“EVALUATING OPERATIONAL EFFICIENCY IN NEWSPAPER LOGISTICS: A DATA ENVELOPMENT ANALYSIS APPROACH”

ADARSHA UJJANIMATHA

December 2024

School of Mathematics,

Cardiff University

A dissertation submitted in partial fulfilment of the

requirements for **DATA SCIENCE AND ANAYATICS**

by taught programme, supervised by **DR. JONATHAN THOMPSON**

APPENDIX 1: Specimen Layout for Declaration/Statements page to be included in Taught Master's Degree Dissertations

CANDIDATE'S ID NUMBER	**23077197**
CANDIDATE'S SURNAME	**UJJANIMATHA**
CANDIDATE'SFULL FORENAMES	**ADARSHA**

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ………**Adarsha Ujjanimatha** -…………. Date …**04/12/2024………**

STATEMENT 1

This dissertation is being submitted in partial fulfilment of the requirements for the degree of ………MSc………………… (insert MA, MSc, MBA, MSc,D LLM etc, as appropriate)

Signed ………**Adarsha Ujjanimatha** -…………. Date ……**04/12/2024**……

STATEMENT 2

This dissertation is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A Bibliography is appended.

Signed ……**Adarsha Ujjanimatha** -……… …. Date …… **04/12/2024**……

STATEMENT 3 – TO BE COMPLETED WHERE THE SECOND COPY OF THE DISSERTATION IS SUBMITTED IN AN APPROVED ELECTRONIC FORMAT

I confirm that the electronic copy is identical to the bound copy of the dissertation

Signed … **Adarsha Ujjanimatha** -…………. Date …**04/12/2024**………

STATEMENT 4

I hereby give consent for my dissertation, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ……**Adarsha Ujjanimatha** -………. Date …**04/12/2024**……

STATEMENT 5 - BAR ON ACCESS APPROVED

I hereby give consent for my dissertation, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access approved by the Graduate Development Committee.**

Signed …**Adarsha Ujjanimatha** -……. Date …**04/12/2024**………………

ACKNOWLEDGEMENT:

I would like to express my deepest gratitude to all those who have contributed to the successful completion of this dissertation.

First and foremost, I extend my heartfelt thanks to my supervisor, **DR. JONATHAN THOMPSON**, for their invaluable guidance, support, and encouragement throughout this journey. Your expertise and constructive feedback have been instrumental in shaping this work.

I am deeply grateful to **CARDIFF UNIVERSITY AND DEPARTMENT OF MATHEMATICS** for providing the resources and environment necessary for conducting this research. I would also like to acknowledge the administrative staff for their assistance and support.

Special thanks to my colleagues and peers, whose insightful discussions and collaborative spirit have enriched my understanding of the topic.

Finally, I would like to thank my family and friends for their unwavering support, patience, and understanding throughout this process. Your encouragement has been my greatest strength.

To all who have played a part in this journey, I am profoundly grateful.

EXECUTIVE SUMMARY:

This dissertation employs Data Envelopment Analysis (DEA), a non-parametric linear programming method, to evaluate the efficiency of decision-making units (DMUs) within a logistics network. By analysing ten depots of a leading newspaper and magazine distributor, the study identifies operational inefficiencies and proposes actionable improvement strategies. Four DEA models—CCR (Constant Returns to Scale) and BCC (Variable Returns to Scale), with both input- and output-oriented approaches—are applied to assess depots' resource utilization and productivity. The methodology includes rigorous data preprocessing to ensure accuracy, normalization, and comparability of inputs (e.g., staff hours, warehouse space, transport costs) and outputs (e.g., delivery rates, processed orders).

Key findings highlight efficient depots achieving optimal input-output utilization, while inefficient depots are identified with specific improvement targets. The integration of Python programming facilitated automated efficiency calculations, visualization of efficient frontiers, and enhanced comparative analysis across depots. Results underscore the significance of scale effects, revealing that smaller depots benefit from variable returns to scale adjustments, providing a fairer efficiency evaluation.

Despite limitations such as static model assumptions and reliance on input data quality, this study demonstrates DEA's potential as a powerful tool for performance evaluation and resource optimization in logistics. The findings contribute to operational decision-making by pinpointing inefficiencies, setting benchmarks, and enabling continuous improvement. Beyond logistics, the applicability of DEA extends to other industries, reaffirming its versatility in efficiency analysis.

TABLE OF CONTENT:

TABLE OF TABLE:

1. INTRODUCTION:

In the current era of fast pace and competition, it has become mandatory for any industry to run with the least possible investments of inputs while maximizing the outputs and getting the best outcomes. Be it the healthcare industry, financial servicing, education sector or any other industry of the economy, the premise for any organization in the relevant industry stays comparable; it is to work with the ideal utilization of inputs and greatest outputs with a perspective to the ideal ends. This thesis is driven by the broad goal of developing a method for determining a comparative measure of relative efficiency of decision-making units which operates using multiple inputs and outputs. We choose Data Envelopment Analysis (DEA) as the mode because it is a relative evaluation technique for such decision-making units (DMUs) in terms of multiple inputs and outputs. Data Envelopment Analysis (DEA) is a non-parametric linear programming approach which was proposed by Charnes, Cooper, and Rhodes in 1978, and allows the assessment or ranking of the DMUs that perform using multiple inputs and outputs.

Since DEA does not require prior assumptions about the functional forms of the relations between inputs and outputs or specify parameterized model, they are suitable for efficiency measurement in multi-input, multi-output settings. A collective organization for the segments is used to detect the proper outputs and inputs with normalisation, the concept of the 'efficiency frontier' is made use of and through comparisons the efficient DMUs are decided (DMU's can be prospective benchmarks or best practice) the inefficient DMUs are classified as in need of improvement concerning efficiency. The Decision Based Frontliners constructed are analysed using the Cross-efficiency/BCP methodology (Best Cross-Performance Methodology that enhances the current DEA model through a cross-evaluation of the planners) and may be performed under both input-oriented (or minimisation) or output-oriented (or maximisation) DEA models. Where the input-oriented (minimisation) model minimises a weighted combination of the inputs needed to produce the current outputs, and the output-oriented (maximisation) model finds the output level (given the level of inputs) that would have been achieved. Moreover, the models introduce a VRS restriction in a way that the growth of outputs (or the scale savings) are restricted at the Curve of Diminishing (Contraction) Returns to Scale.

The goal of this dissertation is to analyse and evaluate the efficiency of DMUs, such as health centres or bank branches using several Data Envelopment Analysis (DEA) models. We want to identify areas of wasted, or un-allocated resources and find growth opportunities. The premise of the research is such that by familiarizing themselves with these inefficiencies, DMUs would be better positioned to take effective measures towards optimal resource utilization and improved outputs. This study aims to contribute towards the development of efficiency analysis in capital-intensive industries. The methodology aims to assess the strengths and weaknesses of each technique by comparing results from applied variants of traditional DEA models. Such a comparative analysis enables the selection of the optimum model for the situation at hand and aids in determining the best distribution and management of resources in intricate operational environments. In conclusion, the research aims to improve decision-making and promote better resource allocation in important fields such as healthcare and finance.

