, G.H., Holmes, G.M., D'Alpe, C., Strunk, L.A., McGee P. & Slifkin, R.T. (2006) Financial Indicators for Critical Access Hospitals. *The Journal of Rural Health, 22*(3), 229-236.

Samoilenko, S., & Osei-Bryson, K.M. (2008). Increasing the Discriminatory Power of DEA in the presence of the sample heterogeneity with cluster analysis and decision trees. *Expert Systems with Applications, 34*, 1568-1581.

Sharma, K.R., Leung, P. & Zaleski, H.M. (1997) Productive Efficiency of the Swine Industry in Hawaii: Stochastic Frontier vs. Data Envelopment Analysis. *Journal of Productivity Analysis, 8,* 447-459.

Sharma, M.J., & Jin, Y.S. (2011). Self Organizing Maps Infusion with Data Envelopment Analysis. Retrieved on 24.05.2012, Available under: http://cdn.intechweb.org/pdfs/13293.pdf

Ulucan, A. (2002). ISO500 Şirketlerinin Etkinliklerinin Ölçülmesinde Veri Zarflama analizi Yaklaşımı: Farklı Girdi Çıktı Bileşenleri ve Ölçeğe göre Getiri Yaklaşımları ile Değerlendirmeler. *Ankara Üniversitesi SBF Dergisi, 57*(2), 185-202.

Ulus, F., Kose, O., Ertek, G., & Sen, S. (2006). *Financial benchmarking of transportation companies in the New York Stock Exchange (NYSE) through data envelopment analysis (DEA) and visualization*. 4th International Logistics and Supply Chain Congress, İzmir, Turkey.

Weill, L.(2004). Measuring cost efficiency in European banking: a comparison of frontier techniques. *Journal of Productivity Analysis, 21*, 133–152.

Wen, H.J., Lim, B., & Huang, H.L. (2003). Measuring e-commerce efficiency: a data envelopment analysis (DEA) approach. *Industrial Management & Data Systems, 103*(9), 703-710.

Wongchai, A., Tai C.F., & Peng, K.C. (2011) An Application of Super-efficiency and Tobit Method

for Financial Efficiency Analysis of Food Industrial Companies in Taiwan. *Emergency Management and Management Sciences (ICEMMS), 2nd IEEE International Conference.*

Yang, Z. (2010). Cross System Bank Branch Evaluation Using Clustering and Data Envelopment Analysis. *Lecture Notes in Computer Science (LNCS), 62*15, 238-245.

Yeh, Q.J. (1996). The Application of Data Envelopment Analysis in Conjunction with Financial Ratios for Bank Performance Evaluation. *The Journal of Operational Research Society, 47*(8), 980-988.

yWorks. http://www.yworks.com/en/index.html.

Zhou, Z., Huang, Y., & Hsu, M.K. (2008). Using Data Envelopment Analysis to Evaluate Efficiency: An Explanatory Study of the Chinese Hotel Industry. *Journal of Quality Assurance in Hospitality & Tourism, 9*(3), 240-256.

KEY TERMS & DEFINITIONS

Financial Benchmarking: The benchmarking of companies using financial data.

Data Envelopment Analysis (DEA): A non-parametric analytical method for benchmarking a group of entities.

Information Visualization: The field of computer science that works with the visualization of large-scale complex data for discovering new useful knowledge.

Financial Ratios: Ratios used in finance; they are computed as the ratio of two financial metrics.
K-Means Clustering: An unsupervised machine learning / data mining method for clustering a set of entities into clusters of similar entities.