UNIVERSIDADE CATÓLICA PORTUGUESA

Benchmarking and Analysis of Store Performance in a Retail Group

Ana Catarina Lourenço Mendonça

Católica Porto Business School

March 2019

UNIVERSIDADE CATÓLICA PORTUGUESA

Benchmarking and Analysis of Store Performance in a Retail Group

Final work in Project modality submitted to Universidade Católica Portuguesa to obtain the Master's Degree in Management (specialization in Business Analytics)

by

Ana Catarina Lourenço Mendonça

under the supervision of

Professora Doutora Conceição Silva

Católica Porto Business School

March 2019

"If you can measure it, you can manage it"

Robert S. Kaplan

Acknowledgments

This dissertation has been, beyond question, one of the most challenging projects that I have undertook, for the commitment, rigor and dedication required, but also for the enriching journey. In fact, this project has brought me theoretical knowledge that I could apply in a real-life context, offering me a close perception of the firm's activity.

First of all, my main and greatest acknowledgement goes to my supervisor Professor Conceição Silva, for accepting the supervision of this project and for helping and supporting me when things seem to be adverse. The guidance regarding the way forward in the development of the work and the knowledge transmitted to me during this period was of major importance for the accomplishment of the project.

Secondly, I would like to thank my family and my friends for the all the support, understanding and encouragement in this last step of my education.

And finally, I would like to express my deepest gratitude to my supervisor at DESFO, SA, Pedro Figueiredo, for allowing me to use DeBorla Group data in this study, for all the challenges proposed in order to make my project better and, above all, for the availability and continuous support of my decisions in this dissertation.

Abstract

An everyday need of retail firms operating in saturated markets, where they face fierce competition, is the accurate and unbiased analysis of store performance. The aim of this work is to analyse and benchmark store efficiency and propose targets for store performance improvement in a Portuguese Retail Group. Panel data for 27 stores in the period 2015 to 2017 has been used to allow the assessment of store efficiency and the setting of improvement goals for the inefficient units, while identifying adequate efficiency drivers.

The methodologies and literature review allowed the identification of the techniques to be applied in the study: 1) Data Envelopment Analysis, to measure stores relative efficiency and to set improvement targets to the stores under analysis; 2) Tobit Regression Model, to identify the predominant factors leading to efficiency. Literature review and analysis of the available dataset enabled the selection of the variables to be used when applying each of the above techniques.

The work undertaken led to a number of important conclusions, regarding different aspects of this real life multivariable problem: 1) identification of efficient and inefficient stores; 2) store efficiency distribution by geographical location; 3) evolution of store efficiency over the time; 4) definition of performance targets for the inefficient stores; 5) benchmark highest performing units against lowest performing units with different indicators and 6) identification and quantification of the environmental factors influencing store efficiency.

Key-words: Store Performance, Store Efficiency, Retail Analytics, Retail Performance Benchmarking, Data Envelopment Analysis, Tobit Regression Model.

Resumo

As empresas de retalho que operam em mercados saturados, com uma concorrência feroz, têm a necessidade constante de fazer uma análise precisa e imparcial do desempenho das suas lojas. O objetivo deste trabalho é analisar e comparar a eficiência das lojas um grupo de retalho em Portugal propondo objetivos de melhoria de desempenho. Foram utilizados dados em painel para 27 lojas referentes ao período de 2015 a 2017, para avaliar a eficiência das lojas e estabelecer objetivos de melhoria para as unidades ineficientes, identificando também os fatores com maior impacto na eficiência.

As metodologias e a revisão de literatura permitiram a identificação das técnicas a aplicar no estudo: 1) *Data Envelopment Analysis*, para medir a eficiência relativa e definir objetivos de melhoria para as lojas em análise; 2) *Tobit Regression Model*, para identificar os fatores predominantes que conduzem à eficiência. A revisão de literatura e a análise da amostra disponível suportaram a seleção das variáveis a serem utilizadas na aplicação de cada uma das técnicas.

O trabalho realizado conduziu um conjunto de conclusões importantes, relativamente a diferentes aspetos deste problema real multivariável: 1) identificação das lojas eficientes e ineficientes; 2) distribuição da eficiência das lojas por localização geográfica; 3) evolução da eficiência das lojas ao longo do tempo; 4) definição de objetivos de desempenho para as lojas ineficientes; 5) Comparação das lojas de elevado e baixo desempenho relativamente a diferentes indicadores e 6) a identificação e quantificação das variáveis ambientais que afetam a eficiência.

Palavras-chave: Desempenho da loja, Eficiência da loja, Análise de retalho, Comparação de desempenho no retalho, *Data Envelopment Analysis, Tobit Regression Model*

Index

Index of Figures

Index of Tables

Acronyms

DEA	Data Envelopment Analysis
CCR	Charnes, Cooper and Rhodes
BBC	Banker, Charnes and Cooper
VRS	Variable Return to Scale
CRS	Constant Return to Scale
DMU	Decision Making Unit
OLS	Ordinary Least Squares
TE	Technical Efficiency
PTE	Pure Technical Efficiency
SE	Scale Efficiency
QGIS	Geographic Information System
EBITDA	Earnings Before Interest, Taxes, Depreciation and Amortization

Chapter 1 - Introduction

Today's economic conjuncture is uncertain and this leads to a volatile economic profile of families. The response of retail sector to market's volatility is being very dynamic and in constant transformation, as to meet the households' needs everywhere and in an efficient way (i.e. to create more value with fewer resources). Ever increasing competition is a reality that retailers know well and always take into consideration, because it is a major challenge to their profitability. In a saturate retail market, firms have more and more necessity of improving their performance.

In most cases, firms performance assessment is based on operational and financial indicators, composing their financial statements, and the main indicator used to measure performance is profit. However, this approach have a number of limitations (Fernandes, 2007): (1) The profit being the only indicator may biase performance evaluation, because many different factors influence firm's economic activity, such as location, trade area, etc. Therefore, a store that faces a lot of competition may attain less profit comparatively to others, though, if the store is making better use of resources, may achieve a superior performance. (2) An operational or financial ratio has a clear and evident interpretation if it is related to only one resource and one outcome. To seek a more extensive analysis, various ratios should be applied simultaneously. Nevertheless, each ratio could lead to different and contradictory statements on the firm's performance. (3) In order to obtain an unique performance measure, an aggregated ratio should be considered, in spite of the inherent subjectivity in the definition of the weights attributed to each indicator.

Another trend in measuring firm's performance is the comparison between the initially proposed objectives (as set in the planning of a firm's budget) and the results actually achieved. When doing this, account must be taken that those objectives are defined through a number of simplified hypotheses, which could lead to a deviation from reality, thus being another approach featuring limitations (Fernandes, 2007).

In order to decrease the impact of such limitations and obtain a more accurate and unbiased performance evaluation, two widely used methodologies have been applied in this dissertation: Data Envelopment Analysis (DEA) and Tobit Regression Model[1].

1. Research objectives

The thesis is on store performance evaluation and improvement in the retail sector. To achieve this goal, two methodologies have been applied in a real business case. DESFO, SA was the retail sector group under study.

Firstly, Data Envelopment Analysis was the selected tool to evaluate store efficiency. The results provide a score for that efficiency, which allow us to know how efficient the stores are. In order to make a correct and impartial analysis, each store is evaluated by a comparison to its peers. So, could identify which stores are fully efficient and which are not. DEA also provides improvement goals for the inefficient units, in order to drive them to become completely efficient. These so-called improvement goals are linked to the variables - inputs and outputs - that compose the DEA model.

Secondly, a study of the impact that certain variables have on store efficiency, as measured by the previous methodology, should be undertaken.

These variables are not the same that characterize the DEA model, because the aim is to study the effect of variables that the stores do not have under their control. With the help of Tobit Regression Model, a multiple regression model was estimated to provide information on the statistical significance of those variables and

[1] The Data Envelopment Analysis and Tobit Regression Model will be addressed in more detail in the following chapters.

quantifying the impact on efficiency. This allowed us to know which variables should be considered by the firm in the future when opening a new store.

2. Structure

In the first chapter, the purpose and motivation of the thesis is presented, as well as the research objectives and the methodologies used to pursue them. In the second chapter, a brief description of the firm used as real business case is presented. In Chapter 3, the used methodologies are theoretically and mathematically described. Chapter 4 proceeds with a literature review of those techniques and their practical applications to retail. In Chapter 5, a contextualization of the stores and a rigorous description of the variables that compose the available dataset are made. Chapter 6 advances with the development of the DEA model applied to a group of selected stores and with the display of the respective results. In Chapter 7, the environmental variables that impact store efficiency are presented. Last but not least, Chapter 8 outlines the conclusions of the research study and provides contributions for future work.

Chapter 2 – The empirical setting

1. The firm - DESFO SGPS, S.A.

DESFO was founded 30 years ago, in 1984. The name of the company, is a combination of Desicor and Forcargo, which were the first two companies in the group. Initially, the firm was specialized in the production of furniture. Along the years it has diversified production to other products and broaden activities to retail and logistics.

Nowadays, DESFO holds four independent business units: (1) Industry, where manufacturing is driven by efficiency and quality; (2) Retail, focusing on services and products that meet costumer's demand; (3) Logistics, diversifying services in the supply chain; (4) Investments, where the firm commits to new projects and partnerships.

Desfo has 100% Portuguese capital and owns eight different brands: DESICOR and ONESKIN for manufactured product lines, DeBorla and MEGA in retail, FORCARGO and TRANSNAUTICA in logistics, and DESFOINVEST and DEKOINVEST in investments. All the activities are located in Portugal, except those of DEKOINVEST in Romenia and those of MEGA in Angola.

The strategy adopted by the company may be seen as opportunistic, always trying to take advantage of the business opportunities that were appearing over time. Initially, DeBorla started out as a discount store business, with the price playing a decisive role in the competitiveness. With its 2013 Joint Venture, DeBorla has become a household store, focusing on interior design products. Price was no longer the critical factor, with design, market trends and quality being of major importance.

In 1998, DeBorla opened the first store, and with that DESFO began its activity in Retail. DeBorla currently has 34 stores in Portugal, of which five are located in Madeira and Açores. DeBorla currently offers seven product categories, where Kitchen represents 21% of sales, Interior Design 20%, Storage 11%, Textile 10%, Garden 8% and Bathroom 7%.

Chapter 3 – Methodologies

1. Data Envelopment Analysis (DEA)

According to the extensive literature, Data Envelopment Analysis, also called DEA, is a nonparametric, deterministic approach and also a quantitative and analytic tool, that may be used to measure the performance of firms in different types of studies.

It is used to measure performance of Decision Making Units (DMUs), that convert multiple inputs (resources that the firm has in its power) into multiple outputs (results produced by the firm) (Thanassoulis, 2001). Since it handles data inputs and data outputs, it is known as a "data oriented" approach, as stated by (Cooper et al., 2011) and (Cooper et al., 2004). The correct identification of input and output variables in each specific case is obviously crucial so, as a general rule, the inputs and outputs used in a DEA model are chosen according to the specific firm's strategy and objectives with the advice of the relevant literature.

The acceptance and use of DEA has been rapidly increasing. Examples of applications may be found in performance evaluation in many industries, hospitals, universities, cities, courts, and other businesses, including also the performance assessment of countries, regions, etc. (Cooper et al., 2004). Many benchmarking studies use DEA, allowing sources of inefficiency in many profitable firms to be identified.

As emphasized by several authors, namely (Thomas et al., 1998), DEA focus on frontiers, in other words, it identifies the limit of outcome/output attained by each DMU given a set of resources/inputs, instead of portraying a major trend as in Ordinary Least Squares (OLS) method, which gives an average performance.

As reinforced by (Banker & Morey, 1986), the purpose of a DEA model is to determine an efficiency level to each unit in the reference set, based on a peer performance comparison. These peers, called benchmarks, are the best performing units that can be used as role models for comparison in the evaluation of the less efficient units (Thomas et al., 1998). The set of efficient DMUs form a so-called "efficient frontier", allowing the determination of efficient target levels of the inputs and outputs for the less efficient units, called non-frontier units (those levels that would render the units efficient).

(Farrell, 1957) is considered a pioneer in this matter, as he proposes a way of measuring efficiency. Observing that there were many restrictions when combining multiple inputs in any measure of efficiency, he proposed an approach to overcome this problem, extending the concept of productivity to a more general concept of efficiency. This concept is based on the extended Pareto-Koopmans definition (Koopmans, 1951) and is also known as Technical Efficiency: a unit is technically efficient if none of its inputs or outputs could be reduced or improved without, simultaneously, increasing or deteriorating other inputs or outputs. Thus, through the efficiency measurement of a DMU relatively to others similar DMUs, the concept of Relative Efficiency emerges (Cooper et al., 2004). This was formalized by (Charnes et al., 1978) and not by (Farrell, 1957), despite the fact that the definition is in agreement with Farrell's models. The concept of Technical Efficiency was distinguished by (Farrell, 1957) from other concepts of efficiency, like Scale Efficiency.

Extending Farrell's work considering only one output, the DEA methodology was introduced, in its present form, by (Charnes et al., 1978), with multiples inputs and outputs being considered in the analysis. They described DEA as a mathematical programming approach that helps building production frontiers, using only efficient DMUs where the output is maximized.

(Thomas et al., 1998) highlights relevant points that should be taken into consideration in DEA models, which are supported by (Parsons, 1992), (Thurik,

1992) and (Kamakura et al., 1996). First of all, weights should be assigned, when considering multiples inputs and outputs, because they will reflect the relative importance of each variable. Secondly, environmental variables that have an impact on efficiency, such as the location of a retail store or trade area factors, should be considered in the DEA models (directly or indirectly). The third point is that, traditional ordinary least squares, used to establish the input/output relation, is not ideal because it is based on averages. DEA, in fact, is more accurate than typical regression analysis when identifying efficient and inefficient units, since it considers each unit separately and compares each unit only with the most similar efficient units. Furthermore, DEA is much more effective when motivating and rewarding managers, like store managers, in the use of practices that could be observed in a specific store and transferred to other stores. A distinction between inputs under the control of store managers and the ones that only headquarters managers may have under control should obviously be made. Last but not least, more than one outcome should be taken into account, because DMU assessment always regards multiple and, in some cases conflicting, performance measures.

1.1 Basic Formulations of the DEA Methodology

The DEA methodology formulated by (Charnes et al., 1978) generalizes the single-output to single-input classical ratio definition to multiple outputs and inputs.

Consider n DMUs (j=1,…,n), which are referred to an entity evaluated by their ability to convert inputs into outputs, where each of the DMUs consumes m inputs (x_m) and guarantees s outputs (y_s). DEA is formulated as a mathematical programming problem and is that it does not need weights to be pre-assigned to the inputs and outputs. Those weights are optimized by solving the DEA model. The goal is to maximize the ratio between the weighted outputs and the weighted inputs, which can be formulated as a fractional programming model to evaluate the efficiency of each DMU in turn, as shown below:

$$max\ h_{0\ (u,v)} = \frac{\sum_r u_r\, y_{r0}}{\sum_i v_i\, x_{i0}}$$

St:

$$\frac{\sum_r u_r y_{rj}}{\sum_i v_i x_{ij}} \leq 1, \qquad j = 1,\ldots,n,$$

$$u_r \geq \varepsilon, \qquad r = 1,\ldots,s,$$

$$v_i \geq \varepsilon, \qquad i = 1,\ldots,m.$$

Where,

- ε is a very small positive number,
- j is the index number related to each of the n DMUs,
- 0 is the index of the DMU under analysis,
- y_{rj} and x_{ij} are the observed outputs and inputs, respectively, of the DMU_j,
- u_r and v_i represents the assign weights to the output and input, respectively.

With this model an efficiency measure for any DMU is obtained, where the ratio of the total weighted output to the total weighted input for all DMUs is maximized, as shown by the previous equation. This is subjected to the constraint that for the ratio above every DMU must be less than or equal to one, ensuring that no unit will have an efficiency score greater than one. Thus, the efficiency measure of any DMU_j, expressed as h_j^*, assumes a value between 0 and 1, obtained by the optimal solution values of the outputs y_{rj}^* and inputs x_{ij}^*. Therefore, an unit could attain the maximum efficiency score of 1 and, if that is the case, is an efficient unit; on the other hand, if the unit attains an efficiency score less than 1, it is an inefficient unit. The efficient units will form the efficiency frontier, with the inefficiency of the remaining units being measured by the distance to that frontier. Furthermore, through a peers comparison, information is obtained about feasible input's reductions, maintaining the levels of outputs. The input's savings and, on the other hand, the output's increases without increasing the levels of inputs are two approaches for the improvement of inefficient DMUs. It is therefore possible to determine the input's and output's target values for the inefficient DMUs to become fully efficient.

1.2 CCR and BCC models

The DEA model uses two important approaches, both of them analysing the relative efficiency of a DMU. The first approach is the so-called CCR (Charnes, Cooper and Rhodes) model, also called Constant Return to Scale (CRS) model introduced by (Charnes et al., 1978). The second formulation is the BCC (Banker, Charnes and Cooper) model, also called Variable Return to Scale (VRS) model by (Banker et al., 1984).

The selection of any of these models depends on the size variability of the DMUs. In the CCR model the units present a similar size, meaning that maximum efficiency is always achievable irrespective of the size of the DMU. Under the VRS model, only units of similar size are compared as it is assumed that size matters. The choice for one or another approach depends on the empirical setting and whether each of the above assumptions is the most adequate.