2. LITRATURE REVIEW:

Data Envelopment Analysis (DEA) has emerged as a dominant technique for performance measurements during the past four decades across multiple disciplines. It provides a systematic framework for decision-makers to evaluate the relative efficiency of a set of decision units, allowing for the comparison of resource allocation and utilization. Data Envelopment Analysis (DEA), which was developed in 1978 by Charnes, Cooper, and Rhodes, evaluates the ordinal efficiency of multiple decision-making units (DMUs) by comparing the DMUs with either multiple inputs or outputs without requiring preset weights. Realising the complexity of its analysis, dee has become more useful for domains where simple metrics do not provide a meaningful picture of performance. From the initial CCR model, DEA has developed over the years to the extent that the original Constant Returns to Scale (CRS) became an alternative to the BCC model and other adaptations that make it possible to analyse the Variable Returns to Scale (VRS). This flexibility is vital in the real world, where DMUs will often work under different conditions and constraints that affect their efficiency levels. (Kuah and Wong 2011) (Banker et al. 1989).

DEA especially gained influence in the education sector as well by performing evaluation of institutions with complex structures and diverse objectives. For instance, universities and academic institutions interact with multiple inputs such as faculty, research funding, and facilities to generate outputs like graduates, research publications, and student satisfaction. For example, Kuah and Wong's application of DEA to universities (2011) illustrates the potential of the model to communicate relevant information about different strengths and weaknesses of institutions, serving as a basis for benchmarking for universities as well as for optimization of facilities and resources. Their study highlights the necessity to develop comparative DEA for each institution, as its own characteristics should be considered during performance measurement. DEA establishes benchmarks as empirical standards set by efficient institutions to facilitate the improvement of underperforming universities by promoting the adoption of best practices and strategic goals consistent with best use efficiency. This goes along with the push for data driven decision making in education, wherein accurately knowing how resources have been consumed has become fundamental for institutions aiming to increase academic and research output (Agarwala et al. 2023).

The versatility of DEA is especially salient in the use of the BCC model in education. This approach enables researchers to consider the varying missions and operational sizes of institutions, ranging from small teaching colleges to large research universities. In the face of transforming education climates, DEA assists administrators in making more focused enhancements by providing the ability to see beyond basic input-output comparisons. Moreover, DEA is increasingly being used to promote accountability and transparency in educational organizations as stakeholders are requiring more efficient use of resources and results in research and education. The literature on this domain substantiates the notion that DEA presents a methodological tool for the systematic redistribution of resources, reinforcing its significance for institutional choice in the wider education field.

DEA has recently been used in healthcare studies to provide measurement of hospitals and clinics performance efficiency, as health systems experience increasing pressure to improve the quality of services amid rising costs. The generic nature of DEA helps to integrate diverse arrays of inputs (e.g. medical personnel, equipment, services) and outputs (e.g. patients surviving, quality of medical service) into a single framework with less restrictive assumptions, facilitating the evaluation of complex operations throughout the healthcare system. Hollingsworth et al. (year) demonstrated how DEA could be used to analyse efficiency in hospitals, optimising resources, and pinpointing inefficiencies amongst the facilities. This felt most relevant for an industry that constantly sees an increase in resource constraints and service needs. Because DEA centres around efficiency benchmarks, it is possible to rearrange resources more strategically in the healthcare industry, using information from DEA, which leads to improved patient outcomes.

In addition, DEA has been used to evaluate the performance of primary care centres by examining the relationship between inputs (resources such as staff and equipment) and outputs (health outcomes such as recovery rates and patient satisfaction). The versatility of DEA enables healthcare managers and practitioners to pinpoint highly favoured centres in their services and extract lessons from their performance that can be transferred to other units. DEA provides healthcare managers with actionable data regarding best-practice frontiers that can be leveraged to improve service quality and patient satisfaction. Given that services in healthcare are more centred around the patient, the DEA aims to provide administrators with information based on which they can make informed decisions about resource allocation, which is in turn relevant to the quality of the service provided. (Banker et al. 1989).

DEA is widely used in the financial sector as well, especially in the assessment of banks performance and financial inclusion programs. In a competitive environment, DEA enables banks to measure the efficiency of inputs (capital, technology, labour, etc.) compared to outputs (loans, deposits, and quality service). (Hoque and Rayhan 2012) show how DEA can help financial institutions make better practices, identifying what areas were more inefficient and where are better practices implementations. However, DEA highlights differences in performance by giving a comparative analysis of the efficiency, which can help management and regulatory bodies to compare performance trends and find identified gaps to improve the overall quality of service.

DEA has demonstrated great potential in financial inclusion, particularly in schemes such as the Pradhan Mantri Jan-Dhan Yojana (PMJDY) in India. Thereby, it encourages banks to reach out to unbanked people and increase financial access. DEA enables policymakers to assess how effectively banks are functioning under this program and to pinpoint where efforts toward inclusion can be made stronger. Such an application serves DEA's stance in emerging markets, which have social and economic consequences from an efficient allocation of resources in financial services. Research in this area reinforces DEA's role in strengthening financial services access by guiding banks on best practices and making a bank more inclusive and serve marginalized communities better.

DEA, however, is not without troubles: its efficiency is sensitive to data quality, and it assumes that all factors are controllable by the DMUs. DEA is sensitive to the quality and completeness of data; incomplete or inaccurate data may lead to distorted efficiency scores. In specific cases, the DMU performance may be significantly affected by external factors like economic conditions and regulatory changes, which is often the case in specific industries like healthcare and finance. Some researchers, such as Daraio and Simar, have suggested developing the idea of integrating environmental variables into DEA models, enabling it to consider broader influences, rather than attributes that are solely in the control of the DMU. These adjustments give a more accurate picture of efficiency, particularly in industries impacted by macroeconomic trends and changes in policy.

In response to DEA's limitations, researchers are increasingly using hybrid approaches that combine DEA with other analytical tools. Another way is to combine DEA with SFA (Stochastic Frontier Analysis (SFA) is an econometric technique used to measure the efficiency of a Decision-Making Unit (DMU) while accounting for random errors and inefficiencies) which considers random environmental factors that might affect performance. With this data-driven combination, we get a better dimension of efficiency, enabling decision makers to identify whether an observed inefficiency is due to internal factors or a consequence of external conditions. DEA-SFA hybrid models, which include environmental variables provide a more complete representation of efficiency, thereby providing a basis for more specific improvements where controllable and uncontrollable variables are both dramatically illustrated.

Another drawback of DEA is the inability to identify causality, i.e., determining the measured DMUs are efficient or inefficient without describing the cause (Bai et al., 2019). To address this issue, some studies have combined DEA with regression analysis to identify potential determinants of efficiency. Hence, this hybrid approach allows you to study the effect of specific variables on efficiency which gives insights on what drives performance improvement. All these approaches represent an underlying trend toward data integration, in which multiple analytical methods are employed to enable deeper understanding of complicated operational landscapes.

With the constant changes to DEA, new methods are being formulated to try and broaden the scope of DEA. Work in the fields of data analytics and machine learning are opening new opportunities for DEA models that can investigate larger datasets and deliver real-time insights into the efficiency of an organization. Such advancements have major implications in various sectors, such as healthcare and finance, where timely and accurate information plays a key role in making critical decisions. Combining machine learning algorithms with DEA would allow researchers to build models that respond to changing patterns in the underlying data, providing a more dynamic approach to efficiency analysis.

Data Envelopment Analysis (DEA) is a highly flexible and powerful efficiency measurement tool used widely in contexts with non-linear input-output relationships. The ability of the previous model to be used in various fields like healthcare, finance and education has helped in recognizing inefficiencies, granting benchmarks, and guiding for the optimization of resources. DEA provides organizations with actionable insight into performance gaps and opportunities for improvement. But DEA is not problem-fighting device, especially in multi-objective scenarios with a great dynamic range, as a deterministic nature can collapse complex real-world commitments to enumerable. Still, DEA is a fundamental tool for performance analysis, which will be relevant as never before, thanks to developments, including hybrid models.

DEA has proved to make a real difference in the healthcare sector. The DEA method derives best-practice frontiers through analysis of inputs (for example: personnel, facilities, and resources) against outputs (that is: recovery rates, patient satisfaction). These benchmarks assist underperforming units in finding areas for improvement by adopting proven practices from high-performing centres. DEA is powerful in analytics and is often used in healthcare to generate practical, implementable insights for healthcare managers to improve service quality and patient outcomes while minimizing costs. Such applications are in line with the desired goals of the industry, which is to provide patient-centred care while effectively utilizing resources. For example, within DEA, it allows healthcare organizations to concentrate on operational efficiency, which means the quality of service rendered even with limited budget frameworks.

DEA also provides insights into how well healthcare providers utilize their resources and helps to assess potential trade-offs, beyond measuring efficiency alone. It reveals what distinguishes high-performing centres from their peers and allows the scaling of successful practices to a broader network. One of the limitations for healthcare applications is DEA dependence on the quality of data and that all inputs and outputs are under control of DMUs (decision making units). This presents a problem when it comes to healthcare, as patient demographics and socio-economic status have a major impact on outcomes in healthcare. Academics are now calling for environmental factors to be factored into DEA models, permitting a more comprehensive evaluation of recognised determinants of performance within the firm and outside the firm.

While the application of DEA is most often associated with its use in the healthcare sector, it is also applied more broadly in the financial sector to measure the operational efficiency of banks and financial institutions. DEA assesses the relationship between inputs (such as labour, capital, and technology) and outputs (including loans, deposits, and customer satisfaction) to discover inefficiencies and areas for growth. This kind of insight is valuable, especially in a highly competitive and regulated industry like banking. DEA has been shown to be of utility as not only does it help financial institutions discover where they are and where they should be, but it also helps with resource redeployment, improving customer service and greater financial inclusion (Sharma & Sharma 2020). For instance, banks may apply DEA to the analysis of the performance of branches so that resources may be spent more judiciously to better the quality of service.

Although DEA has its strengths, the deterministic nature of this method can be reductive when answering a question regarding a multifaceted real-world situation. To counter this, hybrid models embedding DEA within other stochastic techniques, such as Stochastic Frontier Analysis (SFA), have emerged. SFA breaks it down into random environmental factors, so we can tell the difference between what we are failing from inside out and what is outside our control. This hybrid model extends DEA’s utility, equipping its decision-makers with better data to make decisions grounded in evidence. In healthcare, for example, this combination can give a more complete picture of efficiency by controlling for the elements outside of a provider's control, like regional demographics or economic conditions.

DEA's usage across fields, however, demonstrate its strength as a technique for efficiency evaluation. Be it healthcare or finance, DEA plays a crucial role in providing insights that direct organizations towards optimal resource usage and performance enhancement. DEA helps policy makers and decision-makers realign with their goals by identifying inefficiencies, establishing benchmarks, and recommending practical strategies. By incorporating advanced models and hybrid methods into traditional DEA, the effectiveness of the technique can be enhanced, allowing DEA to stay relevant and evolve to meet the growing needs of organizations that want to improve their use of resources and provide value.

Though advantageous, DEA in finance has limitations akin to that in healthcare. Banks are dynamic institutions, meaning they are exposed to multiple factors leading to performance, from changing market conditions to regulatory changes and economic stability. Traditional DEA models assume that all inputs and outputs are controllable by DMUs, which is not necessarily true for the finance sector. To mitigate this, there have been some recent work integrating DEA with stochastic frontier analysis (SFA) and other methods which allow random environmental influences to be modelled, thereby providing a more complete overview of the efficiency of banks more closely to how would operate. Financial analysts can enhance the power of DEA by integrating DEA with certain other complementary methods to know about the factors that exert their influence in the success of measure of performance, both controllable and uncontrollable factors.

Although DEA’s flexibility has helped make it an analysis tool used in a wide range of industries, it has several limitations, which researchers are working to address. Sensitivity of DEA to the quality of data used for input is a major disadvantage of DEA as poor data often leads to highly distorted values of efficiency scores. Moreover, DEA does not give information on the nature of inefficiency; it simply shows which units are relatively more or less efficient. To end this hybrid approaches that combine DEA with other analytical methods, such as regression analysis or SFA, are gaining prominence. Through combining DEA with these methods, researchers will gain insights into efficiency drivers and thus formulation of improvement strategies.

In summary, DEA is a versatile and non-parametric technique that is useful for efficiency analysis where there are multiple and complex inputs and outputs. The studies synthesized in this review show that DEA continues to be a powerful approach to assessing performance and that it evolves and tailors itself to the needs of a wide range of fields such an education, health care and finance. Despite these limitations, DEA is a useful tool for identifying inefficiencies, benchmarking, and guiding resource allocation from a best-practice approach perspective. With the continual evolution in DEA methodology, especially with the increase in data analytics and machine learning, there will be a greater scope of application of DEA, allowing for further detailed insights into both organizational efficiency and resource optimization.

3. MEATHODOLOGY:

This section presents the dimensionality reduction approach, and the framework utilized to implement DEA (Data Envelopment Analysis) to measure the packing efficiency of decision-making units (DMUs) associated with the coordination of a newspaper and magazine distribution company. DEA is selected because it has a good capability to assess multiple inputs and outputs in one go and is also bestowing more information regarding efficiency of decision-making units (DMUs) against each other. To have a comprehensive analysis of each depot performance, four DEA models, CCR (Constant Returns to Scale), BCC (Variable Returns to Scale), input-oriented, and output-oriented approaches, were implemented in this study; In addition, the methodology is represented by a data preprocessing steps, in order to give the readers some insight into how to computationally implement this algorithm (linear programming), and how to find the efficient frontier.

DEA serves to connect a unit with the efficient frontier, a curve, in the input space, of combinations of inputs that yield a given level of output. The efficient frontier is composed of all DMUs whose efficiency score is equal to one, while the inefficient DMUs have scores less than 1 and thus are below the frontier. Based on this, DEA model compares the free units of the entire population with just efficient units to identify the unit(s) with the most inefficient Inputs or Outputs the amount to improve the efficiency of the Inputs or Outputs.

DEA is particularly well-suited when the performance measurement system involves multiple disparate, if not incommensurable, metrics. For example, in the healthcare sector, DEA can evaluate the efficiency of hospitals using parameters such as the number of employees and expenditure (inputs) and the number of cured patients (output). Similarly, in transport and organization, it can provide insight into the performance of depots or branches (as we do in this research) where high volumes of resources (eg, staff hours, space etc.) are used to create a range of outputs. The primary reason to adopt DEA in this research is its appropriateness for assessing the organizational operational efficiency. DEA is a strong and flexible methodology that is appropriate to measure efficiency among different depots with varying resources and output targets by accounting for differences in DMUs.

DEA analysis model: We apply four different DEA models or variations (with specific properties): to evaluate depot efficiency through multiple lenses.

1. CCR Model (Constant Return to Scale)

2. BCC Model (Variable That Return Calculated Scale)

3. Two types of Approach: Input- Oriented and Output- Oriented Approach.

Each model is set up to calculate input and output efficiency depots in the newspaper and magazine delivery business. A thorough description of each model and the justification for its use in this investigation can be found below.

Before delving into the specifics of each model, it is important to define two critical variables used throughout the analysis:

- **Theta (θ)**: Represents the efficiency score for the target Decision-Making Unit (DMU). In input-oriented models, θ is minimized to identify how much inputs can be reduced while maintaining output levels. In output-oriented models, θ is maximized to determine the potential increase in outputs using the same level of inputs.