Both models can follow two type of orientations leading to an input-oriented model or an output-oriented model. In the first case, the goal is to minimize the inputs keeping the outputs constant. In the latter case, the objective is to maximize the outputs keeping the inputs constant. The efficiency scores always fall in a range between 0 and 1, regardless if it is a CRS or a VRS model.

In brief, two aspects have to be considered in order to proceed with this method and to measure efficiency in an accurate way. Firstly, it is important to choose between an input-oriented and an output oriented model. Secondly, a decision on the returns-to-scale is needed, a firms activity may be characterized as having Constant Return to Scale or Variable Return to Scale. In order to compare both approaches, the CCR and BCC models have to be calculated.

(Banker et al., 1984) developed a linear programming problem, in order to compute the efficiency scores as follows:

BCC model: Input Oriented[2]

$$e_{j_0} = \max \theta_{j_0} = \sum_{r=1}^{s} u_r y_{rj_0} + u$$

St:

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0$$

$$\sum_{i=1}^{m} v_i\, x_{i0} = 1$$

$$u_r \geq \ \varepsilon, \qquad r = 1, \dots, s$$
$$v_i \geq \varepsilon, \qquad i = 1, \dots, m$$

Where:

- ε is a very small positive number
- u is free

By using dual linear programming, the BCC model presented above can be developed as an "Envelopment Model":

[2] In this section, only the BCC model with an input orientation will be presented, because those models will be the ones used in this study.

$$e_{j_0} = min\ \theta_{j_0} - \ \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+)$$

St:

$$\sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- = \theta x_{i0}, \qquad i = 1, \dots, m,$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = y_{r0}, \qquad r = 1, \dots, s,$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, s_i^-, s_r^+ \geq 0, \qquad \forall i\ , j, r$$

Where:

- s_i^- and s_r^+ are the slack values of the inputs and outputs, respectively
- θ_k(k=1,…,n) is the efficiency value of the DMUk
- ε is a very small positive number

This model identifies the peers of the DMU under analysis (DMU_{j_0}), allowing the DMU relative efficiency (θ_{j_0}) and optimal inputs and outputs values to be measured. This optimal values are given by a linear combination ($\sum_{j=1}^{n} x_{ij}\lambda_j$, $\sum_{j=1}^{n} y_{rj}\lambda_j$) of the benchmarks of DMU_{j_0} – the efficient units that dominate the inputs and outputs of DMU_{j_0}.

Taking into account the above mentioned definition of efficiency and applying it to these models, a DMU is completely efficient, if $\theta^* = 1$ and $s_i^{-*} = s_r^{+*} = 0$. However, a DMU is weakly efficient if at least one slack is different from zero, $s_i^{-*} \neq 0$ and/or $s_r^{+*} \neq 0$, for any i and r (Cook & Zhu, 2005).

The technical efficiency of the DMU_{j_0} is attained by reducing the inputs by the value permitted by the slack. So, the input target value is given by the product of the observed input with the efficiency score obtain by the DEA model, minus the slack value. On the other hand, the technical efficiency of DMU_0 can be achieved by increasing the outputs by the value permitted by the slack. In this case, the output target value is given by the product of the observed output with the efficiency score obtained, minus the slack value.

The values of the output and input targets are measured as following:

$$\widehat{x_{i0}} = \theta^* x_{i0} - s_i^{-*} \qquad i = 1, \ldots, m$$

$$\widehat{y_{r0}} = y_{r0} + s_r^{+*} \qquad r = 1, \ldots, s$$

Through this model, information can be provided about the slacks of each input s_i^{-*} and output s_r^{+*}. As mentioned above, if a DMU has a slack that is not null, it means that that unit is weakly efficient, and hence that value corresponds to the amount of input reduction and output increase feasible, so that unit becomes fully efficient.

Mathematically, the difference between the BCC model and the CRS model can be demonstrated as following: a constraint is withdrawn to the former model, $\sum_{j=1}^{n} \lambda_j = 1$, which implicates the removal of the variable u in the dual problems.

1.3 Scale effects

As previously mentioned, two approaches may be used to evaluate the performance of the DMUs: CRS and VRS.

The differences between a CRS frontier and a VRS frontier are shown in Figure 1. The CRS frontier is represented by a straight line, where DMU_E is the only efficient unit belonging to it. However, in the case of the VRS frontier, a convexity constraint is added to the model, $\sum_{j=1}^{n} \lambda_j = 1$, which leads to the frontier being piece-wise linear and determined by the set of units B, E and C, as shown by Figure 1.

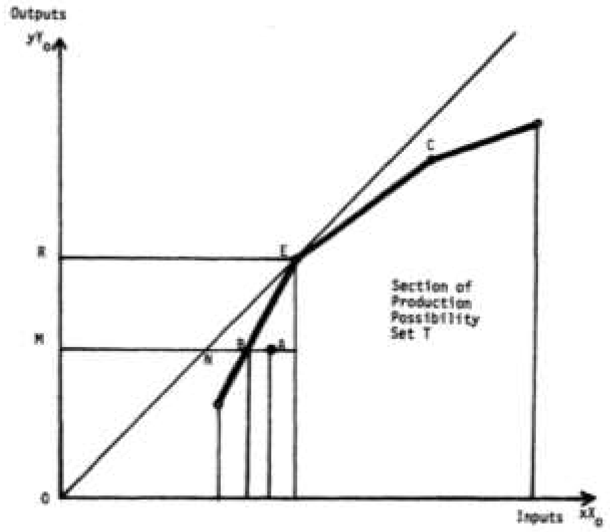

Figure 1 – CRS and VRS frontiers according to (Banker et al., 1984)

When looking at the VRS frontier, DMUs B and C produce their outputs with a different scale, comparatively to DMU E. More concretely, DMU B shows increasing returns to scale, because an increase of an input leads to a proportionally larger increase in the output; unlike, DMU C shows decreasing returns to scale, for the fact that an increase of an input leads to a proportionally smaller increase in the output. So, with the CRS model we estimate a technical and also a scale efficiency, opposing to the VRS model that does not include the scale efficiency. It comes out from Figure 1 that:

CRS efficiency ratio = Technical Efficiency (TE) of DMU A = $\frac{MN}{MA}$.

VRS efficiency ratio = Pure Technical Efficiency (PTE) of DMU A= $\frac{MB}{MA}$.

Scale Efficiency represents the impact of the scale of the operation of a DMU, and is measured by the ratio between the CRS efficiency ratio and the VRS efficiency ratio, $\frac{TE}{PTE}$.

Scale Efficiency (SE) ratio of DMU A = $\frac{CRS}{VRS} = \frac{MN}{MB}$.

According to the expressions above, the sources of technical inefficiency of a DMU can be the result of: an inefficient operation (PTE), an unproductive scale (SE), or both (PTE and SE).

To understand what type of returns to scale characterize the efficient frontier, we can compute scale effects. If the DMU shows evidence of scale effects, then the wise choice is to proceed with VRS, otherwise CRS should be pursued (Dyson et al., 2001).

2. Tobit Regression Model

In order to analyse the drivers that impact the performance of a firm, the most common model used is a regression model. However, as stated by (Ko et al., 2017), a general regression model as the Ordinary Least Squares (OLS), is not able to analyse the factors that impact efficiency measured by DEA. In fact, with the efficiency value limited to a range between 0 and 1 the OLS could lead to biased estimates or invalid inferences. This is corroborated by the fact that the values estimated by OLS regression models could assume values inferior to 0 or superior to 1, thus out of the limited range that characterize the efficiency scores [0, 1]. Tobit Regression Model (Tobin, 1958) can overcome the above limitation, since it accounts for the possibility of truncated dependent variables.

(Coelli et al., 2005) proposed the identification of the factors that impact and are determinant to the DMUs efficiency, using a two stage approach. In a first step, the efficiency scores of the DMUs are obtained through a DEA model and, in a second step, the most impactful variables in those scores are found by using a Tobit Regression Model. Here, the dependent variable is characterized by the efficiency scores and the independent variables are the ones that could influence the dependent variable.

Many studies point out that the variables used in the DEA model are inherently dependent on the efficiency scores. Hence, the estimates developed in this second-stage will be biased and inconsistent (Yu & Ramanathan, 2008). In order to overcome this problem, authors like (Simar & Wilson, 2007) and (Coelli et al., 2005) suggest the application of a bootstrapping technique allowing best practices to improve performance.

As stated earlier, this model is useful when the dependent variable of this model is restricted to a certain range of values. The model is represented as:

$$y_i^* = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + u_i$$
$$if\ y_i^* \leq 0, then\ y_i = 0$$
$$if\ y_i^* \geq 1, then\ y_i = 1$$
$$if\ 0 < y_i^* < 1, then\ y_i = y_i^*$$

Where:

- y_i is the efficiency score of the DMU_ithat is calculated by DEA,
- x_{1i} , x_{2i} , ... are the values of independent variables of the DMU_i,
- β_1, β_2, ... are the coefficients of the independent variables, which will determine the expected effect of the independent variables on efficiency.

This method is very important to identify the drivers that have influence on the efficiency score (y_i), that it is our dependent variable, and their impact. The results of this model provide information about two important parameters. The p-value and the R^2, as it will be shown in practice in Chapter 7. The **p-value** of the independent variables represents their statistical significance: if a variable is statistically significant (p-value < level of significance), means that the variable has impact on efficiency. The $\boldsymbol{R^2}$ represents the percentage of the efficiency variability caused by the model, more concretely, through the independent variables that compose the regression model.

Chapter 4 - Literature Review

1. Studies on retail performance

DEA has been used in many application domains, one of which is retail. In this chapter, the methodologies used to analyse retail performance will be reviewed.

(Almohri et al., 2018) identify the drivers that impact the performance of an automotive dealership in comparison with similar dealerships and propose to rely on optimization techniques as to offer dealerships tailored recommendations. To achieve such objective, three techniques were applied. First, a clustering technique with filtered data from internal and external sources was used in order to obtain a group of similar stores where performance evaluation could be estimated more accurately. Second, an effective Finite Mixture of Regressions technique was used to undertake model-based clustering and to model store performance. This technique was based on competitive learning and clustering stores into a number of homogeneous groups, allowing effective performance modelling and the identification of practices to be used as benchmarks for less efficient stores. Finally, an optimization model was used to tailor recommendations for individual stores in the clusters, as to enable simultaneous improvement of store profitability and sales.

(Vyt & Cliquet, 2017) use many approaches to develop store performance standards, in order to fairly distribute rewards for managers as a function of the performances they achieve. First, they rank the stores in two ways: a store ranking by retailers and by using a two-step DEA method; second, they determine the so-called store's profile. In a first step, the retailers are clustered along the retail chain using two criteria: location (stores are classified as belonging to a rural or a urban area) and sales area (stores are classified as having more or less than 1500m^2). Then, another variable (sales, in this case) was used to compare the stores within each cluster. As to allow for comparative analysis, efficiencies for every store and for each cluster have been estimated. Finally, a two-step DEA model was used: initially, a DEA model assessed the individual efficiency of stores; then an Ordinary Least Squares (OLS) regression was used to know which variables impact on efficiency scores.

Likewise this last study, many authors adopt a two-stage procedure to evaluate performance. (Barros, 2006), (Vyt, 2008) and (Xavier et al., 2015), measured retailers efficiency in Portuguese hypermarkets and supermarkets, in a French supermarket retail chain and in a Portuguese fast fashion retailing sector, respectively. All have adopted a DEA model to evaluate the relative efficiency of each store and regression models to find out which variables impact on the efficiency score extracted by the previous stage. However, the first author used a Tobit Regression Model, the second author used OLS regression, and the third used a quantile regression technique. One should note that they have all used a bootstrap technique, because the efficient scores are correlated to the explanatory variables of the regression model, in order to avoid biased and inconsistent results.

In (Gandhi & Shankar, 2016), the authors wanted to know the current level of performance of Indian retailers and to find out how they could plan and improve their operations and profitability in terms of the company, the store, the merchandise category, or even ate the sku (Stock Keeping Unit) level. For such purpose, they have used two methods: Strategic Resource Management (SRM) and DEA. First, two generic Indian retailers (Shoppers Stop and Trent) were compared using the SRM model, and then they have been benchmarked with a greater retailer (Walmart). In order to validate these results, a DEA model has been used to assess the efficiency of 11 retailers, including the two considered in the previous analysis.

(Yu & Ramanathan, 2008) and later (Yu & Ramanathan, 2009) studied the performance of firms in the retail sector. The first study has been undertaken in the UK and deals with different types of business, like food retailing, home appliances retailing and department store retailing; the second case has targeted retailers in

China. In both cases, the same three methodologies have been used: Data Envelopment Analysis (DEA), with the objective of measuring the technical and scale efficiency of the retailers, Malmquist productivity index (MPI), analysing changes in the patterns of efficiency during the period in question, and a bootstrapped Tobit Regression Model, that has the power to calculate the impact of environmental variables on the efficiency levels measured by the DEA model. The studies reached the conclusions that only ten out of 41 UK retailers are technically efficient and only seven out of 61 Chinese retailers are technically efficient; another conclusion was that 50% of the UK retail firms and 37.3% of the Chinese retail firms have registered progress in terms of MPI during the period under analysis; and finally, that three out of five environmental factors have significant influence in the UK retail efficiency while among the Chinese retail firms only one out of five environmental factors has that impact.

2. DEA applied to the benchmark of retail stores

Benchmarking is a powerful method to study competition between firms in the retail sector, with DEA being instrumental to analyse the efficiency of comparable firms (Yu & Ramanathan, 2008).

This section will review a set of studies that have used this method, identifying which type of data and which models were used to fulfil the objective of assessing the relative efficiency of a store.

(Barros & Alves, 2003) and (Barros, 2006) studied the efficiency of Portuguese supermarkets and hypermarkets. In the first case, 47 retail stores were analysed using a cross-section data for the year 2000; the second study targets 22 retail stores using panel data between 1998 and 2003, with 132 observations being considered. Both studies consider an output-oriented model, with a Variable Return-to-Scale (VRS) hypothesis, for the fact that the scale size is controllable by the retail chain’s central management. For comparative terms, a Constant Return-to Scale (CRS) hypothesis was used and the ratio between these efficiency scores, measured by VRS and CRS, has provided the Scale Efficiency measure. The results have shown that scale economies are a dominant source of inefficiency.

(Vyt, 2008) measured the relative efficiency of 38 stores of a French supermarket retail chain. In this case, three DEA models were used. In all models, an output-oriented model under VRS was used. In the first model, the exogenous variables were considered as fixed. In the second model, all inputs were allowed to vary. The third model, consisted on a two-step approach, with the first step using DEA using with store variables only, and the second step an Ordinary Least Square (OLS) regression. The results have shown that the first two models have poor discriminant power, because all the stores were classified as efficient, regardless of the fact that, in the second model, one store was considered as inefficient. With the third model, only 15 supermarkets were assessed as efficient.

(Perrigot & Barros, 2008) have also assessed the technical efficiency of French retailers. In this study, data from 11 stores with a panel data between 2000 and 2004 grouping 55 observations has been used. A output-oriented model with a VRS hypothesis was considered with the objective of maximizing their production. They have also used a CRS hypothesis and the conclusion was that both hypotheses lead to high average efficiency scores, which means that the dominant source of efficiency is scale, contrary to the conclusions of the previous studies. Apart from these two models, two more have been considered: cross-efficiency and super-efficiency DEA models, allowing to discriminate between the efficiency units given by CRS and VRS models.

Measuring stores performance is not only about assessing supermarket and hypermarket stores, but it can be done with other retail industries. For example, (Xavier et al., 2015) have used data from 40 clothing stores, both from winter and summer collections, between 2010 and 2013. An input-oriented model was chosen, because of the fact that retailers had relatively less control on the outputs. Both CRS and VRS models were considered and the results have shown that there were not problems of scale, because scale efficiency was higher than pure technical efficiency.

(Lau, 2013) has targeted a major Australian retailer, assessing six stores in the retailer's network in the year 2009. An input oriented model was used, with both VRS and CRS hypotheses to compare store performance. The results show that in both hypotheses the same efficient store is elected, and that low efficiency scores are due mainly to scale inefficiency. A slack analysis is also undertaken with the purpose of measuring the target input and output values of the stores.

(Thomas et al., 1998) measure the performance of 500 domestic outlets selling "moderately priced home furnishings and household items". Two procedures have been used to measure the performance of those stores. Initially, an output-oriented model was used with two outputs and 16 inputs, discriminating the inputs that store managers have under control from those they do not control at all. The conclusion was that seven of the stores are fully efficient, being positioned on the efficient frontier. Additionally, the authors wanted to identify the critical success factors (CSF) and, for such matter, they proceeded with a multivariate analysis of variance (MANOVA). For that purpose, the inputs were measured based on the efficiency level determined by the DEA model and subsequently divided into quartiles using that criteria. Then, the differences between the outputs in all of the quartiles were evaluated. The results of this test showed that there were differences in both outputs across the quartiles ($F = 80.41$, $\rho > 0.001$). In brief, the results of the MANOVA suggested that there were specific factors associated with high store efficiency.

As stated previously (Yu & Ramanathan, 2008) and later (Yu & Ramanathan, 2009), used DEA to evaluate the performance of 41 UK and 61 Chinese retailers, respectively, both using a panel data. They chose to calculate VRS, CRS and Scale efficiencies separately, using an output-oriented model. The results point out that in both approaches, the firms that are found to be fully efficient with the CRS model, are also completely efficient with the VRS model, which means that those retailers are fully efficient in a scale efficient level, though in the later study the number of efficient firms is higher. This means that these firms generate the maximum outputs for a given set of inputs, unlike the inefficient ones that were not able to do it.

2.1 Factors that influence performance

The variables selected as inputs and outputs, when applying DEA, and the explanatory variables considered in the Tobit Regression Model are of utmost importance. In both methods, variable selection should reflect the firm's goals, objectives, key performance indicators (KPIs) and sales, and also take into consideration the variables used in previous studies.

According to the literature, the more commonly used variables as outputs are: **"Sales"** ((Vyt, 2008); (Vyt & Cliquet, 2017); ((Barros, 2006); (Yu & Ramanathan, 2008); (Yu & Ramanathan, 2009); (Xavier et al., 2015); (Gandhi & Shankar, 2016); (Uyar et al., 2013); (Barros & Alves, 2003); (Yu & Ramanathan, 2009); (Ko et al., 2017); (Gandhi & Shankar, 2014); (Moreno & Sanz-Triguero, 2011); (Donthu & Yoo, 1998); (Perrigot & Barros, 2008); (Thomas et al., 1998)), **"EBITDA"** ((Barros, 2006); (Yu & Ramanathan, 2008); (Yu & Ramanathan, 2009); (Xavier et al., 2015); (Uyar et al., 2013); (Barros & Alves, 2003); (Moreno & Sanz-Triguero, 2011)), and **"Profit"** ((Uyar et al., 2013); (Gandhi & Shankar, 2014); (Perrigot & Barros, 2008); (Thomas et al., 1998)).

When choosing which inputs should be applied, the literature distinguishes controllable (also known as discretionary) from uncontrollable (also known as non-discretionary) variables. Controllable inputs are those that each of the DMUs has under control, unlike the uncontrollable inputs, which are dependent on central management or other external actors. Thus, the authors end up choosing the first ones, because it is through them that DMUs can gain competitive advantage: **"Number of employees"** ((Barros, 2006); (Yu & Ramanathan, 2008); (Yu & Ramanathan, 2009); (Gandhi & Shankar, 2016); (Uyar et al., 2013); (Barros & Alves, 2003); (Ko et al., 2017); (Perrigot & Barros, 2008); (Vyt, 2008); (Vyt & Cliquet, 2017); (Thomas et al., 1998); (Moreno & Sanz-Triguero, 2011); (Thomas et al., 1998)), **“Sales area”** ((Yu & Ramanathan, 2009); (Gandhi & Shankar, 2016); (Uyar et al., 2013); (Pestana Barros & Alves, 2003); (Ko et al., 2017); (Vyt, 2008); (Vyt & Cliquet, 2017); (Ko et al., 2017); (Donthu & Yoo, 1998)), **“Number of stores”** ((Barros, 2006); (Gandhi & Shankar, 2014)), “**Assets”** ((Barros, 2006); (Yu & Ramanathan, 2008); (Xavier et al., 2015); (Perrigot & Barros, 2008); (Gandhi & Shankar, 2014); (Moreno & Sanz-Triguero, 2011)), **“Salaries”** ((Xavier et al., 2015); (Uyar et al., 2013); (Barros & Alves, 2003); (Gandhi & Shankar, 2014); (Thomas et al., 1998); (Moreno & Sanz-Triguero, 2011); (Thomas et al., 1998)), **“Rental cost”** ((Xavier et al., 2015); (Ko et al., 2017); (Vyt, 2008); (Vyt & Cliquet, 2017); (Thomas et al., 1998)), and **“Inventory”** ((Gandhi & Shankar, 2016); (Uyar et al., 2013); (Thomas et al., 1998); (Moreno & Sanz-Triguero, 2011) (Barros & Alves, 2003)).

Although many authors put aside the uncontrollable inputs (also known as environmental variables), some authors consider them, since they may strongly affect the stores’ performance. Examples of those variables are: **“Location”** ((Barros, 2006); (Yu & Ramanathan, 2008); (Yu & Ramanathan, 2009); (Donthu & Yoo, 1998)), **“Retail characteristics”** ((Yu & Ramanathan, 2008); (Yu & Ramanathan, 2009)), **“Competitors within the trade area”** ((Ko et al., 2017); (Vyt, 2008); (Vyt & Cliquet, 2017))**, “Ownership”** ((Barros, 2006); (Yu & Ramanathan, 2008); (Yu & Ramanathan, 2009); (Gandhi & Shankar, 2014)), **“Store age”** ((Uyar et al., 2013); (Gandhi & Shankar, 2014); (Ko et al., 2017); (Barros & Alves, 2003); (Thomas et al., 1998); (Yu & Ramanathan, 2008); (Yu & Ramanathan, 2009); (Thomas et al., 1998)), **“Population within the trade area”** ((Ko et al., 2017); (Vyt, 2008); (Vyt & Cliquet, 2017); (Thomas et al., 1998); (Uyar et al., 2013)), and **“Potential Market”** ((Vyt, 2008); (Vyt & Cliquet, 2017)).

Chapter 5 – Description of the store set and dataset

The aim of this chapter is to present and characterize the stores targeted in this study, and also to describe the variables in the dataset. This step of the work is very important since we needed to capture in detail the reality of the DeBorla retail stores that were the subject of the project. The period of analysis is from January 2015 to December 2017.

1. Contextualization and description of the stores in the dataset

From a total of 34 DeBorla stores, only 27 were considered in the dataset. The exclusion of seven stores was based in two criteria – location and time – since five of the stores are not located in Continental Portugal, and the other two only opened after the end of the period of analysis.

All stores in the sample are in what can be called a similar location, in the suburbs of urban centres. The 27 stores present an average sales area of 1645.1 m^2, ranging from the smallest one with 823.8 m^2 to the largest with 2946.8 m^2.

During the period of analysis, the stores remained unchanged in what regards the majority of their structural characteristics, like Retail, Store Visibility, Parking Space and Area.

2. Variables that compose the dataset

The dataset used is built by 18 variables, organized into six categories, classified as shown in Table 1:

Geography	Demography	Competition	Structure	Management	Results
District	Population within trade area	Number of "Gato Preto" stores within trade area	Age	General Store Costs	Sales
Municipality		Number of "Area" stores within trade area	Area	Salaries	EBITDA
		Number of "Espaço Casa" stores within trade area	Store Visibility	Rent	
		Number of "CASA" stores within trade area	Parking Space		
		Distance to nearest competitor	Retail		

Table 1 - Categorization of variables of the dataset

Following to the categorization of the variables, it is important to make a detailed description of each of them, which is presented in Table 2. This table portrays the meaning of each variable, the type (discrete numeric variable, continuous numeric variable, nominal categorical variable, ordinal categorical variable), the units of measurement and finally some "Notes", whenever an additional explanation is found helpful.

Variable	Description	Type	Unit	Notes
Population within trade area	Number of residents in a radius of x km	Numeric - Discrete	Millions of people	[1]
Retail	1 – the store is located in a retail space; 0 – stand alone store	Categorical - Nominal	Number	
Age	Store age (last period of analysis 31-12-2017)	Numeric - Continuous	Years	
Sales	Total value of store sales, in the period January 2015 – December 2017	Numeric - Continuous	Euros	
General store costs	Total value of store general costs, in the period January 2015 – December 2017	Numeric - Continuous	Euros	
Salaries	Total value of store personnel costs, in the period January 2015 – December 2017	Numeric - Continuous	Euros	
Area	Average sales area, in the period January 2015 – December 2017	Numeric - Continuous	m^2	
EBITDA	Total value of store EBITDA, in the period January 2015 – December 2017	Numeric - Continuous	Euros	
Rent	Total value of store rent, , in the period January 2015 – December 2017	Numeric - Continuous	Euros	
Number of "Gato Preto" stores within trade area	Number of Gato Preto stores in a radius of x km	Numeric - Discrete	Number	[2]
Number of "Area" stores within trade area	Number of Area stores in a radius of x km	Numeric - Discrete	Number	[2]
Number of "Espaço Casa" stores within trade area	Number of Espaço Casa stores in a radius of x km	Numeric - Discrete	Number	[2]
Number of "CASA" stores within trade area	Number of CASA stores in a radius of x km	Numeric - Discrete	Number	[2]
Store Visibility	Indicates if the store has visibility, assuming the following values: 2) when the store is near a main road (having significant traffic during the week) and DeBorla brand is clearly identified by who is passing near the store; 1) if one of the previous conditions is satisfied; 0) if none of the above conditions is verified.	Categorical - Ordinal	Number	
Parking Space	Indicates if the store has parking space, assuming the following values: 0) when the store has no private parking but there is parking nearby; 1) when the store has its own park, but there may be parking difficulties; 2) when the store has a private parking.	Categorical - Ordinal	Number	
District	Name of the district where the store is located	Categorical - Nominal	n.a	
Municipality	Name of the municipality where the store is located	Categorical - Nominal	n.a	
Distance to the nearest competitor	Distance of each DeBorla store to the closest competitor	Numeric - Continuous	km	[3]

Table 2 - Detail of the variables that compose the dataset

From the Table above, the dataset is composed by 13 numerical variables (five discrete and eight continuous), and by five categorical variables (two ordinal and three nominal).

Assuming that most of the variables are clearly defined in Table 2, further analysis is now presented in the cases signalled in the Notes column.

[1] In order to determine the "population within trade area", an open-source Geographic Information System (QGIS) software was used. The following steps were followed:

1. The coordinates (latitude, longitude) of each DeBorla store where inserted (Figure 2);
2. A "buffer" was created for each DeBorla store, as a surrounding area with a radius of 34 km (Figure 2), setting a demarcated area of influence;
3. The number of total residents[3] within the "buffer" around of each DeBorla store (see Figure 3, Table 20 in Appendix B.1), was estimated.

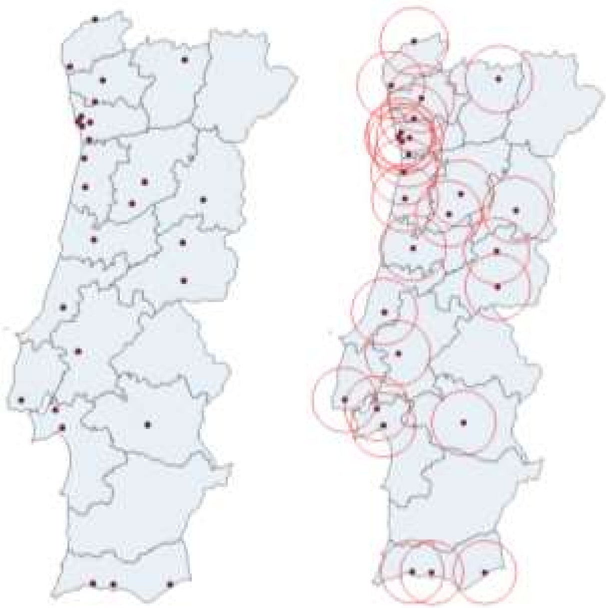

Figure 2 - Locations and respective surrounding areas for each DeBorla store

[3] Number of total residents: this information was provided by Census 2011, the last with the available compatible format data for QGIS.

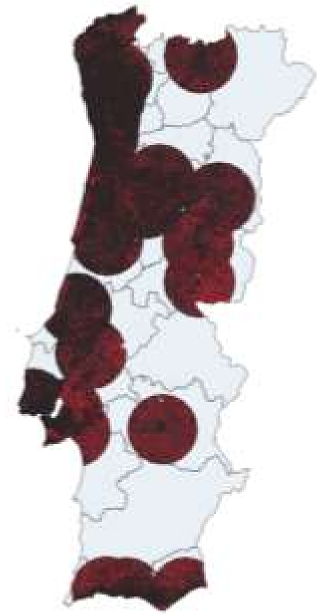

Figure 3 - Population within the trade area

The radius of 34 km of the "buffers" was derived from an estimate on how far DeBorla customers are willing to drive in order to reach each store. To answer this question the QGIS software was used, to estimate the surrounding area, in a process encompassing three different steps:

1. The coordinates (latitude, longitude) of the address of each DeBorla customer were inserted in the software, this information being provided by the firm (Figure 4);
2. The distance between the address of each customer and the location of DeBorla stores was measured. In order to allow that, the firm provided information on the stores that each customer goes to. The distance traveled by each customer to the respective store could then be calculated;
3. The average distance travelled by the customers to each DeBorla store was finally calculated (see Table 22 in Appendix B.2)[4].

[4] All data was made available by DeBorla, both regarding customers and stores, in an anonymized format, therefore complying with the EU General Data Protection Regulation (GDPR) and RGPD in Portugal.

Figure 4 - Locations of DeBorla stores and their customers

As shown by Table 21, three of the stores (DMU 1, DMU 13 and DMU 24) are outliers, because of the unusually large average distances travelled by their customers. All these stores are located in Algarve, south of Portugal, and the explanation for being outliers is that their customers take the advantage of the fact that they are on vacation to buy in those DeBorla stores. For that reason, these outliers will not be taken into consideration into the determination of the average distance travelled by customers. As we can see in Table 3, the average travelled distance for all DeBorla stores, excluding the outliers (DMU 1, DMU 13 and DMU 24), is quite similar to the median travelled distance for all stores, including the outliers.

Average distance (excluding DMU 1, DMU 13, DMU 24)	41 km
Median distance (all stores)	34 km

Table 3 - Comparison between average distance and median distance travelled by customers

So, the radius of the buffers was calculated using the median of the distances travelled by the customers to access the stores.

[3] To determine the "number of each of the competitors within trade area", the QGIS software was used, going through the following steps:

1. The coordinates (latitude, longitude) of all stores, DeBorla and different competitors were inserted (Figure 5);
2. The number of each of the competitors within each buffer of each DeBorla store was calculated (Table 22, Appendix B.3).

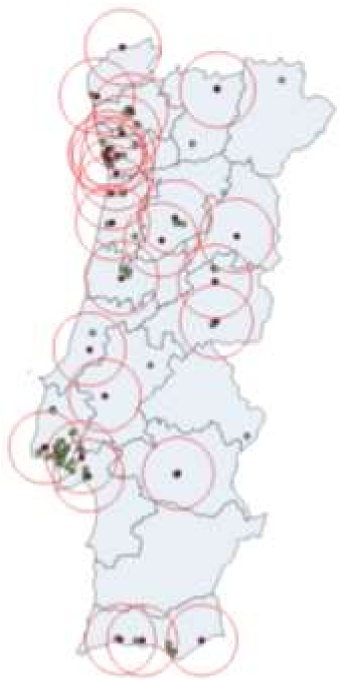

Figure 5 - Location of DeBorla stores (red) and competitors stores (green)

[4] To determine the "distance to the nearest competitor", the QGIS software was used, by:

1. Inserting the coordinates (latitude, longitude) of all stores, DeBorla and competitors (Figure 5);
2. Calculating a distance matrix for the nearest competitors of each DeBorla stores and the respective distance (Table 23, Appendix B.4).

Chapter 6 - Store Performance Evaluation

The aim of this chapter is to assess store performance in DeBorla group, by determining their efficiency scores through a DEA model, as well as to set improvement goals to the stores[5]. The model is expected to be able to evaluate the capacity of the stores to generate output based on the existing inputs that contribute the most to the profitability of the stores.

In the present analysis, 27 stores are considered with panel data of three consecutive years, with a total of 81 observations. In fact, only 71 observations are being analysed, because some of the DMUs were not operating along all those years, namely the stores that opened in 2015 and 2016.

Regarding the DEA model used, both Constant Return to Scale (CRS) and Variable Return to Scale (VRS) models have been considered, in order to analyse different types of efficiency: technical efficiency (TE), pure technical efficiency (PTE) and scale efficiency (SE). Further than that, it was considered of major relevance to know which returns to scale characterize each of the stores, as well as to identify the existence of scale effects by using a non-parametric test. An input-oriented DEA model was used, because store managers have clearly more control on their inputs than on their outputs.

---

[5] The DEA model was run with the Efficiency Measurement System (EMS) software.

1. Selection of variables

The selection of the variables is a very important step in the development of the DEA model. This was done, taking into consideration an extensive literature review[6] and also the objectives that were set by the firm.

In this process we first considered only discretionary variables, those that the store management can have some control upon, and excluded contextual variables related to the market and the competition.

Table 4 shows a correlation matrix[7] between the discretionary variables chosen to be included in the model. Through this matrix we are able to infer that there is a strong positive correlation between the variables selected as inputs (General Store Costs, Salaries, Area and Rent) and the variables selected as outputs (Sales and EBITDA). This means that these output variables depend positively on the input variables contributing to their growth. However, both outputs have a strong correlation, which makes sense because Sales is a very important variable that contributes positively to EBITDA. We have therefore decided to opt for only one of these variables, and Sales was chosen because it is the output exhibiting the strongest relation with the other variables. A decision was also made of not including Rent in the model, even though it shows a strong correlation with the output Sales. This is because some of the stores belong to the company and do not pay a Rent and if this variable were to be included in the model it would distort the results.

---

[6] Chapter 4 and Appendix A, put forward evidence that the variables used in the present study are fully corroborated by the literature.

[7] The correlation matrix was constructed using software R, a free software package for statistical computing, specifically the Pearson Correlation analysis that studies the correlation or linear dependence between two variables. The results belong to an interval between -1 and 1 (-1 means that the variables have a perfect negative correlation; 0 means that the variables do not depend linearly from to one another; 1 means that the variables have a perfect positive correlation).

Since the variables to be included in the model are those that contribute the most to store profitability, the discretionary output included in the model was **Sales** and the discretionary inputs were **General Store Costs, Salaries and Area**.