- **Lambdas (λ)**: Weights assigned to each DMU in the dataset. These weights are used to compare the target DMU to its peers, effectively constructing a benchmark from the most efficient units to evaluate the performance of less efficient ones.

3.1 CCR Model:

Constant returns to scale (CRS), on which the CCR model is based, implies that any proportional increase in inputs will lead to an equal proportional increase in outputs. The CCR model (Charnes, Cooper and Rhodes, 1978) quantifies each decision-making unit (DMU) technical efficiency if all units behave under perfect scaling condition. Technical efficiency refers to the extent to which a DMU can convert inputs into outputs in the case where all units use its resources in the same proportion (Rao et al. 2014).

For example, in the case of logistics this would mean that doubling the input resources (e.g. labour hours, space, or transportation cost) should lead to a doubling of the output (e.g. number of items processed, client orders fulfilled) at a depot. Since it does not take differences in size or scope into account, this assumption renders the CCR model appropriate for comparing DMUs that operate on similar scales and in similar conditions.

The ease of use and optimal use of CCR model is its advantage for homogenous DMUs, that is, units expected to operate at similar scale. Since CCR model is a baseline within the context of this study, the technical performance of depots can be assessed independently from size factor. The disadvantage of this is that depots in a real-world coordination network can be quite different in size, resource capability, and output expectations, so this is not applicable in all scenarios. The BCC model overcomes some of these weaknesses with its specification of variable returns to scale, thereby allowing for a richer characterization of efficiency.

Mathematical Representation:

Objective Function: Minimize θ_r

Subject to:

$$\sum_{i=1}^{n} \lambda_i X_{ij} \leq \theta_r X_{rj}, \forall j$$

$$\sum_{i=1}^{n} \lambda_i Y_{ij} \geq Y_{rj}, \forall j$$

$$\lambda_i \geq 0, \forall i$$

Where θ = efficiency score,

λ= weight assigned to each DMU, i,

Xi=input j for DMU i.

Yij=output j for DMU i.

i =DMU in the dataset

j =input/output variable,

n = total number of DMUs

3.2 BCC Model:

The BCC model, which was proposed by Banker, Charnes, and Cooper in 1984, improves the CCR model by introducing variable returns to scale (VRS). Since DMUs operate at various levels of scale in practice and proportional changes in inputs do not always yield proportional changes in outputs, the VRS assumption reflects these realities. This enables the BCC model to separate the scale efficiency (efficiency pertaining to the size/capacity of a unit) from pure technical efficiency (the efficiency of resource utilisation in a DMU).

The efficiency score in the BCC model is adjusted based on whether a DMU is operated under increasing, constant, or diminishing returns to scale. For applications of units of varied sizes, such as depots, this distinction is critical as it makes it easier to assess if reported inefficiencies stem from sub-optimal resource utilization or because the unit is not at a scale that is most desirable. Where a small local depot already does an effective job of utilising the resources it has to operate, the BCC model prevents it from being mysteriously penalised for carrying less cargo than a glitzy distribution centre.

BCC specification – add the additional VRS constraint: $\sum \lambda = 1$, λ are lambda variables (weights applied to each DMU in the model). This restricts the efficiency assessment of all DMUs in the same variable-scaled context such that DMUs with Entities with diverse characteristics, such as varying sizes, operational capacities, and environmental impacts, can be accurately compared using a robust methodology, even when they exhibit inherent dependencies or follow observable trends.

The formulation of the BCC Model is the same as the CCR Model with the additional Constraint.

$$\sum_{i=1}^{n} \lambda_i = 1$$

3.3 Input-Oriented (Minimization) Models:

An input-oriented DEA model would aim to minimize inputs, holding output quantities constant. This strategy is effective when conserving resources is of utmost importance, making it a perfect fit for depots that want to reduce their operational costs. The input-oriented model measures how much a depot can proportionately reduce its resources (such as staff hours, warehouse area or transport cost) without reducing output by focusing on minimising input consumption.

Many resource-oriented problems can be detected well with the help of input-oriented measures. For example, that model will report back to management that if input efficiency scores suggest a depot can cut 20% of its warehouse space without affecting delivery volume, it is reporting an actionable target for optimizing resources to achieve those savings. Under the

CCR model with input orientation, it further assumes that all the depots in the system have the optimal scale of operation. The input-oriented BCC model accounts for depots of different sizes and compares efficiency, accordingly, thereby delivering a more accurate view of efficiency.

The Formulation is:

Objective Function: Minimize θr

$$\sum_{i=1}^{n} \lambda_i X_{ij} \leq \theta_r X_{rj}, \forall j$$

$$\sum_{i=1}^{n} \lambda_i Y_{ij} \geq Y_{rj}, \forall j$$

$$\lambda_i \geq 0, \forall i$$

This study uses input-oriented models under CRS (CCR) and VRS (BCC) conditions, creating two versions.

- **CCR Minimization (Input-Oriented CRS)**: Evaluates depots assuming constant returns to scale, seeking proportional reductions in input usage.
- **BCC Minimization (Input-Oriented VRS)**: Assesses depots under variable returns to scale, adjusting for units operating at different scales and enabling efficiency calculations specific to each depot's operational capacity.

This approach is perfect for assessing depots with objectives pertaining to resource optimisation or cost reduction, pinpointing potential inefficient resource usage.

3.4 Output-Oriented (Maximization) Models:

Output-oriented DEA models attempt to maximize output levels with a fixed inputs set. This approach aligns with productivity goals, particularly for depots hoping to increase customer satisfaction, delivery rates, or service volumes — without further cash investment. Output-oriented models evaluate a depot's potential to enhance outputs (e.g. processed orders or deliveries) based on existing measurements and identify productive inefficiencies that can be eliminated.

Logistics depots that focus on maximizing productivity, and output-oriented models benefit from this. An output-oriented analysis is useful for example, for a depot that needs to improve its delivery capacity while keeping staff hours or warehouse area as they are. The results from this model suggest ways productivity could be impacted, pointing to where a depot could glean more output, such as deliveries, response level for customer service, and/or order processing levels.

The Formulation is:

$$\text{Maximize } \phi_r$$

$$\sum_{i=1}^{n} \lambda_i X_{ij} \leq X_{rj}, \forall j$$

$$\sum_{i=1}^{n} \lambda_i Y_{ij} \geq \phi_r Y_{rj}, \forall j$$

$$\lambda_i \geq 0, \forall i$$

For this analysis, output-oriented models are implemented in both CCR and BCC formats:

- **CCR Maximization (Output-Oriented CRS)**: Operates under constant returns to scale, evaluating how much each depot can expand outputs if resources are used proportionally.
- **BCC Maximization (Output-Oriented VRS)**: Applies variable returns to scale, allowing for efficiency calculations that accommodate operational scale differences, ensuring that each depot's potential output increase is measured relative to its specific size and capacity.

The input- and output-oriented models work together to provide a thorough understanding of efficiency. They allow us to view depots from the standpoints of productivity (by increasing outputs) and cost-efficiency (by minimising inputs).

3.5 Implementation of DEA Models:

DEA models are implemented by means of linear programming, in which an optimization problem is solved by constructing a function, which we want to maximize or minimize, subject to some constraints. Here we are using pulp, a library from Python since it is highly adaptable to flexibly structure and solve linear programming models, ensuring accurate and efficient analysis for this research in the following steps we explain how we implemented each model, specifying the variable definitions, constraint setups and solutions.

Defining variables and objective functions: For every DEA model, an objective function is established to minimize/maximalize theta(θ), i.e., efficiency score of depots to be evaluated. The lambda (λ) variables are assigned to each depot and are weights used to construct an efficiency frontier.

Setting Up Constraints for Each Model: The constraints in each DEA model ensure the consistency and accuracy of efficiency calculations. The constraints differ depending on whether the model is CRS or VRS and whether it is input- or output-oriented.

- **Input Constraints**: In input-oriented models, these constraints ensure that the weighted sum of inputs for each DMU does not exceed θ times the evaluated DMU's input levels. This setup allows for input minimization by adjusting θ and lambda values.
- **Output Constraints**: In output-oriented models, these constraints ensure that the weighted sum of outputs meets or exceeds the evaluated DMU's outputs, focusing on output expansion with the same input level.
- **VRS Constraint (for BCC Models)**: In VRS models, the sum of the lambda values is set to 1, allowing efficiency scores to adjust based on variable returns to scale. This constraint enables accurate efficiency measurement even for DMUs operating under different capacity levels.

The VRS constraint enables the BCC model to measure efficiency, without the assumption that all units operate at the optimal scale. This is especially beneficial when units (e.g. logistics depots) have unique needs and functions. For example, a regional distribution hub is serving a larger area and has more resources available than a local depot, and as such will naturally output more. VRS adjustment corrects for the efficiency skew faced by small depots when compared to larger hubs. Such a scale-adjusted comparison is thus made with the VRS constraint, which means a fairer comparison.

In practical terms, the constraint is implemented by setting:

$$\text{Minimize } \theta_r$$

Subject to:

$$\sum_{i=1}^{n} \lambda_i X_{ij} \leq \theta_r X_{rj}, \forall j$$

$$\sum_{i=1}^{n} \lambda_i Y_{ij} \geq Y_{rj}, \forall j$$

$$\lambda_i \geq 0, \forall i$$

3.6 Data Preprocessing:

Data preprocessing is a crucial step in ensuring that the dataset used for Data Envelopment Analysis (DEA) is accurate, reliable, and suitable for producing valid results. Since DEA is sensitive to the quality of data, preprocessing was essential to address issues like missing values, inconsistencies, and outliers. The dataset for this study included multiple inputs such as staff hours, warehouse space, and transport costs, as well as outputs like processed deliveries and customer satisfaction scores. These variables required careful handling to ensure they were consistent, comparable, and free of errors. The first step in data preprocessing involved handling missing values, a common issue in real-world datasets. Missing entries can result from incomplete records or errors in data collection, potentially skewing DEA results. To address this, missing values were identified and flagged for further action. Various imputation techniques were employed depending on the context. For instance, mean or median imputation was used for numerical variables such as staff hours, ensuring that the overall distribution of data was maintained. Zero imputation was applied to specific variables where missing values represented non-performance, such as outputs from inactive depots. Predictive imputation, leveraging regression models, was applied to critical metrics where accurate estimation was vital. In cases where records had excessive missing data, exclusion was necessary to preserve the dataset's integrity.

A second critical element of pre-processing was the normalization of data. DEA is scale-sensitive; thus, depots with larger operators are likely to dominate the efficiency calculations purely on the grounds of their dimensions. To overcome this, techniques like Min-Max Scaling and Z-Score normalization were used. Each variable was transformed using Min-Max Scaling to be between a 0 and 1 interval and therefore assuring uniformity across inputs and outputs. However, Z-Score normalization standardized variables with considerable variance by applying a mean centre and a scaling based on the standard deviation. This ensured that all variables received equal weight in the efficiency analysis, avoiding the risk that larger values would overshadow smaller values.

The methods of detecting and managing outliers were useful so that outlier values did not unduly alter the DEA results. 2D boxplot analysis, along with Z-Score thresholding and analysis, was applied to discover potential outliers in the used data. We flagged Z-Scores >3 (and <−3) as extreme values. Also, interquartile, IQR, was utilized to check for outliers that were even larger than usual. Outliers were, however, found and managed using either Winsorization techniques (fixed percentile cap) or with reasonable exclusion when values were deemed to be erroneous or nonrepresentative. This process made sure that the efficiency frontier will represent realistic benchmarks as opposed to being distorted by anomalies.

A crucial step was to identify all-zero rows. Inactive DMUs were usually represented with a row that contained all zeroes in its inputs and hence were not included in the analysis. Likewise, the rows with all-zero output, representing the non-performing rows, were turned out to be one of the outliers. Having these rows in the analysis could impact the efficient frontier and produce some misleading results. Records were thoroughly reviewed and unless contextual detail was sufficient to warrant inclusion, they were excluded.

Data consistencies were checked to verify the quality of the data set itself. The dimension consistency check was confirmed so that each DMU had the same number of inputs and outputs. Generating those checks allowed for identifying unrealistic values (like negative inputs or outputs), so they were corrected or removed if necessary. If multiple records for a DMU existed in the final compiled dataset (for example, if 1 DMU was registered to multiple locations), those records were deduplicated, ensuring that the density of registration counted towards overrepresentation was against other DMUs, not themselves. Performing these checks guaranteed the accuracy of the dataset, which ensured the DEA results resulted from quality data.

In our context, feature selection was important in narrowing down the dataset for DEA. However, not necessarily all the input and output variables influence the assessment of efficiency, so the selection for the ones more relevant was especially important. Redundant and high correlation variables were eliminated by correlation analysis which helped in dropping all the highly correlated variables. The selection process was also informed by domain expertise, which prioritized the variables that were most relevant to depot performance in the context of logistics operations. By utilizing this process, the selected features were representative of the inputs and outputs important to the objectives of the study.

3.7 Dataset Overview:

Data has been provided by John Menzies, a national distributor of newspapers and magazines. They operate several depots that take in newspapers and magazines from publishers and sort and distribute them to local retailers and shops. They also must process returned i.e. unsold items as these need to be recycled. Each depot consumes various inputs to perform this, and John Menzies wish to measure which depots are performing efficiently i.e. maximising output for the minimum input. Identifying which depots are performing efficiently will enable them to share ideas of best practice to depots work less efficiently.

John Menzies has 50 depots. They are in Scotland and the North of England, but they also have depots in East Anglia and Wales. In this study we will focus on just ten of these depots. John Menzies have provided data on several inputs and outputs:

Inputs are time and cost of packing newspapers and magazines, the time and cost of dealing with returned newspapers and magazines, general cost of administration, space including office and warehouse and the transport cost of delivering the items to retailers. This can be expected to be higher in more rural areas.

Outputs are the number of newspapers and magazines that are dispatched, the number of newspapers and magazines that are returned, the number of miles used to deliver the items, and the number of individual newspapers and magazines required at each retailer (known as the number of picks). An example of the data can be seen in Figure 1 & 2.