	General Store Costs	*Salaries*	*Area*	*Rent*	*EBITDA*
General Store Costs	1.000				
Salaries	0.786	1.000			
Area	0.641	0.697	1.000		
Rent	0.642	0.657	0.418	1.000	
EBITDA	0.484	0.560	0.339	0.202	1.000
Sales	**0.801**	**0.875**	**0.587**	**0.704**	**0.801**

Table 4 - Correlation matrix supporting the selection of discretionary variables

Table 5 shows the mean and the median of the discretionary variables which are used in the DEA model for all stores. **Sales** represents the total value of the items sold, expressed in euros, between January 2015 and December 2017; **General Store Costs** and the **Salaries** represent the total value of the respective charges, expressed in euros, for the same period; and **Area** represents the store area used for the sale of the products, expressed in m^2, during the same period.

Variables	*Mean*	*Median*
Sales	1 522 576	1 585 963
General Store Costs	61 950	54 985
Salaries	182 233	189 878
Area	1 645	1 740

Table 5 - Mean and median of the discretionary variables of the DEA model

2. Identification of the Return to Scale type

In order to identify the Return to Scale type it was important to analyse the possibility of the existence of scale effects. As previously mentioned, CRS efficiency is measured by the product of PTE and SE and, on the other and, VRS efficiency is measured only by PTE. Table 24 in Appendix C.1 presents the efficiency scores with CRS and VRS for each of the 71 observations and the respective scale efficiency. A non-parametric test was realized on this data, the **Wilcoxon signed-rank test**[8], that compares the distribution of two different samples, in order to find out if both samples came from populations with identical distributions or not. This test provides two hypotheses: the null hypothesis (H0) assumes that both efficiency distributions are the same; the alternative hypotheses (H1) assumes that both efficiency distributions are different. The results of the test have shown evidence that H0 should be rejected, because the p-value was approximately zero. This means that scale efficiency is significant and therefore DeBorla stores are characterized by Variable Return to Scale (VRS).

3. Identification of the inefficiency sources

At this point, it was necessary to know which inefficiency sources characterized each store. As mentioned in Chapter 3, the sources of technical inefficiency could be: an inefficient operation, an unproductive scale or both. Table 24 Appendix C.1 presents the technical inefficiency sources for each store, as measured by the comparison between PTE and SE: if SE is greater than PTE this means that the store has an inefficient operation; on the other hand, if PTE is greater that SE this means that the store has an unproductive scale. When PTE equals SE they both have a value

[8] The Wilcoxon signed-rank test was run using the R software.

of 100%, meaning that there are no technical inefficiency sources. Table 6 shows the segregation of the store observations according to the technical inefficiency source: out of 71 observations, 46 are characterized by having inefficient operations and 21 unproductive scales, with four being efficient. It also comes out that the prevailing technical inefficiency source is Inefficient Operation.

Efficient	Inefficient operations	Unproductive scale	Total
4 (6%)	46 (65%)	21 (30%)	71 (100%)

Table 6 - Technical inefficiency sources for all observations

4. Efficiency Results assuming VRS without environmental variables

In the present section, the DEA results are presented and analysed, considering only discretionary variables. According to the conclusions of the previous section, the following results and respective interpretation are based on the assumption of Variable Return to Scale.

Figure 6 presents the store efficiency distribution obtained applying the VRS approach, using the 71 observations. Table 7 shows summarized statistics of the 71 stores efficiency observations of this model: the mean value for of all stores and for the inefficient ones, the median for all the stores, the lowest observed efficiency score, the standard deviation and the number of efficient stores.

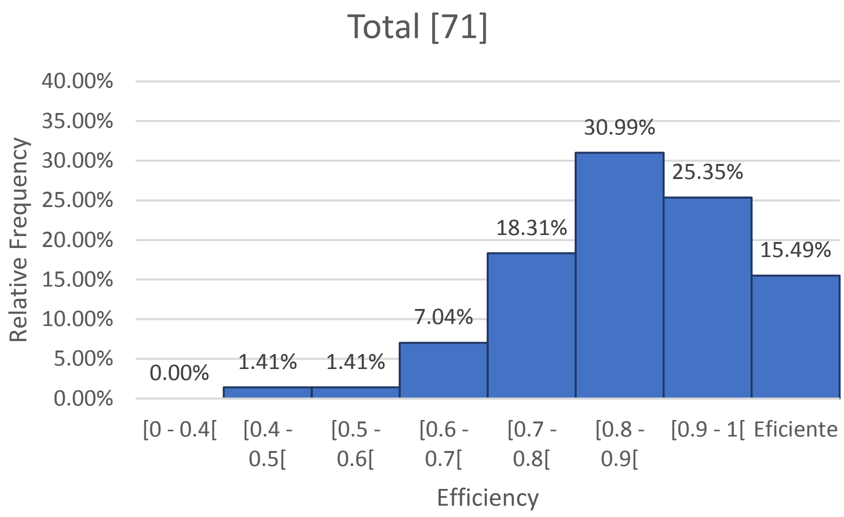

Figure 6 – Store efficiency distribution with VRS

Mean [all stores]	**Mean [inefficient stores]**	**Median [all stores]**	**Minimum**	**Standard deviation**	**Number of efficient stores**
85.92%	83.33%	87.28%	48.31%	0.1175247	11

Table 7 - Summarized Statistics of Store Efficiency observations

According to these results, there is a large number of observations whose efficiency score is very high, and a small number that have a small efficiency score. In fact, most observations (approximately 31%) reach 80% to 90% efficiency and about 15% of them are fully efficient. Additionally, Figure 6 shows that this sample portrays a negative asymmetry, which is corroborated by the fact that the mean is inferior to the median (85.92% < 87.28%).

Table 24, column 2, in Appendix C.1, presents the results of the efficiency scores obtained for the 71 observations assuming VRS. In a total of 71 stores, 60 are inefficient with an average efficiency of 83.33%, which means that those stores could potentially decrease their inputs by an average of 16.67%, in order to become fully efficient stores.

5. Comparative analysis of individual store performance

In order to compare the efficiency of the 27 stores, Figure 7 presents a ranking of the stores efficiency. Since the efficiency measured by DEA without environmental variables is related to 71 observations, the bar graph shows the average efficiency of each store over the period in analysis (2015-2017). The store with the higher average efficiency was DMU 26, a fully efficient store during the period, the only one with an average score of 100%. There is a very large discrepancy between most stores and DMU 5, since this one was only 48% efficient, on average. However, since this store opened only in 2017 this low average efficiency is likely to be a result of a period of analysis equal to the initial launching period of the store. The same interpretation may be considered for DMU 27. This store presented one of the lowest performance levels and also opened in 2017. In that sense, it is somewhat surprising that DMU 6 managed to reach the middle of the ranking since, likewise DMU 5 and DMU 27, only opened in 2017.

The stores exhibiting a high level of performance (above 90% average efficiency) were DMU 26, DMU 1, DMU 20, DMU 7, DMU 16, DMU 21, DMU 18, DMU 11, DMU 12 and DMU 23. All 27 stores have achieved at least an average efficiency above 50%, except for DMU 5.

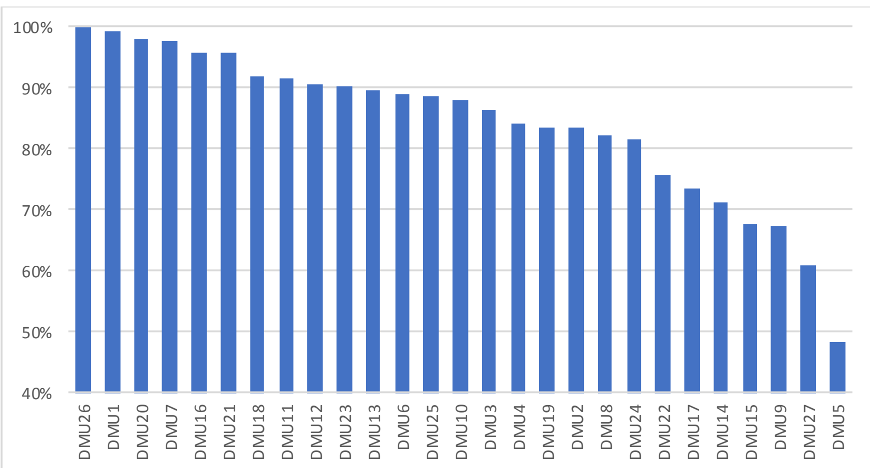

Figure 7 -Average efficiency ranking of all the stores analysed

Table 25, Appendix C.2, presents the yearly and average efficiency values for each store, information supporting the previous and follow analyses in this section.

In order to find out if there is any geographic/location pattern related with the efficiency of DeBorla stores, an analysis of the average efficiency of those stores is made, clustering them by region (Table 8).

Table 32, in Appendix C.6, provide information on the Municipalities, Districts and Regions of Continental Portugal, where each of the 27 stores are located.

Region	Average efficiency	Number of stores
North (w/out Porto Met)	92%	5
Metropolitan Area of Porto	79%	5
Centre (w/out Lisbon Met)	88%	9
Metropolitan Area of Lisbon	80%	3
Alentejo	84%	2
Algarve	90%	3

Table 8 – Average stores efficiency by region

The region showing a higher average efficiency is the North (excluding the Porto Metropolitan Area), with approximately 92%. Both Metropolitan Areas (Porto and Lisbon) show a similar average efficiency (about 79% and 80%, respectively), with the stores located in Porto Metropolitan Area having the lower level of performance of all regions. North and Algarve were the regions showing similar average efficiency (approximately 92% and 90%, respectively). The Centre was the region exhibiting a good performance (approximately 88%), regardless the fact that DMU 5, located in Castelo Branco, was the unit with the worst average efficiency.

Figure 8 shows the average efficiency of the stores in each Municipality[9], as portrayed by a colour code.

[9] Each store is located in a different Municipality, so the average efficiency of the Municipality is the same of the store located in that same Municipality.

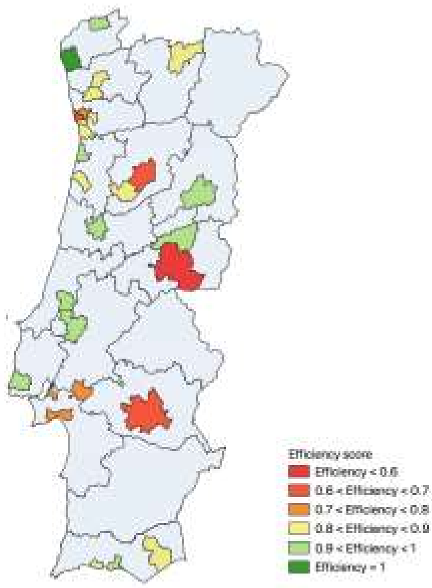

Figure 8 – Location of DeBorla stores according to their average efficiency

The figure does not display any pronounced geographical pattern. It may nevertheless be of great value if analysed taking into account the specific inputs and outputs for each store complemented with specific DeBorla contextual knowledge on the operation of that store. Just as an example, the large difference in performance between stores located in Castelo Branco and Fundão is probably explained by the fact that Fundão is a consolidated operation (12 years old), while Castelo Branco has just started (first year). Another example is the store located in Albufeira, which has top performance (99%) and this may be a result of the specificities of the potential market: high density of real estate related with tourism activities (hotels, villas, restaurants,…), complemented with the local residents and the visitors from other countries and regions.

6. Store performance over time

An important perspective of analysis is to look at the evolution of store performance over time, by looking at the efficiency evolution in the period 2015-2017. Figure 9 shows that the 27 stores achieved the highest performance in 2016 (approximately 89% of average efficiency), with small yearly changes (between 2% and 3%).

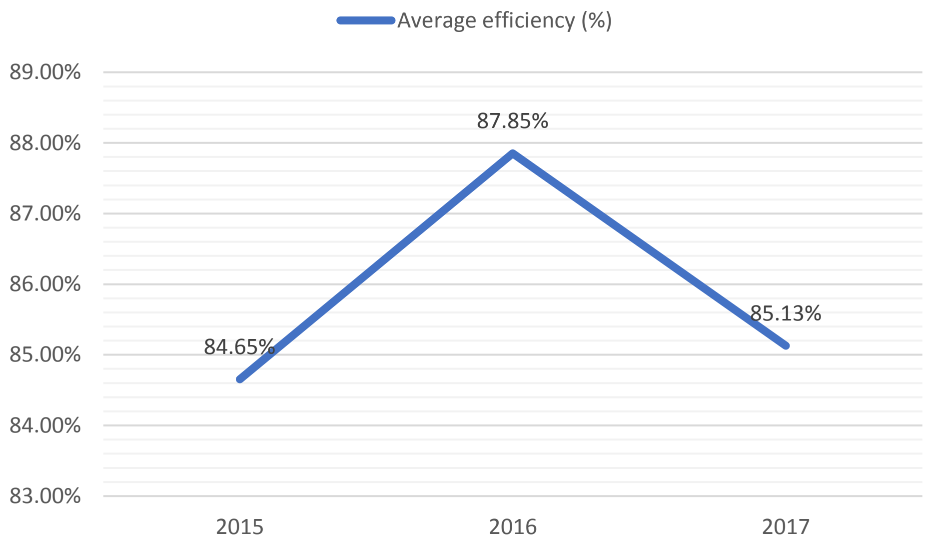

Figure 9 - Average store efficiency between 2015-2017

The efficiency evolution along the period of analysis is available for each store, allowing efficiency variation patterns for selected stores to be analysed. For a better understanding, Table 9 provides information on Efficiency Score, General Store Costs, Salaries and Sales[10] for each of a number of selected stores (DMU 4, DMU 17, DMU18, DMU 25 and DMU 15).

[10] The variable Area was not considered in this analysis, because it is has a constant value over the period of analyis.

	VRS			General Store Costs			Salaries			Sales		
DMU	**2015**	**2016**	**2017**	**2015**	**2016**	**2017**	**2015**	**2016**	**2017**	**2015**	**2016**	**2017**
DMU 4	76.81%	84.2%	91.39%	49 503	46 953	44 566	180 764	198 171	193 057	1 328 532	1 451 828	1 540 676
DMU 17	64.39%	74.2%	81.80%	64 192	58 858	53 817	201 046	198 653	202 330	1 393 095	1 572 119	1 680 254
DMU 18	94.26%	100%	81.20%	45 706	40 134	68 778	189 916	190 593	194 112	1 734 787	1 744 456	1 726 186
DMU 25		100%	77.27%		17 682	47 924		33 332	161 663		240 694	1 023 012
DMU 15		59.66%	76.02%		44 790	127 087		148 185	380 065		796 422	2 604 112

Table 9 –Score efficiency, General Store Costs, Salaries and Sales values for selected stores

DMU 4 and DMU 17 show an overall positive growth of efficiency in the period, of about 19% for DMU 4 and 27% for DMU 17. The cause for this growth was an increase in **Sales**, approximately 15% and 20%, respectively, and a simultaneous decrease in costs, more concretely, in **General Store Costs**, approximately 10% and 17%, respectively. Despite an increase in Salaries in both stores, that increment was not influent enough to have a negative effect on the global performance improvement.

DMU 18 presents an overall negative growth in efficiency of 14%. The store first shows an efficiency increase of 6% (from 94.26% to 100%), becoming fully efficient; then a rapid decrease is observed, of approximately 19% (from 100% to 81.20%). The variable that had a greater effect on this growth was **General Store Costs**, with a 12.19% decrease in the first year and a 71.37% increase in the second year, explaining the abrupt fall in DMU efficiency. Likewise in DMU 4 and DMU 17, there were variables that could not generate a sufficient impact to revert the performance decrease: this is the case of Salaries and Sales.

DMU 25 and DMU 15 only opened in 2016. Their performance evolution between 2016 and 2017 exhibit similar variation rates, but in opposite directions: DMU 25 presents a negative efficiency evolution of approximately 23% (from 100% to

77.27%), while DMU 15 presents a positive efficiency growth of 27% (from 59.66% to 76.02%). DMU 25 presented low values for **Salaries and General Store Costs** and modest **Sales** in 2016, which is perhaps quite normal in the opening year. However, the abrupt increase in General Store Costs and in Salaries in 2017, around 2.7 times higher, were not compensated by the increase in Sales.

The evolution of DMU 15 was in the opposite direction, because it presented an enormous increase in **Sales** (227%), which more than compensated for the increase in costs.

7. Performance improvement goals without environmental variables

It is now timely to analyse the values of the performance improvement goals for the inputs and outputs of each store, as determined by VRS. Those values, also called input and output targets are obtained using the formulae presented in Chapter 3. In Appendix C.4, Table 27 to Table 30 show both the current values and the target values for the 27 stores for each of the years and for each of the input and output variables: General Store Costs, Salaries, Area and Sales, respectively. So, in each of these tables, the inefficient units, that is those with an efficiency inferior to 100%, have different values for the current and the target values: each current value is multiplied by the efficiency score, where then the slack is subtracted. For the efficient units, those where efficiency equals 100%, the target values are obviously equal to the current ones.

Table 10 summarizes the potential improvements for all units analysed. General Store Costs, Salaries and Area have a potential of reduction of their values as large as € 281 564, € 197 270 and 9 571 m^2, respectively, and Sales can potentially increase in a total of € 127 662.

General Store Costs	Salaries	Area	Sales
€ 281 564	€ 197 270	9 571 m^2	€ 127 662

Table 10 - Potential reductions of inputs and potential increases of output (aggregate values for all stores)

Likewise, Figure 10 presents the total potential gains for each input and output of the 71 observations. It is shown that General Store Costs can potentially decrease 21.26%, Salaries 16.01% and Area 23.90% while Sales can potentially increase 0.12%. The corresponding potential economic impact amounts to about € 600 000, without taking into account the savings in Area reductions. While Area is the input with greater potential to reduce, although with a more difficult impact estimation, the total savings resulting from these global potential reductions are quite significant.

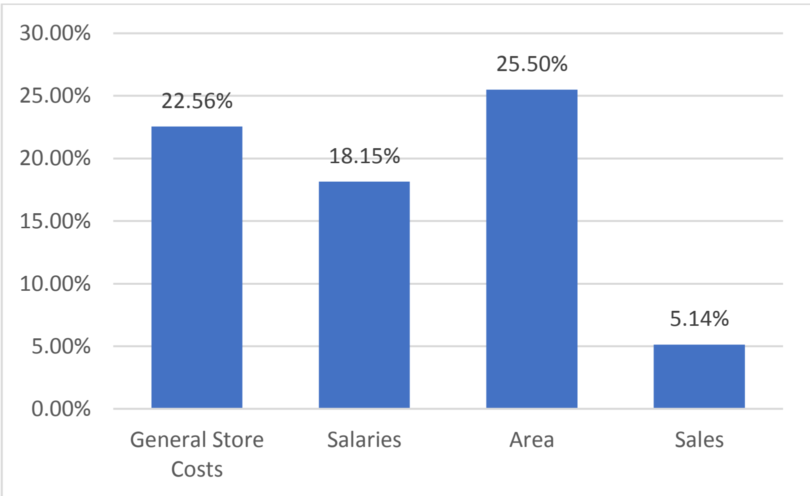

Figure 10 - Potential gain of each variables (aggregate values for all the stores)

Table 31, in Appendix C.5, shows the percentage of potential gains for each of the 27 stores during each year of the period of analysis. Table 26, in Appendix C.3, shows the benchmarks for each of the 27 stores and the specific contribution of each benchmarks (λ) for the efficiency measurement.

While some stores have a balanced potential for improvement in all inputs, other stores exhibit a higher potential for improvement in some specific inputs. That means that these stores' inefficiency is mainly due to the inefficient use of one or more of its inputs. Three specific DMUs will now be analysed.

For example, for **DMU 13 in 2016** the main source of inefficiency is **Salaries**, since in this input the store exhibits a much larger potential for reduction (21%) than in the other inputs (1%). Table 11 presents further information on DMU 13 in 2016 showing the target values for General Store Costs and Area, 1% smaller than the real values, and the target values for Salaries, 21% smaller than the real values. This unit is also compared with its peers (those contributing with more than 30% (λ) for the computation of the target levels). In this case DMU 1 in 2016 is the selected benchmark, that obtained a similar output (0.6% smaller) with 41% lower salaries.

DMU	λ	Efficiency	General Store Costs		Salaries		Area		Sales	
			Current Value	Target	Current Value	Target	Current Value	Target	Current Value	Target
DMU 13 2016		98.95%	€ 45 916	€ 45 434	€ 206 049	€ 163 222	1 278	1 265	€ 1 692 753	€ 1 692 753
DMU 1 2016	0.62	100.00%	€ 44 823		€ 146 015		938		€ 1 591 679	

Table 11 - Efficiency, input and output values for store DMU 13 in 2016 and selected benchmark

Another interesting example is **DMU 27 in 2017**, whose main sources of inefficiency are **General Store Costs** and **Area**. These inputs exhibit a very high potential for reduction, approximately of 58% and 61%, respectively. Table 12, also shows that the target value for Salaries was 39% smaller than the current value and those for General Store Costs and Area were 58% and 61% smaller, respectively, than the current values. This is confirmed when comparing this unit with its peers. According to the table, the benchmark DMU 1 in 2016 presented 30% higher Sales with 50% less General Store Costs, 23% less Salaries and 63% less Area.

DMU	λ	Efficiency	General Store Costs		Salaries		Area		Sales	
			Current Value	Target	Current Value	Target	Current Value	Target	Current Value	Target
DMU 27 2017		60.95%	€ 89 735	€ 37 529	€ 189 878	€ 115 731	2 555	995	€ 1 228 578	€ 1 228 578
DMU 1 2016	0.73	100.00%	€ 44 823		€ 146 015		938		€ 1 591 679	

Table 12 - Efficiency, input and output values for store DMU 27 in 2017 and selected benchmark

DMU 14 in 2016 exhibit a balanced potential for improvement in all inputs (approximately 28%). Table 13, shows that the target values of all inputs were 28% smaller than the real values. According to this table, DMU 1 in 2016 and DMU 18 in 2016, the selected benchmarks, presented a similar magnitude output, but using less resources. For instance, DMU 1 in 2016 attained higher Sales with less inputs, with special relevance for Area (about 54% smaller). However, this is not necessarily observed in all variables for both peers: DMU 18 in 2016 was the peer that attained higher Sales but with a larger Area.

DMU	λ	Efficiency	General Store Costs		Salaries		Area		Sales	
			Current Value	Target	Current Value	Target	Current Value	Target	Current Value	Target
DMU 14 2016		71.79%	€ 50 585	€ 36 315	€ 191 056	€ 137 159	2 034	1 460	€ 1 304 783	€ 1 304 783
DMU 1 2016	0.36	100.00%	€ 44 823		€ 146 015		938		€ 1591 679	
DMU 18 2016	0.33	100.00%	€ 40 134		€ 190 593		2 290		€ 1 744 456	

Table 13 - Efficiency, input and output values for store DMU 14 in 2016 and selected benchmarks

8. Benchmarking highest performing units against lowest performing units

In order to find out if there is a common pattern in the factors that lead to the 100% efficiency level of the peers used as benchmarks, different types of analysis have been made for which specific indicators have been calculated. In the first place, **sales per square meter** of store area was selected as a measure of effective used of space. Then **sales per worker**, a measure of the productivity of labour, was considered. Since we do not have the information on the number of workers per store, a proxy has been used. Assuming that the total monthly cost of an average worker is € 1100 per month (gross salary, social security costs and insurance), the total annual cost would be € 15 400. Dividing the total amount of salaries by € 15 400, a virtual number of store workers may be found and then the sales per virtual store worker calculated. Other calculated indicator was the **percentage of sales (revenues) spent in each store to cover the General Store Costs**, this being in some way related to the weight of externals suppliers and services.

Using these indicators, the direct performance comparison between the highest and the lowest performing units has been undertaken. Figure 11 depicts how the highest and the lowest performing units stand against each other, using a radar graph with normalized values. There is a pattern showing that the top performance units (represented only be the fully efficient units) attained higher sales per square meter and sales per virtual worker, while using lower general store costs per sales. Once again, this shows that the efficient units achieve similar output levels as the most inefficient ones, but using lower inputs.

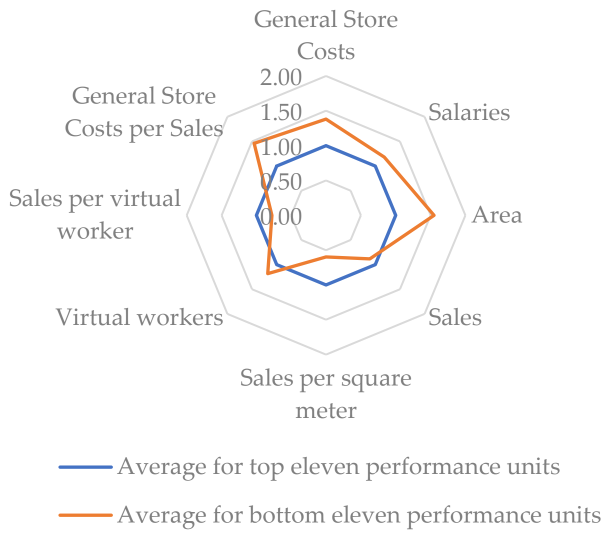

Figure 11 - Radar graph with inputs and outputs of the highest and lowest performing units

The information given by the radar graph is complemented by the one in Table 14.

	General Store Costs	Salaries	Area	Sales	Sales / Area	Virtual workers	Sales / virtual workers	General Store Costs / Sales
Average for top eleven performance units	45 575	148 089	1 286	1 323 679	1 055	10	131 230	3.83%
Average for bottom eleven performance units	62 946	175 371	1 994	1 174 395	630	11	101 896	5.59%

Table 14 - Inputs and outputs average values for the highest and lowest performing units

This chapter closes with an analysis of a different type to benchmark the top performing units against the bottom performing units regarding operational profitability. Since DeBorla does not allocate their financial investments to the individual stores and the investments in the older stores are partial or completely depreciated comparatively to the recent ones, the store profit margin did not appear as an adequate indicator. A good proxy to profit is the operational result, EBITDA margin (EBITDA / sales), assessing DeBorla stores operating profitability.

Table 15 and Table 16 present Sales, EBITDA and EBITDA margin, respectively, for the highest and the lowest performing units.

DMU	Efficiency	Sales	EBITDA	EBITDA margin
DMU1 2015	100%	1 357 836	282 850	21%
DMU1 2016	100%	1 591 679	348 737	22%
DMU12 2017	100%	1 224 897	175 069	14%
DMU16 2016	100%	550 554	17 210	3%
DMU18 2016	100%	1 744 456	303 232	17%
DMU20 2017	100%	1 916 818	403 185	21%
DMU25 2016	100%	240 694	14 505	6%
DMU26 2015	100%	909 900	106 393	12%
DMU26 2016	100%	980 418	131 156	13%
DMU26 2017	100%	1 031 993	147 987	14%
DMU7 2016	100%	3 011 226	478 757	16%
Average for top eleven performance units	**100%**	**1 323 679**	**219 007**	**15%**

Table 15 - Sales, EBITDA and EBITDA margin for the eleven top performance units

Two outliers – DMU 16 in 2016 and DMU 25 in 2016 – stand out in Table 15 for their relatively poor EBITDA margin. Both stores have very small sales, suggesting that the lack of dimension of those two efficient units leads to low EBITDA margins. All the other efficient units have an EBITDA margin above 10%, showing a strong correlation between 100% efficient units and high EBITDA margins. Even considering the two outliers, the average value of the EBITDA margin is 15%, nearly twice that of the inefficient units.

DMU	Efficiency	Sales	EBITDA	EBITDA margin
DMU14 2015	74%	1 197 750	239 586	20%
DMU22 2015	73%	1 564 922	122 207	8%
DMU14 2016	72%	1 304 783	196 809	15%
DMU13 2015	71%	1 445 765	86 200	6%
DMU9 2017	70%	1 167 321	129 652	11%
DMU14 2017	69%	1 355 419	210 719	16%
DMU9 2016	65%	825 728	42 950	5%
DMU17 2015	64%	1 393 095	93 095	7%
DMU27 2017	61%	1 228 578	112 417	9%
DMU15 2016	60%	796 422	31 892	4%
DMU5 2017	48%	638 566	-48 518	-8%
Average for bottom eleven performance units	**66%**	**1 174 395**	**110 637**	**8%**

Table 16 - Sales, EBITDA and EBITDA margin for the eleven bottom performance units

In Table 16, DMU 5 in 2017 stands out as the lowest performing unit and with the lowest negative EBITDA margin (-8%). Additionally, the five less efficient units have an EBITDA margin below 10%. The average value of the EBITDA margin is 8%.

Using this information, a matrix positioning the top performing units and the bottom performing units according to their level of efficiency and EBITDA margin was built (Figure 12).

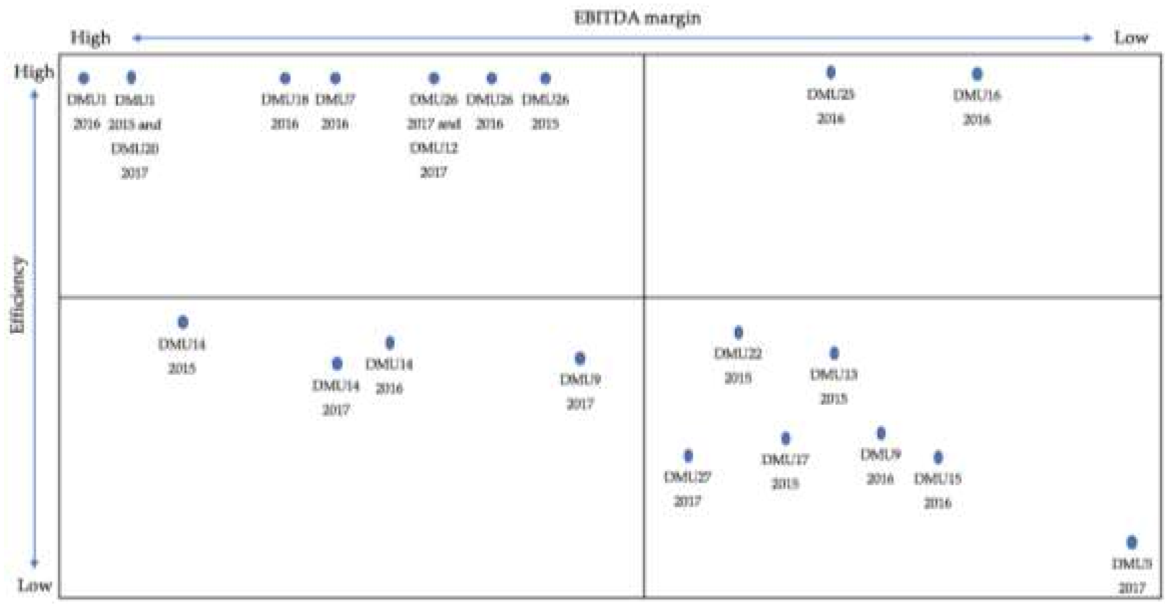

Figure 12 - Matrix positioning the top performance units and bottom performance units

With high efficiency and low EBITDA margin, we have the above referred outliers, DMU 16 in 2016 and DMU 25 in 2016. On the high efficiency/high EBITDA quadrant we may consider one cluster with three units with decreasing EBITDA margin values of 21% and 22% and six units with EBITDA margin values from 17% to 12%. In the low efficiency/low EBITDA margin quadrant, apart from the outlier DMU 5 in 2017 with the lowest efficiency and the lowest EBITDA margin, the units are quite scattered. Finally, the four units in the low efficiency/high EBITDA margin quadrant are positioned quite close to the top.

The normalized radar graph of Figure 13 confirms the conclusions above, showing that the top performance units (represented only be the fully efficient units) attained higher EBITDA and EBITDA margins, with slightly higher values of average sales. This results from the lower operating expenses (Salaries and General Store Costs), that led to a higher profitable operation.

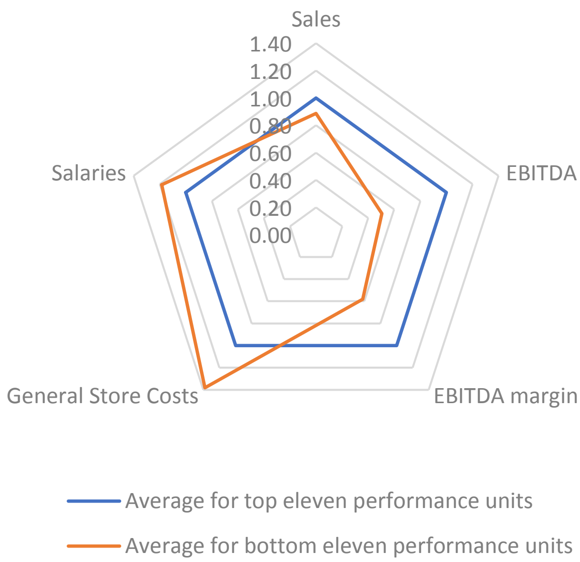

Figure 13 - Radar graph with General Store Costs, Salaries, Sales, EBITDA and EBITDA margin for the highest and lowest performing units

Chapter 7 - Impact of the environmental variables on store efficiency

The main purpose of this Chapter is to understand the impact of the environmental variables on store efficiency. The Tobit regression model (run through the R software) was the tool used in the analysis. The dependent variable of the model is the "Efficiency Score", measured by the DEA model without environmental variables, and the independent variables are: "Population within trade area", "Location", "Store Visibility", "Parking Space"[11], "Age", "Number of "Gato Preto", "Number of Espaço Casa", "Number of Area", "Number of "CASA", and "Distance to the nearest competitor". With this set of variables, a model showing the efficiency variations caused by increments in the independent variables was estimated.

The results of the estimated model are represented in Table 17. The estimated model has an $R^2 \approx 51\%$, meaning that about 51% of the efficiency variation is explained by the independent variables and the remaining 49% of that variation is explained by other factors.

[11] Store Visibility and Parking Space are categorical variables that assume the values: 0, 1 and 2. In order to estimate a correct regression model, each variable had to be converted into two independent variables, each of them assumed as a dummy variable. For instance, Visibility 1 is a dummy variable, assuming the value of 1 when the observations are with store visibility of 1 and 0 otherwise; Visibility 2, assumes the value 1 when store visibility is 2; The base case is store visibility of 0. The same logic characterizes variable Parking Space.