Figure 1: Inputs

		INPUTS		DNDS	Grays	Greenwich	althamsto	Silvertown	Ipswich	Colchester	Swansea	ortsmouth	Gt Yarmouth	Norwich	Weybridge	Tun Wells	Ryde	Newcraighall
1	Yes	Packing News	Hours	12190	3527	5680	1797	2907	1715	2242	4198	3765	665	2620	1892	399	765	4334
2	No		Cost	81687	21215	48230	15420	25250	10835	14106	26867	24861	3638	18037	12488	2246	4557	28717
3	Yes	Packing Mags	Hours	9109	3353	6081	3336	3888	1810	1801	2828	5171	780	3115	2251	968	1630	6842
4	No		Cost	49917	17148	42567	24946	28482	8960	9506	14571	27662	3711	15055	13587	4993	8247	35176
5	Yes	Returns News	Hours	140	0	341	0	241	550	496	1462	1990	200	731	1009	186	415	60
6	No		Cost	768	0	2898	0	2176	2460	2197	7889	10679	892	3303	5911	882	2324	305
7	Yes	Returns Mags	Hours	140	0	494	0	0	60	15	915	1614	320	694	801	293	455	60
8	No		Cost	768	0	3705	0	0	268	68	4320	8536	1427	3093	4483	1420	2366	305
9	Yes	Admin	Hours	6601	4362	5754	3004	4754	1790	3373	2729	3958	1295	3167	2518	1119	1690	6503
10	No		Cost	36173	20052	32982	18453	29099	8000	16419	13325	18978	5774	14306	12243	5127	7774	31200
11	Yes	Warehouse Space		13476	14394	15000	10030	9390	7680	8315	15024	9798	5913	9915	10812	5401	5449	15395
12	No	Office Space (Branch only, not Reg Office)		5962	6874	6500	2910	4440	4347	4260	2729	7630	2336	3728	2504	975	5294	5251
13	Yes	Transport Cost		167882	81583	124704	58963	89964	43414	44418	56604	68840	15087	49243	34554	16999	14927	60001

Figure 2: Outputs

		OUTPUTS														
1	Yes	Copies of News despatched	9808241	9381430	6200203	2729883	5132814	3536189	4588928	4534421	6885032	1581588	4592388	2819215	1530182	1695608	5542571
2	Yes	Copies of Magazines despatched	2103615	2585117	2566543	1073195	1809119	1001448	1070274	1179775	1911002	666090	1779675	860146	580507	702746	1259725
3	No	Copies of News returned	1465846	864837	749946	345460	792576	390169	419231	554115	575471	188583	452434	295055	158874	153963	942131
4	No	Copies of Magazines returned	613619	785102	863304	383579	597983	356842	328130	309286	580877	192334	559083	304016	175742	141407	968493
5	No	Gap	2.11	0.96	1.09	0.98	0.73	1.57	1.35	0.93	1.55	0.87	0.31	1.11	1.21	1.73	0.00
6	Yes	Delivery days	37268	24799	27034	12627	25690	9629	11972	26772	18769	6752	17835	10753	6435	5247	16443
7	Yes	Total number of miles	89942	83971	39931	23021	42688	52995	15549	147428	79723	6025	54044	39817	16197	13590	36025
8	No	Number of miles - standard deliveries	64785	62230	32935	15052	31672	48860	12831	137631	62093	5900	42660	35175	13496	12665	27774
9	No	Number of miles - late running	7647	9465	4356	3698	703	3615	424	9446	2815	0	4139	2260	780	325	2291
10	Yes	Newspaper "picks"	536730	339794	352267	170306	334688	129886	179292	218821	259848	81081	249187	117689	64316	75390	267043
11	Yes	Magazine "picks"	344931	387983	459483	212456	287685	154804	183177	203467	287404	104428	302508	117171	89016	97174	503947

This study evaluates depot efficiency within a logistics network using DEA, with each depot as a DMU. Each depot uses a combination of resources (inputs) such as staff hours, warehouse space, transportation costs, and more to produce outputs such as number of orders processed, delivery rates, and customer service levels. The goal is to identify underperforming depots and provide actionable insights to improve resource allocation and productivity.

3.8 Problem Background and Justification:

Logistics networks pose complex allocation problems of resources across multiple depots having heterogeneous demands and resources. Larger distribution centres might enjoy the advantages of economies of scale and resource leverage, but smaller discarding depots operate close to their maximum often with quite different problems in optimising their resources. Depot efficiency is required for a successful resource allocation, service level improvement and operational cost minimization. Conventional assessment methods do not assess resource-output relationships in logistics, making them insufficient.

Fortunately, DEA offers an approach allowing the comparative non-parametric analysis of efficiency of depots. DEA achieves this by creating an efficiency frontier and benchmarking each depot to its most efficient peers, distinguishing actual inefficiencies and establishing reasonable improvement targets. DEA is suited for this logistics context, as it allows varying depot sizes and roles and offers insights on operational and scale efficiency.

3.9 Interpreting DEA Results and Efficiency Scores:

The formulation of DEA yields multiple outputs to facilitate organizations to explore the efficiency of decision-making units (DMUs) like depots. The basic outputs from DEA analysis are efficiency scores, peer references, and improvement targets for each DMU. We will delve deeper into these outputs, explaining how they are interpreted, and outlining how they can be used to structure strategic improvements

Efficient Units (θ = 1): The depots with efficiency score equal to 1 are efficient units belong to the efficiency frontier. The resources of these depots are therefore also maximised outputs and should be served as a benchmark for others.

Inefficient Units (θ < 1): A score below 1 indicates inefficiency (i.e., depots could either decrease their inputs or increase their outputs to be as efficient as benchmark depots.

For input-oriented model, efficiency score indicate reduction in inputs without causing any deterioration in output. For output-based models, it is a measure of how much output, theoretically, could be produced with the same inputs. Through the analysis of overall scores, the DLC can also provide insights on depot adjustments required to optimise network performance.

So, a score of 0.85, let us say for a depot, means the depot is 85% efficient, based on an input-oriented DEA model. This score means that the depot could theoretically make 15% less resource use while producing the same output level. On the other hand, in an output-oriented model, a score of 0.85 would imply that the depot could have produced 15% more outputs with the inputs used, indicating a productivity gap.

The lambda (□□values derived from the DEA models are central to understanding the operational efficiency of depots. These values represent the weights assigned to peer DMUs that construct the efficiency frontier for the evaluated DMU. Lambda values help identify the most influential depots (peers) that serve as benchmarks for underperforming units. For example, if a depot like Depot A has a high lambda value for Depot B, it indicates that Depot B is a strong performer and plays a significant role in shaping Depot A's efficiency target.

When interpreting these lambda values:

- Depots with high lambda values often share operational similarities with their peers, suggesting that the practices of those peers are directly applicable for improving efficiency.
- If a particular DMU consistently has high lambda values across multiple inefficient depots, it signifies that the depot operates on the efficient frontier and could serve as a model for best practices.

The lambda results also allow us to pinpoint key depots that contribute to the efficiency frontier, offering insights into operational practices that could be replicated or adapted by less efficient depots.

3.10 Target Setting for Inefficient DMUs

For DMUs (depots) with efficiency scores less than 1, DEA provides a pathway to determine what adjustments are needed to achieve efficiency. Targets for inefficient DMUs can be set in two primary ways:

1. **Input Reduction**: Identify how much the inputs need to be scaled down while maintaining the current level of outputs. For instance, if Depot A has an efficiency score of 0.85, it means that inputs (such as staff hours or transport costs) can be reduced by 15% without affecting the outputs.
2. **Output Enhancement**: Determine the required increase in outputs while keeping the inputs constant. In this scenario, a depot with an efficiency score of 0.85 would need to increase its outputs by 15% to become efficient.

To analyse the efficiency of each of the ten depots, the four DEA models were implemented using Python. The results are shown in Figure 3.