	Estimate	St. Error	Z value	P-value
Intercept	0.9271	0.0546	16.992	0.0000
Population	0.3976	0.1006	3.954	0.0000
Retail	0.1218	0.0374	3.255	0.0012
Visibility 1	0.0877	0.0453	1.938	0.0527
Visibility 2	-0.0881	0.0476	-1.853	0.0639
Parking 1	-0.2032	0.0667	-3.044	0.0023
Parking 2	-0.2304	0.0674	-3.420	0.0006
Age	0.0054	0.0032	1.667	0.0955
Gato Preto	0.0205	0.0175	1.172	0.2413
Espaço Casa	-0.0760	0.0143	-5.325	0.0000
Area	-0.1646	0.0719	-2.290	0.0220
CASA	-0.0124	0.0166	-0.677	0.4984
Distance	-0.0019	0.0014	-1.352	0.1765
P-value	0.0000			
R^2	0.5084			

Table 17 - Tobit regression model results (aggregate of all analysed stores)

It is also shown that the estimated model has a p-value = 0.000, which means that at least one environmental variable has impact on the efficiency score. In other words, the model highlights the variables that explain better the efficiency score obtained through the DEA model, namely: Population within trade area, Retail, Store Visibility, Parking Space, Age, Number of Espaço Casa within trade area and Number of Area within trade area. Those variables have a p-value below the level of significance (5% for some variables and 10% for other) meaning that they are statistically significant in the estimated model. It can also be observed that those variables have different impacts on efficiency.

For instance, **Population within trade area** – with a statistical significance of 5% - has a positive impact in efficiency: as population increases in 1 million residents, it will affect the efficiency in 0.3976 (approximately 40%); **Retail** – with a statistical significance of 5% - also affects positively the efficiency: whenever a store is located in a retail space, the efficiency will increase in 0.1218 (approximately 12%). This may be the result of store location in a Retail being more attractive to costumers due to the higher traffic generated; **Age** – with a statistical significance of 10% - also has a positive impact in efficiency: as the age of the store increases, the efficiency will increase in 0.0054 (approximately 0.5%), possibly because a more consolidated business.

The **Number of Espaço Casa within trade area** and the **Number of Area within trade area** – both with a statistical significance of 5% - have a clearly negative effect in efficiency: as the number of Espaço Casa and the number of Area increases, efficiency will decrease in 0.0760 and 0.1646, respectively (approximately -7,6% and -16%, respectively). Through these estimates it is clear that the presence of Area stores has a more harmful effect on efficiency than that of Espaço Casa stores.

Store Visibility and Parking Space have different interpretations, since they are characterized as categorical variables. **Store visibility of 1** – with a statistical significance of 10% - has a positive impact on efficiency of stores being located on places with a store visibility of 1. Instead, the impact of **Store visibility of 2** – with a statistical significance of 10% - seems to have a negative effect on efficiency, when compared to the base case of visibility 0. This may be the result of the fact that all stores with visibility of 2 are usually very well located and therefore visibility is not really a factor influencing efficiency. However, those variables may be capturing effects of other environmental variables not considered in the model. **Parking Space of 1** and **Parking Space of 2** – both with a statistical significance of 5% - have a negative impact on the efficiency of stores located on places with parking space of 1 or 2, when compared to the base case of parking 0. Following the previous interpretation, these estimates may be a result of stores with a parking space of 1 or even 2 being well located and therefore parking will not be an impactful factor. As before, one should be aware that these variables may be capturing other effects beyond those considered in the regression model.

After analysing how the statistically significant variables explain the efficiency of DeBorla stores, it might be interesting to know which variables are more likely to determine efficiency. Figure 14 shows how the different environmental variables impact the efficiency of the highest and lowest performing units. For instance, units with the highest performance have clearly less **Population** within trade area and a significantly lower number of competitors – **Espaço Casa and Area** – within trade area, than those with the lowest performance. They are also slightly older units. The normalized radar graph (Figure 14) also shows that there is not much difference regarding Retail, Store Visibility and Parking Space, leading to the conclusion that they are not predominant factors in determining efficiency, probably because a minimum level of quality for those factors is always assured

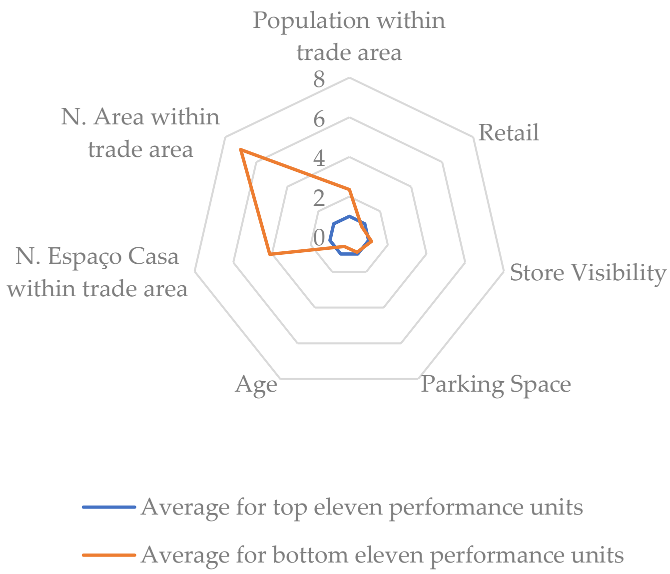

Figure 14 – Impact on efficiency of the environmental variables -benchmarking top with bottom performers

The information given by the radar graph is complemented by the one in the Table 18.

<table>
<tr><th></th><th>Population within trade area</th><th>Retail</th><th colspan="2">Store Visibility</th><th colspan="2">Parking Space</th><th>Age</th><th>N. Espaço Casa within trade area</th><th>N. Area within trade area</th></tr>
<tr><td rowspan="2">Average for top eleven performance units</td><td rowspan="2">477 258</td><td>1</td><td colspan="2">1</td><td colspan="2">2</td><td rowspan="2">6</td><td rowspan="2">2</td><td rowspan="2">0</td></tr>
<tr><td>(91%)</td><td>(82%)
Visibility 1</td><td>(18%)
Visibility 2</td><td>(18%)
Parking 1</td><td>(82%)
Parking 2</td></tr>
<tr><td rowspan="2">Average for bottom eleven performance units</td><td rowspan="2">1 123 213</td><td>1</td><td colspan="2">1</td><td colspan="2">2</td><td rowspan="2">4</td><td rowspan="2">6</td><td rowspan="2">1</td></tr>
<tr><td>(73%)</td><td>(64%)
Visibility 1</td><td>(36%)
Visibility 2</td><td>(36%)
Parking 1</td><td>(64%)
Parking 2</td></tr>
</table>

Table 18 - Average values of the environmental variables for the top and bottom performing units

Chapter 8 - Conclusions

1. Results

The objective of the present thesis was to analyse store performance in a well-established retail group in Portugal. Data Envelopment Analysis was used to evaluate the efficiency of DeBorla stores and set performance improvement goals; Tobit Regression Model, was used to find out which environmental variables may have impact in the efficiency of those stores.

In the DEA model applied, Sales was the selected output, and General Store Costs, Salaries and Area, were selected as the most adequate inputs. A comparison between CRS and VRS model, has allowed us to infer that DeBorla stores exhibited scale effects, leading to the conclusion that the activity of these stores revealed variable return to scale. These means that, since the output – sales– varies in different proportions according to the dimension of the stores, they had to be comparable in size when assessing their efficiency. VRS was therefore the DEA model used to evaluate the performance of DeBorla stores.

The VRS model used a dataset composed of 71 observations (each observation for one store in one year of operation in the period 2015-2017). The results have shown that 11 observations attained fully efficiency (100% efficient) and the remaining 60 have shown inefficiency (< 100% efficient). Out of those inefficient observations, 21 referred to an unproductive scale and 46 to an inefficient operation, being this latter the predominant source of technical inefficiency. The majority of the inefficient observations have nevertheless exhibited a good performance, attaining an average efficiency of 83.33%, approximately, meaning that those units could become fully efficient if they achieve a reduction of their inputs in about 16.67%.

Ranking the stores by the three years average efficiency has shown that **DMU 26** was the only 100% efficient DeBorla store, that is fully efficient over the whole period of analysis. **DMU 5** was found to be the unit with the lowest average efficiency (48.31%). Besides knowing where each store stands against each other, it appeared to be interesting to know about the influence of specific locations. The analysis concluded that the North was the region of Continental Portugal where stores had the highest average efficiency (approximately 92%) and, perhaps surprisingly, both the Metropolitan Areas (Porto and Lisbon) were the regions where stores had the lowest average efficiency (about 79% and 80%, respectively). A map displaying the location of DeBorla stores according to their average efficiency, by Municipality, did not reveal any clear geographical pattern.

The analysis of the store performance evolution over time concluded that 2016 was the year presenting a higher average efficiency (87.85%), followed by 2017 (85.13%) and 2015 (84.65%). A group of specific stores were analysed regarding their performance evolution during the period of analysis (2015-2017), with interesting links being established between input and output variables and the likely causes for the observed variations.

With regard to the inefficient units, an assessment of their performance improvement goals suggested that **General Store Costs** and **Area** were the inputs with the largest margin for improvement. Some of those stores have a balanced potential for improvement in all inputs, but others exhibit a higher potential for improvement in specific inputs. Selected stores were analysed in order to determine the inputs being used inefficiently.

Benchmarking the highest performing units against the lowest performing units with selected indicators led to the following conclusions: the top performance units have a more effective use of space, have more productive labour, spend less of their revenues to cover the General Store Costs and have a higher profitable operation.

Finally, the Tobit regression model has allowed the identification of **nine environmental factors** which have been shown to be statistically significant to explain the efficiency variation of DeBorla Stores: Population within trade area, Retail, Age, Number of Espaço Casa and Number of Area within trade area, Visibility and Parking. By comparing the top performing units with the bottom performing units, a conclusion could be drawn that **Population**, number of **Espaço Casa** and number of **Area** and **Age** are the environmental variables more likely to determine store efficiency.

This work has provided evidence of the power and usefulness of statistical analysis tools, like DEA and Tobit Regression models, in the assessment and benchmark of retail store performance.

2. Contributions and Future Work

This thesis allowed performance analysis of 27 stores of the DeBorla retail group during the period from 2015 to 2017 to be undertaken. The limited time frame to develop the work, as well as constraints regarding the store information made available, led to limitations both in the scope and in the depth of the analysis. The exploitation of available dataset and the highly valuable discussions with both the thesis supervisor and the firm Controlling Manager suggested other paths still to be explored. Some of those paths are now briefly presented.

a) Assessing the impact on performance of alternative variables

While the dataset was being built, a set of environmental variables were considered in the study, namely the number of direct competitors of DeBorla stores. Considering also the impact of indirect competitors on store performance might help us to come closer to reality. Similarly, when selecting the variables used in both methodologies, it is possible that some variables that were not considered could lead to different results. For the DEA model it might be interesting adding the store's inventory and the value of assets for each store as inputs, as well as the number of customers for each store as output. Regarding the Tobit Regression Model, one could also consider the seniority of store employees (number of years of experience).

b) In depth store analysis

An analysis of the evolution of store efficiency over the period 2015-2017 and the definition of performance targets for inputs and outputs of the inefficient units were made for specific stores. However, a more in depth analysis for each store appears to be of fundamental relevance. Among other, possible retail indicators to be used as variables could be: Effectivity (Retail Conversion Rate); Average Sale; Gross Margin; Inventory Turnover and Customer Retention. This would certainly require further information to be provided by DeBorla group to complete the existing dataset.

c) Store product/shelf segmentation analysis

Another potentially relevant analysis that could not be done, although suggested by DeBorla Group, was the analysis of the area of the store shelf allocated to each product category – *kitchen, Interior Design, Storage, Textile, Garden* and *Bathroom* – for the efficient units. The knowledge and good practices acquired by the benchmarks would then be transferred to the inefficient units, with the objective of improving the performance of those stores.

d) Optimizing the location of future stores

The analysis of store performance undertaken by this study should be an instrument to support decision making by the management in different domains. An analysis leading to the selection of feasible optimal locations for future DeBorla stores has obviously a great potential interest. A methodology supporting the assessment of the "quality" of potential locations selected by DeBorla group, considering the environmental characteristics of their efficient stores, could be developed making full use of QGIS.

e) Extending the present analysis to other retail groups

Finally, similar type studies may possibly be undertaken in other retail areas, both analysing individual stores and benchmarking them within the same group of stores or against competitors.

References

Almohri, H., Chinnam, R. B., & Colosimo, M. (2018). Data-Driven Analytics for Benchmarking and Optimizing Retail Store Performance.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis*. *Management Science*, *30*(9), 1078–1092.

Banker, R. D., & Morey, R. C. (1986). Efficiency Analysis for Exogenously Fixed Inputs and Outputs. *Operations Research*, *34*(4), 513–521.

Barros, C. P. (2006). Efficiency measurement among hypermarkets and supermarkets and the identification of the efficiency drivers: A case study. *International Journal of Retail & Distribution Management*, *59*, 602–614.

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

Coelli, T. J., Prasada Rao, D. S., O'Donnell, C. J., & Battese, G. E. (2005). *An Introduction to Efficiency and Productivity Analysis* (2nd ed). Springer.

Cook, W. D., & Zhu, J. (2005). Data Envelopment Analysis. *Modeling Performance Measurement*, (2002), 1–27.

Cooper, W. W., Seiford, L. M., & Zhu, J. (2004). Data envelopment analysis: History, Models and Interpretations. *Handbook on Data Envelopment Analysis*, 1–39.

Cooper, W. W., Seiford, L. M., & Zhu, J. (2011). Chapter 1. Data Envelopment Analysis. *Handbook on Data Envelopment Analysis, Second Edition*, *164*(2002), 1–39.

Donthu, N., & Yoo, B. (1998). Retail productivity Assessment using Data Envelopment Analysis.pdf. *Journal of Retailing*, *74*(1), 89–105.

Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V, Sarrico, C. S., & Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, *132*(2), 245–259.

Farrell, M. . J. . (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*, *120*(3), 253–290.

Fernandes, M. C. R. B. V. (2007). *Desenvolvimento de um sistema de avaliação e melhoria de desempenho no sector do retalho*. Faculdade de Engenharia da Universidade do Porto.

Gandhi, A., & Shankar, R. (2014). Efficiency measurement of Indian retailers using Data Envelopment Analysis. *Measuring Business Excellence*, *42*(6), 500–520.

Gandhi, A., & Shankar, R. (2016). Strategic resource management model and data envelopment analysis for benchmarking of Indian retailers. *Benchmarking*, *23*(2), 286–312.

Jorge Moreno, J. de, & Sanz-Triguero, M. (2011). Estimating technical efficiency and bootstrapping Malmquist indices: Analysis of Spanish retail sector. *International Journal of Retail and Distribution Management*, *39*(4), 272–288.

Kamakura, W. A., Lenartowicz, T., & Ratchford, B. T. (1996). Productivity assessment of multiple retail outlets. *Journal of Retailing*, *72*(4), 333–356.

Ko, K., Chang, M., Bae, E. S., & Kim, D. (2017). Efficiency analysis of retail chain stores in Korea. *Sustainability (Switzerland)*, *9*(9), 1–14.

Koopmans, T. C. (1951). Activity Analysis of Production and Allocation. *Economic Journal*, *62*(247), 625–628.

Lau, K. H. (2013). Measuring distribution efficiency of a retail network through data envelopment analysis. *International Journal of Production Economics*, *146*(2), 598–611.

Manuel Xavier, J., Ferreira Moutinho, V., & Carrizo Moreira, A. (2015). An empirical examination of performance in the clothing retailing industry: A case study. *Journal of Retailing and Consumer Services*, *25*, 96–105.

Parsons, L. J. (1992). Productivity Versus Relative Efficiency in Marketing: Past and Future? In *Research traditions in marketing* (pp. 169–196). Amsterdam: Kluwer.

Perrigot, R., & Barros, C. P. (2008). Technical efficiency of French retailers. *Journal of Retailing and Consumer Services*, *15*(4), 296–305.

Pestana Barros, C., & Alves, C. A. (2003). Hypermarket retail store efficiency in Portugal. *International Journal of Retail & Distribution Management*, *31*(11), 549–560.

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, *136*(1), 31–64.

Thanassoulis, E. (2001). *Introduction to the Theory and Application of Data Envelopment Analysis: A Foundation Text with Integrated Software*. Springer US.

Thomas, R. R., Barr, R. S., Cron, W. L., & Slocum, J. W. (1998). A process for evaluating retail store efficiency: A restricted DEA approach. *International Journal of Research in Marketing*, *15*(5), 487–503.

Thurik, R. (1992). Applied econometrics and productivity analysis in marketing. In *Research Traditions in Marketing* (pp. 197–200). Amsterdam: Kluwer.

Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. *Journal of the Econometric Society*, *26*(1), 24–36.

Uyar, A., Bayyurt, N., & Dilber, M. (2013). Evaluating operational efficiency of a bookshop chain in Turkey and identifying efficiency drivers. *International Journal of Retail & Distribution Management*, *41*(5), 331–347.

Vyt, D. (2008). Retail network performance evaluation: a DEA approach considering retailers' geomarketing. *The International Review of Retail*, *18*(2), 235–253.

Vyt, D., & Cliquet, G. (2017). Towards a fairer manager performance measure: a DEA application in the retail industry. *International Review of Retail, Distribution and Consumer Research*, *27*(5), 450–467.

Yu, W., & Ramanathan, R. (2008). An assessment of operational efficiencies in the UK retail sector. *International Journal of Retail & Distribution Management*, *36*(11), 861–882.

Yu, W., & Ramanathan, R. (2009). An assessment of operational efficiency of retail firms in China. *Journal of Retailing and Consumer Services*, *16*(2), 109–122.

Appendix

Appendix A - Attach to Chapter 4

Table 19 summarizes a number of studies identified in the literature review regarding the assessment of store performance. The Table present the methodologies applied, the DMUs, the inputs and outputs used for the DEA model and the explanatory variables for the Tobit Regression model.

Studies	Methodologies	DMUs	Inputs	Outputs	Explanatory Variables
(Barros, 2006)	DEA Tobit Regression Model	22 hypermarkets in Portugal	Labour, Capital	Sales, Operational results, Value added	Share, Outlets, Ownerships, Regulation, Location
(Yu & Ramanathan, 2008)	DEA Tobit Regression Model	41 retail firms from UK	Total assets, Shareholders funds, Employees	Turnover, Profit before taxation	Head office location, Types of ownership, Years of incorporation, Legal form, Retail characteristics
(Yu & Ramanathan, 2009)	DEA Tobit Regression Model	61 retailers from China	Total selling floor space, Employees	Sales, Profit before taxation	Head office location, Firm nationality, Years of incorporation, Ownership type, Retailing characteristic
(Manuel Xavier et al., 2015)	DEA	40 Portuguese clothing retail stores	Total salaries and wages, Rental costs, Investments in assets	Sales volume, Earnings before income taxes and amortization	
(Gandhi & Shankar, 2016)	DEA	11 Indian retailers	Number of employees, Square foot area, Inventories	Sales	
(Uyar et al., 2013)	DEA Tobit Regression Model	79 Turkey bookshops chain	Area, Population, Inventory, Employee, Salaries, Other costs	Sales, Profit	Education of manager, Experience of manager, Age of shop
(Pestana Barros & Alves, 2003)	DEA	47 Portuguese retail outlets	Full-time employees, Part-time employees, Cost of labour, Absenteeism; Area of outlets, Number of point of sales, Age of	Sales, Operational results	

			outlet, Inventory, Other costs		
(Gandhi & Shankar, 2014)	DEA Tobit Regression Model	18 Indian retailers	Cost of labour, Capital Employed	Sales, Profit	Outlets, Ownership, Age since incorporation, Mergers and acquisitions
(Ko et al., 2017)	DEA Tobit Regression Model	32 Korean retailers	Store size, Number of items, Number of employees, Rental cost	Sales revenue, Number of customers	Store age, Number of items per unit area, Number of items per employee, Trade area index, Number of competitor stores
Vyt, 2008)	DEA Tobit Regression Model	French supermarket retail chain	Sales area, Number of employees, Space allocation	Turnover	Percentage of seniors, Unemployment rate, Percentage of second homes, Percentage of individual homes, Number of inhabitants, Competition index for supermarkets, Competition index for convenient stores, Competition index
(Perrigot & Barros, 2008)	DEA Tobit Regression Model	11 French retailers	Labour, Capital, Total costs	Turnover, Profit	Trend, Square trend, Quoted, M&A, Group, International
(Vyt & Cliquet, 2017)	DEA Ordinary Least Squares (OLS) model	38 stores of the French of the supermarket retail chain	Size, Labour, Shelf space allocation of the category	Turnover	Percentage of seniors, Unemployment rate, Percentage of second homes, Percentage of individual homes, Number of inhabitants, Competition index for supermarkets, Competition index for convenient stores, Competition index for hard discounters

Table 19 - Framework of the use of Data Envelopment Analysis and Tobit Regression Model studies (built out of in the literature review)

Appendix B - Attach to Chapter 5

B.1 Values of the variable "population within the trade area" for each DeBorla store (in millions of people)

DMU	Total resident population (in millions)
DMU 1	0.284230
DMU 10	2.363051
DMU 11	0.135895
DMU 12	0.128513
DMU 13	0.204775
DMU 14	2.046684
DMU 15	1.903503
DMU 16	0.109465
DMU 17	2.436173
DMU 18	1.388100
DMU 19	1.987515
DMU 2	0.562485
DMU 20	0.452105
DMU 21	0.321255
DMU 22	1.115262
DMU 23	2.568520
DMU 24	0.192082
DMU 25	0.363822
DMU 26	0.587158
DMU 27	0.278578
DMU 3	1.208178
DMU 4	1.878897
DMU 5	0.082176
DMU 6	0.091514
DMU 7	0.477897
DMU 8	2.208883
DMU 9	0.097413

Table 20 - Population within the trade area

B.2 Values of the average distance travelled to each DeBorla Store

DMU	Average distance traveled (km)
DMU 1	135.04
DMU 10	18.92
DMU 11	59.81
DMU 12	78.01
DMU 13	124.87
DMU 14	23.35
DMU 15	29.73
DMU 16	63.91
DMU 17	40.27
DMU 18	24.02
DMU 19	18.81
DMU 2	33.89
DMU 20	40.43
DMU 21	40.44
DMU 22	36.5
DMU 23	29.08
DMU 24	171.25
DMU 25	59.72
DMU 26	33.8
DMU 27	40.14
DMU 3	29.38
DMU 4	17.76
DMU 5	68.54
DMU 6	75.72
DMU 7	41.51
DMU 8	17.88
DMU 9	58.1

Table 21 - Average distance traveled by clients to each DeBorla Store

B.3 Number of stores of each competitor within the trade area

Table 22 shows the number of Gato Preto, Espaço Casa, Casa and Area stores within each DeBorla store trade area.

DMU	Number of Gato Preto stores	Number of Espaço Casa stores	Number of CASA stores	Number of Area stores
DMU 1	4	2	3	0
DMU 10	6	14	3	1
DMU 11	1	1	0	0
DMU 12	0	0	0	0
DMU 13	2	1	2	0
DMU 14	4	12	2	1
DMU 15	4	11	1	1
DMU 16	0	1	0	0
DMU 17	13	12	8	2
DMU 18	5	5	1	1
DMU 19	4	11	1	1
DMU 2	2	3	1	0
DMU 20	1	0	0	0
DMU 21	0	0	0	0
DMU 22	8	6	6	1
DMU 23	13	10	7	3
DMU 24	2	2	2	0
DMU 25	1	2	1	0
DMU 26	2	1	2	0
DMU 27	1	1	1	0
DMU 3	2	6	3	0
DMU 4	5	10	1	1
DMU 5	1	1	0	0
DMU 6	0	1	0	0
DMU 7	2	2	0	0
DMU 8	4	12	2	1
DMU 9	0	1	1	0

Table 22 - Number of Gato Preto, Espaço Casa, Area and CASA stores within the trade area

B.4 Store distance to the nearest competitor

Table 23 shows the nearest competitor for each DeBorla store and the respective distance.

DMU	Nearest competitor	Distance (km)
DMU 1	Gato Preto	0.177
DMU 10	Espaço Casa	0.162
DMU 11	Espaço Casa	11.364
DMU 12	Gato Preto	36.037
DMU 13	Gato Preto	2.549
DMU 14	Espaço Casa	2.613
DMU 15	Gato Preto	0.466
DMU 16	Espaço Casa	0.458
DMU 17	CASA	0.142
DMU 18	Espaço Casa	10.249
DMU 19	Gato Preto	1.447
DMU 2	CASA	0.270
DMU 20	Gato Preto	14.784
DMU 21	Espaço Casa	45.099
DMU 22	Gato Preto	0.902
DMU 23	CASA	0.121
DMU 24	CASA	0.068
DMU 25	CASA	20.693
DMU 26	CASA	0.270
DMU 27	CASA	5.631
DMU 3	Gato Preto	0.671
DMU 4	Gato Preto	8.407
DMU 5	Espaço Casa	3.709
DMU 6	Espaço Casa	1.184
DMU 7	Gato Preto	6.157
DMU 8	Espaço Casa	3.583
DMU 9	CASA	1.447

Table 23 – Store distance matrix to the nearest competitor

Appendix C - Attach to Chapter 6

C.1 DEA model results without environmental variables for all 71 observations

Table 24 shows the efficiency scores, applying the DEA model without environmental variables, with CRS and VRS hypotheses, as well as the Scale Efficiency and the Source of Technical Inefficiency, for all 71 observations.

DMU	CRS	VRS	Scale Efficiency	Source of Technical Inefficiency
DMU1 2015	96.81%	100.00%	96.81%	Unproductive scale
DMU1 2016	100.00%	100.00%	100.00%	Efficient
DMU1 2017	91.37%	97.69%	93.53%	Unproductive scale
DMU10 2015	64.61%	84.34%	76.61%	Unproductive scale
DMU10 2016	72.67%	91.83%	79.14%	Unproductive scale
DMU10 2017	67.79%	88.35%	76.73%	Unproductive scale
DMU11 2015	65.05%	94.81%	68.61%	Unproductive scale
DMU11 2016	66.08%	92.37%	71.54%	Unproductive scale
DMU11 2017	62.00%	87.28%	71.04%	Unproductive scale
DMU12 2015	73.79%	86.70%	85.11%	Unproductive scale
DMU12 2016	74.88%	84.69%	88.42%	Inefficient operation
DMU12 2017	87.98%	100.00%	87.98%	Unproductive scale
DMU13 2015	66.67%	71.22%	93.61%	Inefficient operation
DMU13 2016	98.09%	98.95%	99.13%	Inefficient operation
DMU13 2017	97.28%	98.30%	98.96%	Inefficient operation
DMU14 2015	68.82%	73.60%	93.51%	Inefficient operation
DMU14 2016	67.38%	71.79%	93.86%	Inefficient operation
DMU14 2017	66.42%	68.82%	96.51%	Inefficient operation
DMU15 2016	49.69%	59.66%	83.29%	Inefficient operation
DMU15 2017	62.54%	76.02%	82.27%	Inefficient operation
DMU16 2016	59.31%	100.00%	59.31%	Unproductive scale
DMU16 2017	57.74%	91.80%	62.90%	Unproductive scale
DMU17 2015	63.42%	64.39%	98.49%	Inefficient operation
DMU17 2016	73.91%	74.22%	99.58%	Inefficient operation
DMU17 2017	81.76%	81.80%	99.95%	Inefficient operation

DMU18 2015	94.19%	94.26%	99.93%	Inefficient operation
DMU18 2016	100.00%	100.00%	100.00%	Efficient
DMU18 2017	80.87%	81.20%	99.59%	Inefficient operation
DMU19 2015	77.21%	78.92%	97.83%	Inefficient operation
DMU19 2016	88.22%	90.03%	97.99%	Inefficient operation
DMU19 2017	79.89%	81.32%	98.24%	Inefficient operation
DMU2 2015	81.93%	82.04%	99.87%	Inefficient operation
DMU2 2016	82.27%	84.08%	97.85%	Inefficient operation
DMU2 2017	82.95%	84.07%	98.67%	Inefficient operation
DMU20 2015	88.28%	94.46%	93.46%	Unproductive scale
DMU20 2016	93.46%	99.25%	94.17%	Unproductive scale
DMU20 2017	100.00%	100.00%	100.00%	Efficient
DMU21 2015	91.02%	92.29%	98.62%	Inefficient operation
DMU21 2016	96.98%	97.50%	99.47%	Inefficient operation
DMU21 2017	96.89%	97.78%	99.09%	Inefficient operation
DMU22 2015	71.43%	73.37%	97.36%	Inefficient operation
DMU22 2016	73.71%	75.25%	97.95%	Inefficient operation
DMU22 2017	77.19%	78.39%	98.47%	Inefficient operation
DMU23 2015	86.98%	87.98%	98.86%	Inefficient operation
DMU23 2016	88.44%	88.45%	99.99%	Inefficient operation
DMU23 2017	86.96%	94.32%	92.20%	Unproductive scale
DMU24 2015	75.43%	77.52%	97.30%	Inefficient operation
DMU24 2016	79.70%	82.47%	96.64%	Inefficient operation
DMU24 2017	84.39%	84.64%	99.70%	Inefficient operation
DMU25 2016	64.38%	100.00%	64.38%	Unproductive scale
DMU25 2017	59.09%	77.27%	76.47%	Unproductive scale
DMU26 2015	65.10%	100.00%	65.10%	Unproductive scale
DMU26 2016	78.00%	100.00%	78.00%	Unproductive scale
DMU26 2017	73.83%	100.00%	73.83%	Unproductive scale
DMU27 2017	57.69%	60.95%	94.65%	Inefficient operation
DMU3 2015	88.81%	89.72%	98.99%	Inefficient operation
DMU3 2016	85.04%	85.87%	99.03%	Inefficient operation
DMU3 2017	82.07%	83.58%	98.19%	Inefficient operation
DMU4 2015	71.35%	76.81%	92.89%	Inefficient operation
DMU4 2016	79.04%	84.22%	93.85%	Inefficient operation
DMU4 2017	87.60%	91.39%	95.85%	Inefficient operation
DMU5 2017	38.34%	48.31%	79.36%	Inefficient operation
DMU6 2017	55.28%	89.13%	62.02%	Unproductive scale
DMU7 2015	94.46%	94.65%	99.80%	Inefficient operation
DMU7 2016	100.00%	100.00%	100.00%	Efficient
DMU7 2017	97.59%	98.13%	99.45%	Inefficient operation
DMU8 2015	75.50%	75.99%	99.36%	Inefficient operation

DMU8 2016	82.15%	82.81%	99.20%	Inefficient operation
DMU8 2017	88.07%	88.23%	99.82%	Inefficient operation
DMU9 2016	56.91%	65.02%	87.53%	Inefficient operation
DMU9 2017	64.89%	69.67%	93.14%	Inefficient operation

Table 24 - CRS, VRS, Scale Efficiency and Sources of technical efficiency results for all observations

C.2 DEA model efficiency results, assuming VRS, without environmental variables, for all 27 stores

Table 25 shows the efficiency scores, resulting from the application of the DEA model without environmental variables, assuming VRS, for each period under analysis and for all 27 stores.

DMU	2015	2016	2017	Average Efficiency
DMU 1	100.00%	100.00%	97.69%	99.23%
DMU 10	84.34%	91.83%	88.35%	88.17%
DMU 11	94.81%	92.37%	87.28%	91.49%
DMU 12	86.70%	84.69%	100.00%	90.46%
DMU 13	71.22%	98.95%	98.30%	89.49%
DMU 14	73.60%	71.79%	68.82%	71.40%
DMU 15		59.66%	76.02%	67.84%
DMU 16		100.00%	91.80%	95.90%
DMU 17	64.39%	74.22%	81.80%	73.47%
DMU 18	94.26%	100.00%	81.20%	91.82%
DMU 19	78.92%	90.03%	81.32%	83.42%
DMU 2	82.04%	84.08%	84.07%	83.40%
DMU 20	94.46%	99.25%	100.00%	97.90%
DMU 21	92.29%	97.50%	97.78%	95.86%
DMU 22	73.37%	75.25%	78.39%	75.67%
DMU 23	87.98%	88.45%	94.32%	90.25%
DMU 24	77.52%	82.47%	84.64%	81.54%
DMU 25		100.00%	77.27%	88.64%
DMU 26	100.00%	100.00%	100.00%	100.00%
DMU 27			60.95%	60.95%
DMU 3	89.72%	85.87%	83.58%	86.39%
DMU 4	76.81%	84.22%	91.39%	84.14%
DMU 5			48.31%	48.31%
DMU 6			89.13%	89.13%
DMU 7	94.65%	100.00%	98.13%	97.59%
DMU 8	75.99%	82.81%	88.23%	82.34%
DMU 9		65.02%	69.67%	67.35%

Table 25 – Efficiency of each of the 27 stores per year and store average efficiency

C.3 DEA model benchmarks results, assuming VRS, without environmental variables

Table 26 presents the benchmarks for each of the 27 stores during each year of the period of analysis and the specific contribution of each benchmark (λ) for the efficiency measurement.