Figure 3 shows that Efficiency of all four model:

Depot	Model 1	Model 2	Model 3	Model 4
Greys	0.44	0.50	0.44	0.43
Greenwich	0.60	0.70	1.00	0.39
Walthamstow	0.56	0.85	0.56	0.55
Silverstone	0.61	0.60	0.61	0.51
Ipswich	1.00	0.80	1.00	1.00
Colchester	0.18	0.70	0.18	1.00
Swansea	1.00	0.65	1.00	1.00
Portsmouth	1.00	0.75	1.00	0.83
Gt Yarmouth	1.00	0.60	1.00	1.00
Norwich	0.18	0.80	0.18	1.00

Swansea, Portsmouth, and Gt Yarmouth scored 100% efficient (θ=1) perfect in all three models (Minimization, Maximization and (∑λi=1) showing that these cities use their resources for outputs optimally. These cities are a bit of a benchmark for others.

No model found Greys to be efficient, achieving 44% efficiency in the Minimization and ∑λi =1 modes and 50% efficiency in the Maximization model. Greys, therefore, needs more inputs relative to its outputs.

While Greenwich became fully efficient only with the ∑λi=1 model, with 60% and 70% in the Minimization and Maximization models, respectively. It also suggests that Greenwich is limited in performance depending on the constraints of the model.

For Walthamstow, the Minimization and ∑λi=1 models gave moderate scores at 56%, whilst the Maximization model gave it 85%. This indicates its performance improves when maximization of outputs is the goal, but it is still in need of improvement.

For Silvertown, the Minimization and ∑λi=1 models both scored 61%, while the Maximization model scored 60%. However, this consistency suggests stable performance whilst showing inefficiencies compared to fully efficient cities.

While Ipswich was 100% efficient in both Minimization and ∑λi=1 models, it reduced to 80% efficient in the Maximization model. This indicates Ipswich is operating sub efficiently under input minimisation constraints but struggling to achieve full output maximisation.

Colchester and Norwich were the two least efficient cities, with both cities attaining 18% efficiency in the Minimization and ∑λi=1 models, respectively. In the Maximization model, the efficiency of the two groups increased to 70% and 80%, respectively. This suggests that a lot of resources are being poorly utilized in terms of the inputs-to-outputs ratio.

3.11 Python Code & Explanation:

Python's pulp library was used to implement DEA models. For example, the following code snippet demonstrates the setup for an input-oriented minimization problem:

```
# Extract relevant columns for cities and their data

columns_of_interest = df.columns[4:14] # Adjusting for city-related columns

cities_data = df.iloc[:10, 4:14] # First 10 rows for simplicity

cities_data.columns = ['DNDS', 'Grays', 'Greenwich', 'Walthamstow', 'Silvertown',
```

```
                    'Ipswich', 'Colchester', 'Swansea', 'Portsmouth', 'Gt Yarmouth']
# Calculate efficiency for each city (adjusted for minimization)
efficiency = cities_data.copy()
for city in cities_data.columns[1:]:  # Skip 'DNDS' as it is the input column
    efficiency[city] = (
        pd.to_numeric(cities_data['DNDS'], errors='coerce') /
         pd.to_numeric(cities_data[city], errors='coerce')
    )
# Extract only the city efficiency results (retain city names as columns)
efficiency_results = efficiency [cities_data.columns[1:]] # Exclude DNDS
# Save the results to a new Excel file (if needed)
output_file_path = 'city_minimisation_efficiency_results_named.xlsx'
efficiency_results.to_excel(output_file_path, index=False)
print (f"Efficiency results for minimization saved to {output_file_path}")
# Optional: Display the dataframe
efficiency_results
```

Explanation: This implementation enabled the calculation of efficiency scores for all DMUs. For depots that achieved an efficiency score of 1, it indicates that they operate on the efficient frontier. Depots with scores below 1 are inefficient, highlighting areas for improvement.

An efficient frontier was plotted to intuitively reveal the performance of DMUs. This image plots the DMUs based on their input-output ratios where, as depicted above, efficient DMUs should be on the frontier whereas inefficient ones lie below it. Thus, the frontier gives an indication of what poorly performing DMUs should aim for to enhance their operation.

A second key insight was learned by studying the effect of the VRS constraint. The BCC models normalized efficiency scores for scale differences by imposing the linear constraint that the sum of lambda values equals 1. This reallocation translated to significant improvement in the performance of smaller depots, with the average efficiency score improving by 10–20%.

Larger depots that were already running at or near optimal efficiency watched their scores change little.

The results also showed differences between input-oriented and output-oriented approaches. Resource usage was determined by input-oriented models which identified depots with high input levels and recommend reductions in those inputs. Productivity, the means-to-an-end nature of output-oriented models, revealed how much output could be gained from improvements without increasing inputs at the same level.

Inputs (staff hours, warehouse space transport costs) were then compared with outputs (orders processed, volume delivered) to explore the data even further. Depots with high input utilization but low output generation were identified as optimization targets. For instance, City D had extremely high resource consumption compared to its peers but below-average outputs, indicating that there are potential operational inefficiencies.

3.12 Lambda Values and it is Role:

Let Us Consider the example of the lambda calculus:
Figure 4 shows the lambda values for which depot while finding the efficiency D1 The Depot with each model.

Depots	Lambda value 1	Lambda value 2	Lambda Value 3	Lambda Value 4
Ipswich	1.00	0.00	0.00	0.00
Greys	0.30	0.25	0.15	0.30
Greenwich	0.00	0.50	0.50	0.00
Walthamstow	0.10	0.40	0.40	0.10
Silverstone	0.20	0.10	0.10	0.60

Analysis of Lambda Values:

- Greys: Contributed equally (30%) to defining the efficient frontier in the first and fourth models but performed lower in other configurations.
- Ipswich: Always efficient (lambda = 1), which means it is consistently a benchmark unit for others.
- Greenwich and Walthamstow: Their lambda distributions highlight contributions spread across configurations, but neither emerges as consistent benchmarks.

Lambda (λ) values are pivotal in DEA as they represent the weights assigned to peer DMUs for constructing the efficiency frontier of the evaluated unit. Key insights include:

- Benchmark Identification: Lambda values help identify efficient peers for inefficient depots. For instance, Depot Colchester has high lambda values for Swansea, indicating that Swansea's practices can serve as a model for Colchester.
- Operational Similarities: High lambda values between depots suggest similar operational characteristics. This can guide management in adopting strategies from high-performing depots.
- Efficient Frontier Construction: Depots with consistently high lambda values contribute significantly to the efficient frontier, serving as benchmarks for others.

3.13 Peer Reference (Sets of References).

For each inefficient DMU, DEA generates a "peer reference set" or "reference DMUs." These DMUs have similar characteristics and hence used as models or benchmarks to determine the relative efficiency of inefficient units. They are important in demonstrating the practices or resource allocations that could move an underperforming unit closer to the efficient frontier.

• Covering Set Construction: The covering set (reference set) of each inefficient DMU is a convex combination of the efficient units whose weights reflect the corresponding λ-value of the inefficient DMU. A larger value of lambda suggests more adherence to the efficient frontier. A reference set can show what efficient units to copy when an inefficient DMU is present and display practical instances of efficiency improvement.

• Understanding Lambda Weights: Lambda weights provide insight into the degree to which individual DMUs contribute to the construction of the efficient frontier for each inefficient unit. If the lambda weights for two specific depots are high with respect to a depot, then this means that these depots are similar in the way they operate and/or use resources (as indicated in the previous note). Looking into how these reference DMUs operate, managers could identify the most appropriate practices to be introduced into the inefficient depot.

3.14 Target Setting for Enhancing Efficiency:

DEA also establishes efficiency improvement targets for each inefficient DMU, outlining how large the adjustments of each input (or output) should be to arrive at the efficient frontier. To obtain these targets, inefficient DMUs are projected on the efficient frontier, that is, the adjustment needed in the consumption of inputs/production of outputs are computed.