DMU	2015	2016	2017
DMU 1	DMU11 2015	DMU25 2017	DMU1 2015 (0.07), DMU16 (0.74), DMU26 2015 (0.19)
DMU 10	DMU1 2015 (0.06), DMU1 2016(0.15), DMU25 2016 (0.21), DMU26 2016 (0.58)	DM1 2016 (0.36), DMU12 2017 (0.11), DMU16 2016 (0.53)	DMU1 2016 (0.36), DMU16 2016 (0.59), DMU26 2016 (0.05)
DMU 11	DMU1 2015 (0.14), DMU25 2016 (0.27), DMU26 2015 (0.59)	DMU21 2016 (0.21), DMU25 2016 (0.18), DMU26 2015 (0.62)	DMU25 2016 (0.10), DMU26 2015 (0.14), DMU26 2016 (0.77)
DMU 12	DMU1 2016 (0.32), DMU12 2017 (0.42), DMU16 2016 (0.22), DMU25 2016 (0.03)	DMU1 2016 (0.4), DMU12 2017 (0.36), DMU16 2016 (0.17), DMU25 2016 (0.03)	DMU11 2017
DMU 13	DMU1 2016 (0.76), DMU26 2015 (0.08), DMU26 2017 (0.17)	DMU1 2016 (0.62), DMU18 2016 (0.13), DMU20 2017 (0.25)	DMU1 2016 (0.50), DMU20 2017 (0.49), DMU7 2016 (0.01)
DMU 14	DMU1 2016 (0.41), DMU 18 2016 (0.27), DMU25 2016 (0.32)	DMU1 2016 (0.36), DMU12 2017 (0.08), DMU18 2016 (0.33), DMU25 2016 (0.23)	DMU1 2016 (0.68), DMU18 2016 (0.13), DMU25 2016 (0.19)
DMU 15		DMU1 2016 (0.11), DMU18 2016 (0.27), DMU25 2016 (0.62)	DMU20 2017 (0.37), DMU7 2016 (0.63)
DMU 16		DMU10 2015	DMU25 2016 (0.25), DMU26 2015 (0.02), DMU26 2016 (0.73)
DMU 17	DMU1 2016 (0.85),DMU25 2016 (0.15)	DMU1 2016 (0.88), DMU18 2016 (0.10), DMU25 2016 (0.03)	DMU1 2016 (0.57), DMU18 2016 (0.30), DMU20 2017 (0.13)
DMU 18	DMU1 2016 (0.27), DMU18 2016 (0.56), DMU20 2017 (0.18)	DMU13 2015	DMU1 2016 (0.91), DMU7 2016 (0.09)
DMU 19	DMU1 2016 (0.95), DMU7 2016 (0.05)	DMU1 2016 (0.86), DMU7 2016 (0.14)	DMU1 2016 (0.94), DMU7 2016 (0.06)
DMU 2	DMU1 2016 (0.95), DMU7 2016 (0.05)	DMU1 2016 (0.60), DMU20 2017 (0.26), DMU7 2016 (0.14)	DMU1 2016 (0.65), DMU20 2017 (0.16), DMU7 2016 (0.18)
DMU 20	DMU1 2016 (0.30), DMU20 2017 (0.60), DMU7 2016 (0.18)	DMU1 2016 (0.17), DMU20 2017 (0.74), DMU7 2016 (0.09)	DMU13 2015
DMU 21	DMU1 2016 (0.99), DMU7 2016 (0.01)	DMU1 2016 (0.90), DMU7 2016 (0.10)	DMU1 2016 (0.40), DMU20 2017 (0.58), DMU7 2016 (0.02)
DMU 22	DMU1 2016 (0.98), DMU25 2016 (0.02)	DMU1 2016 (0.95), DMU7 2016 (0.05)	DMU1 2016 (0.94), DMU7 2016 (0.65)
DMU 23	DMU1 2016 (0.72), DMU7 2016 (0.28)	DMU1 2016 (57), DMU7 2016 (0.43)	DMU1 2016 (50), DMU7 2016 (0.50)
DMU 24	DMU1 2016 (0.48), DMU12 2017 (0.14), DMU18 2016 (0.37), DMU25 2016 (0.01)	DMU1 2016 (0.44), DMU20 2017 (0.43), DMU7 2016 (0.13)	DMU1 2016 (0.09), DMU20 2017 (0.91), DMU7 2016 (0.00)

DMU 25		DMU17 2015	DMI1 2015 (0.31), DMU25 2016 (0.09), DMU26 2015 (0.08), DMU26 2016 (0.52)
DMU 26	DMU12 2015	DMU10 2017	DMU1 2015
DMU 27			DMU1 2016 (0.73), DMU25 2016 (0.27)
DMU 3	DMU1 2016 (0.74), DMU7 2016 (0.26)	DMU1 2016 (0.77), DMU7 2016 (0.23)	DMU1 2016 (0.68), DMU7 2016 (0.32)
DMU 4	DMU1 2016 (0.51), DMU12 2017 (0.19), DMU18 2016 (0.14), DMU25 2016 (0.16)	DMU1 2016 (0.35), DMU12 2017 (0.46), DMU18 2016 (0.19)	DMU1 2016 (0.41), DMU12 2017 (0.27), DMU18 2016 (0.32)
DMU 5			DMI1 2015 (0.27), DMU25 2016 (0.59), DMU26 2015 (0.14)
DMU 6			DMU25 2016 (0.10), DMU26 2015 (0.57), DMU26 2016 (0.33)
DMU 7	DMU1 2016 (0.08), DMU7 2016 (0.92)	DMU18 2017	DMU1 2016 (0.03), DMU7 2016 (0.97)
DMU 8	DMU1 2016 (0.96), DMU25 2016 (0.04)	DMU1 2016 (0.59), DMU20 2017 (0.39), DMU7 2016 (0.02)	DMU1 2016 (0.84), DMU7 2016 (0.14)
DMU 9		DMI1 2015 (0.49), DMU25 2016 (0.46), DMU26 2015 (0.05)	DMU1 2016 (0.59), DMU12 2017 (0.04), DMU18 2016 (0.06), DMU25 2016 (0.31)

Table 26 - Benchmarks results from the DEA model with VRS, without environmental variables

C.4 Performance improvement goals according to DEA model without environmental variables, assuming VRS

Table 27 presents the current and the target values for the input **General Store Costs**, for each of the 27 stores during each year of the period of analysis.

	General Store Costs					
	Current Values			Target Values		
DMU	2015	2016	2017	2015	2016	2017
DMU 1	45 690	44 823	46 400	45 690	44 823	44 303
DMU 10	39 331	35 882	36 435	33 172	32 951	32 190
DMU 11	44 426	42 662	38 868	35 701	38 363	33 924
DMU 12	40 700	43 851	35 296	35 287	37 138	35 296
DMU 13	79 014	45 916	49 067	44 267	45 434	48 233
DMU 14	47 288	50 585	56 716	34 804	36 315	39 032
DMU 15		44 790	127 087		26 722	96 612
DMU 16		24 400	33 160		24 400	30 441
DMU 17	64 192	58 858	53 817	40 834	43 684	44 022
DMU 18	45 706	40 134	68 778	43 082	40 134	52 366
DMU 19	91 054	82 251	78 963	48 564	55 907	49 317
DMU 2	60 777	68 093	71 497	48 749	57 252	60 108
DMU 20	58 676	56 135	49 670	55 426	55 714	49 670
DMU 21	61 194	59 291	50 193	45 386	53 113	49 079
DMU 22	93 239	88 347	80 159	44 284	48 424	49 894
DMU 23	103 121	99 625	99 527	67 326	78 716	84 238
DMU 24	53 582	69 601	58 325	41 537	57 400	49 366
DMU 25		17 682	47 924		17 682	37 031
DMU 26	41 752	34 543	42 913	41 752	34 543	42 913
DMU 27			89 735			37 529
DMU 3	94 137	101 107	95 602	65 581	62 878	70 074
DMU 4	49 503	46 953	44 566	38 023	39 544	40 729
DMU 5			60 596			28 664
DMU 6			41 560			37 042
DMU 7	125 882	124 424	132 086	117 710	124 424	122 077
DMU 8	60 624	58 014	65 383	43 781	48 042	56 118
DMU 9		54 985	51 269		32 718	35 719

Table 27 - Performance improvement goals for General Store Costs

Table 28 presents the current and the target values for the input **Salaries**, for each of the 27 stores during each year of the period of analysis.

	Salaries					
	Current Values			Target Values		
DMU	**2015**	**2016**	**2017**	**2015**	**2016**	**2017**
DMU 1	127 576	146 015	145 099	127 576	146 015	141 747
DMU 10	138 304	140 416	144 650	116 646	119 383	113 052
DMU 11	109 117	122 147	146 009	103 454	112 827	127 437
DMU 12	161 168	167 426	169 212	139 733	141 793	169 212
DMU 13	202 348	206 049	195 301	144 112	163 222	169 924
DMU 14	165 495	191 056	189 758	121 804	137 159	130 591
DMU 15		148 185	380 065		88 407	239 977
DMU 16		91 062	122 194		91 062	112 174
DMU 17	201 046	198 653	202 330	129 453	147 440	165 506
DMU 18	189 916	190 593	194 112	179 015	190 593	157 619
DMU 19	192 308	181 127	188 055	151 770	163 068	152926
DMU 2	185 340	208 507	209 156	152 053	175 312	175 837
DMU 20	211 301	203 421	191 903	185 474	191 055	191 903
DMU 21	159 154	162 838	178 950	146 884	158 767	174 978
DMU 22	195 967	201 400	196 223	143 781	151 554	153 819
DMU 23	205 313	224 028	239 184	180 634	198 153	206 651
DMU 24	212 894	220 397	222 047	165 036	181 762	187 941
DMU 25		33 332	161 663		33 332	124 917
DMU 26	130 422	138 609	141 784	130 422	138 609	141 784
DMU 27			189 878			115 731
DMU 3	198 339	202 387	225 551	177 949	173 790	184 862
DMU 4	180 764	198 171	193 057	138 845	165 189	166 593
DMU 5			150 113			72 520
DMU 6			138 782			123 696
DMU 7	273 339	268 473	271 542	258 142	268 473	264 860
DMU 8	186 452	200 624	185 193	141 685	166 137	163 396
DMU 9		130 328	164 903		84 739	114 888

Table 28 - Performance improvement goals for Salaries

Table 29 presents the current and the target values for the input **Area**, for each of the 27 stores during each year of the period of analysis.

	Area					
	Current Values			Target Values		
DMU	2015	2 016	2017	2015	2016	2017
DMU 1	938	938	938	938	938	916
DMU 10	1 088	1 088	1 088	917	999	961
DMU 11	980	980	980	929	905	855
DMU 12	1 261	1 261	1 261	1 093	1 068	1 261
DMU 13	1 278	1 278	1 278	910	1 265	1 256
DMU 14	2 034	2 034	2 034	1 374	1 460	1 160
DMU 15		2 947	2 947		1 440	2 187
DMU 16		986	986		986	905
DMU 17	1 914	1 914	1 914	969	1 073	1 428
DMU 18	2 290	2 290	2 290	1 799	2 290	1 092
DMU 19	1 749	1 749	1 749	1 015	1 165	1 030
DMU 2	1 829	1 829	1 829	1 018	1 328	1 335
DMU 20	1 549	1 549	1 549	1 463	1 537	1 549
DMU 21	1 604	1 604	1 604	950	1 108	1 324
DMU 22	1 800	1 800	1 800	942	1 012	1 042
DMU 23	1 849	1 849	1 849	1 398	1 631	1 744
DMU 24	1 915	1 915	1 915	1 484	1414	1 496
DMU 25		1 150	1 150		1 150	889
DMU 26	824	824	824	824	824	824
DMU 27			2 555			995
DMU 3	1 740	1 740	1 740	1 362	1 307	1 454
DMU 4	1 596	1 596	1 596	1 226	1 344	1 459
DMU 5			2 167			1 047
DMU 6			960			856
DMU 7	2 565	2 565	2 565	2 428	2 565	2 517
DMU 8	1 792	1 792	1 792	946	1 206	1 169
DMU 9		1 583	1 583		1 029	1 103

Table 29 - Performance improvement goals for Area

Table 30 presents the current and the target values for the output **Sales**, for each of the 27 stores during each year of the period of analysis.

	Sales					
	Current Values			Target Values		
DMU	2015	2016	2017	2015	2016	2017
DMU 1	1 357 836	1 591 679	1 445 726	1 357 836	1 591 679	1 445 726
DMU 10	937 448	999 476	943 269	937 448	999 476	943 269
DMU 11	788 018	886 681	886 728	788 018	886 681	900 088
DMU 12	1 157 102	1 243 239	1 224 897	1 157 102	1 243 239	1 224 897
DMU 13	1 445 765	1 692 753	1 768 339	1 445 765	1 692 753	1 768 339
DMU 14	1 197 750	1 304 783	1 355 419	1 197 750	1 304 783	1 355 419
DMU 15		796 422	2 604 112		796 422	2 604 112
DMU 16		550 554	722 552		550 554	794 465
DMU 17	1 393 095	157 2119	1 680 254	1 393 095	1 572 119	1 680 254
DMU 18	1 734 787	1 744 456	1 726 186	1 734 787	1 744 456	1 726 186
DMU 19	1 658 434	1 789 303	1 671 859	1 658 434	1 789 303	1 671 859
DMU 2	1 661 712	1 876 619	1 903 448	1 661 712	1 876 619	1 903 448
DMU 20	1 924 614	1 961 505	1 916 818	1 924 614	1 961 505	1 916 818
DMU 21	1 601 673	1 739 555	1 806 773	1 601 673	1 739 555	1 806 773
DMU 22	1 564 922	1 655 937	1 682 071	1 564 922	1 655 937	1 682 071
DMU 23	1 993 001	2 196 177	2 294 570	1 993 001	2 196 177	2 294 570
DMU 24	1 585 963	1 917 878	1 889 786	1 585 963	1 917 878	1 889 786
DMU 25		240 694	1 023 012		240 694	1 023 012
DMU 26	909 900	980 418	1 031 993	909 900	980 418	1 031 993
DMU 27			1 228 578			1 228 578
DMU 3	1 961 829	1 913 704	2 041 943	1 961 829	1 913 704	2 041 943
DMU 4	1 328 532	1 451 828	1 540 676	1 328 532	1 451 828	1 540 676
DMU 5			638 566			638 566
DMU 6			825 774			868 163
DMU 7	2 891 550	3 011 226	2 969 441	2 891 550	3 011 226	2 969 441
DMU 8	1 539 684	1 743 370	1 793 060	1 539 684	1 743 370	1 793 060
DMU 9		825 728	1 167 321		825 728	1 167 321

Table 30 - Performance improvement goals for Sales

C.5 Percentage of potential gains according to DEA model without environmental variables assuming VRS

Table 31 show the percentage of potential gains for the inputs – General Store Costs, Salaries and Area – and for the output – Sales – for each of the 27 stores during each year of the period of analysis.

	General Store Costs			Salaries			Area			Sales		
DMU	**2015**	**2016**	**2017**	**2015**	**2016**	**2017**	**2015**	**2016**	**2017**	**2015**	**2016**	**2017**
DMU 1	0%	0%	5%	0%	0%	2%	0%	0%	2%	0%	0%	0%
DMU 10	16%	8%	12%	16%	15%	22%	16%	8%	12%	0%	0%	0%
DMU 11	20%	10%	13%	5%	8%	13%	5%	8%	13%	0%	0%	1%
DMU 12	13%	15%	0%	13%	15%	0%	13%	15%	0%	0%	0%	0%
DMU 13	44%	1%	2%	29%	21%	13%	29%	1%	2%	0%	0%	0%
DMU 14	26%	28%	31%	26%	28%	31%	32%	28%	43%	0%	0%	0%
DMU 15		40%	24%		40%	37%		51%	26%		0%	0%
DMU 16		0%	8%		0%	8%		0%	8%		0%	9%
DMU 17	36%	26%	18%	36%	26%	18%	49%	44%	25%	0%	0%	0%
DMU 18	6%	0%	24%	6%	0%	19%	21%	0%	52%	0%	0%	0%
DMU 19	47%	32%	38%	21%	10%	19%	42%	33%	41%	0%	0%	0%
DMU 2	20%	16%	16%	18%	16%	16%	44%	27%	27%	0%	0%	0%
DMU 20	6%	1%	0%	12%	6%	0%	6%	1%	0%	0%	0%	0%
DMU 21	26%	10%	2%	8%	3%	2%	41%	31%	17%	0%	0%	0%
DMU 22	53%	45%	38%	27%	25%	22%	48%	44%	42%	0%	0%	0%
DMU 23	35%	21%	15%	12%	12%	14%	24%	12%	6%	0%	0%	0%
DMU 24	22%	18%	15%	22%	18%	15%	22%	26%	22%	0%	0%	0%
DMU 25		0%	23%		0%	23%		0%	23%		0%	0%
DMU 26	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
DMU 27			58%			39%			61%			0%
DMU 3	30%	38%	27%	10%	14%	18%	22%	25%	16%	0%	0%	0%
DMU 4	23%	16%	9%	23%	17%	14%	23%	16%	9%	0%	0%	0%
DMU 5			53%			52%			52%			0%
DMU 6			11%			11%			11%			5%
DMU 7	6%	0%	8%	6%	0%	2%	5%	0%	2%	0%	0%	0%
DMU 8	28%	17%	14%	24%	17%	12%	47%	33%	35%	0%	0%	0%
DMU9		40%	30%		35%	30%		35%	30%		0%	0%

Table 31 - Store potential gains for input and output variables

C.6 Characterization of each of all 27 stores regarding their District, Municipality and Region of Continental Portugal

Table 32 shows a characterization of the location of all 27 DeBorla stores: Municipality, District and Region in Continental Portugal

DMU	District	Municipality	Region
DMU 1	Faro	Albufeira	Algarve
DMU 10	Braga	Vila Nova de Famalicão	North
DMU 11	Castelo Branco	Fundão	Centre
DMU 12	Guarda	Guarda	Centre
DMU 13	Faro	Lagoa	Algarve
DMU 14	Porto	Maia	Metropolitan Area of Porto
DMU 15	Porto	Matosinhos	Metropolitan Area of Porto
DMU 16	Viana do Castelo	Monção	North
DMU 17	Setúbal	Montijo	Metropolitan Area of Lisbon
DMU 18	Aveiro	Ovar	Centre
DMU 19	Porto	Porto	Metropolitan Area of Porto
DMU 2	Aveiro	Aveiro	Centre
DMU 20	Leiria	Porto de Mós	Centre
DMU 21	Santarém	Santarém	Alentejo
DMU 22	Setúbal	Setúbal	Metropolitan Area of Lisbon
DMU 23	Lisboa	Sintra	Metropolitan Area of Lisbon
DMU 24	Faro	Tavira	Algarve
DMU 25	Viseu	Tondela	Centre
DMU 26	Viana do Castelo	Viana do Castelo	North
DMU 27	Viseu	Viseu	Centre
DMU 3	Braga	Braga	North
DMU 4	Porto	Vila Nova de Gaia	Metropolitan Area of Porto
DMU 5	Castelo Branco	Castelo Branco	Centre
DMU 6	Vila Real	Chaves	North
DMU 7	Coimbra	Coimbra	Centre
DMU 8	Porto	Valongo	Metropolitan Area of Porto
DMU 9	Évora	Évora	Alentejo

Table 32 – Store location characterization: District, Municipality and Region of Continental Portugal