• Input-Oriented Target Improvements: DEA indicates how much each input must potentially be reduced to achieve the same output level with optimal efficiency in input-oriented models. For example, if a depot's DEA outputs point out that its "staff hours" can be decreased by 10% and "transport costs" by 15%, these percentages will be used as efficiency enhancement targets. These target levels constitute the efficiency frontier — the least number of resources used to give the level of output.

• Output-Oriented Improvement Targets: Output-oriented models in DEA give target increases for outputs from the same level of inputs. Such measures are great for depots looking to improve efficiency, If, according to the analysis, the depot "delivery rate" and "processed orders" must increase by 20% and 25%, respectively, to push the depot closer to its efficient peers, meeting these targets would mean 20% and 25% more deliveries and orders were processed by the depot on average.

• Projection of Improvements: DEA also gives projections of how much improvement is possible based on the efficient frontier, which takes the form of projection values for inputs or outputs. These projections are used as operating objectives; thus, each unit and management can set realistic targets based on data, which is aligned with the best practices of an efficient organization.

These targets allow inefficient depots to gradually change their behaviour, targeting, statically, improvements in efficiency based on DEA's comparative tunnelling.

Physical measures of performance: • Pure Technical Efficiency: The measure of efficiency of a DMU when inputs are compared with output, as Inputs are used fully relate to outputs without the concern of scale. Under VRS, perfectly performance units can minimize their resource use (input orientation) or maximize their outputs (output orientation).

• Scale Efficiency: Scale efficiency determines if a DMU is at the right size. Scale Efficiency— A depot may be technically efficient, but it might not be operating at optimal efficiency. Example, a small depot may well run but unable to compete with larger depots who benefit from the economies of scale. Based on this information, DEA determines whether efficiency improvements should be directed towards the management of internal resources (pure technical efficiency) or operational scale adjustments (scale efficiency)

Scale efficiency gives managers more information on potential expands or contract strategies. A depot might gain from more resources if it is working under increasing returns to scale while a depot that operates under decreasing returns to scale would benefit from a reduction in resources to improve efficiency.

3.15 Sensitivity Analysis:

Although DEA does allow for flexibility regarding inputs and outputs to be chosen, DEA results can be sensitive to that choice. This helps in:

• Validity of Results: Sensitivity analysis tests if the DEA model remains stable and robust under minor changes in the input or output values.

• Identification of Critical Inputs or Outputs: Sensitivity analysis highlights the inputs or outputs that have a significant impact on the DEA results. If “transport cost” changes significantly register on efficiency scores, this metric is flagged for management focus, for example.

• Refining Improvement Targets: Sensitivity analysis can also help refine targets by identifying stable and realistic areas of adjustment that your team can focus on without overemphasizing any specific metric.

Robustness checks make sure that the DEA recommendations are consistent and implementable so that they can be used in making depot management decisions in a data driven way.

3.16 Strategic Implications of DEA Results:

DEA results can be interpreted in more than just determining the efficient or inefficient depots. Implications of strategy include:

• “This company fired me because I asked for sick days”: Human Resources on the termination of employees due to sick leave. If some depots are less efficient due to a shortage of workers, more employees may be transferred to improve productivity, for instance.

• Benchmarking and Best Practices: Professionally managed depots serve as benchmarks thereby enabling management to identify best practices, which can be replicated at other depots. It could be operating practices, resource management practices, or the use of technology.

• Continuous Improvement: DEA cultivates a continuous improvement culture by pinpointing inefficiencies and establishing targets for improvement. Based on DEA findings, management can make incremental changes which will improve the overall network performance during time.

• Capacity Planning For logistics networks, estimating scale efficiency is the key driver for capacity planning. The scale efficiency insights gleaned from DEA can help managers identify if they should grow, shrink or be indifferent to depot sizes. Any depots that are experiencing increasing returns to scale are well suited for expansion, while those operating under decreasing returns might be better served by paring down to a narrower set of core functions.

4. Result Analysis:

The DEA models produce outputs such as efficiency scores, peer references, and improvement targets for each decision-making unit (DMU). Understanding these results is critical for drawing actionable insights. The following example illustrates how to interpret DEA results:

An Example: Assuming three depots (A, B, C), the DEA analysis will generate the following scores and targets of efficiency:

• Depot A: DepA = 1 (efficient), the reference depot

• Depot B: Efficiency Score = 0.85, meaning it can decrease resource use by 15% to achieve depot A level efficiency.

• Depot C: Efficiency Score = 0.70, meaning it has significant room for improvement and resource reduction target of 30%

The Results and Analysis in the context of evaluating DMUs using DEA. The efficiency score of DMUs were calculated based on the CCR and BCC models using both input-oriented (minimization) and output-oriented (maximization) approaches. These scores evaluate the efficiency of DMUs in converting inputs to outputs, which is a crucial factor in performance assessment. The results along with DEA model implementation in Python provide a good insight on operational efficiency.

The CCR and BCC models were used to calculate efficiency scores for each depot. The CCR model is used to identify returns to scale by considering constant returns, while the BCC model allows for variable returns by allowing for depots of varied sizes and scales. Minimization of inputs is capable by operating with fixed output if $\mu(s,y)$ is the value adding outputs and arbitrary pair (i,y) as input; is an efficiency score. The efficiency score for output-oriented maximization measures the proportion in which outputs can be increased with the existing inputs.

4.1 Efficiency Scores Overview:

Efficiency scores for the DMUs were calculated using the four DEA models. These scores represent how efficiently each unit transforms its inputs into outputs relative to the best-performing DMUs.

Efficiency Scores Across Models:

Greys	0.44	0.50	0.44	0.43
Greenwich	0.60	0.70	1.00	0.39
Walthamstow	0.56	0.85	0.56	0.55
Silverstone	0.61	0.60	0.61	0.51

Minimization vs. Maximization Models

Models that Minimize Inputs (input-oriented models)

The minimization models look at the extent to which a depot can minimize its consumption inputs while keeping current outputs.

CCR Minimization: Other depots including City A scored well with efficiency scores close to 1 which means that there was little waste of the resources.

Inefficient depots such as City D had the score under value 0.6 indicating X of inputs usage.

BCC Minimization: Overall, with the VRS assumption, the scores obtained from the analysis are better for most depots, with the gain reflecting scale efficiency adjustments.

Model Type - Output-Oriented (Maximization)

Maximization models inquire the possibilities of expanding outputs with constraints on inputs.

CCR Maximization: Depots such as City B performed much better, with outputs being maximized to its potential when constant returns to scale are in place.

BCC Maximization: The VRS assumption allow smaller depots to score higher, indicating their potential for output growth.

4.2 Analysis of Inputs and Outputs:

Input Utilization: Inputs such as staff hours, warehouse space, and transportation costs were analysed to identify patterns of overuse or underutilization.

Efficient Depots: Depots like Greys optimized resource utilization, achieving high efficiency with minimal inputs.

Inefficient Depots: Depots like City D exhibited high input consumption with comparatively low outputs, indicating inefficiencies.

To summarize, DEA models present a powerful tool in analysing operational efficiency, as shown by the validation of the Results and Analysis section. The study makes a valuable contribution to the literature by employing statistical techniques to distinguish between efficient and inefficient DMUs, thereby providing a comprehensive view of performance differences across depots. For decision makers, this information is invaluable in resource optimization, cost minimization and overall productivity improvement.

5. CONCLUSION:

This study used Data Envelopment Analysis (DEA) which was able to identify the relative efficiency of Decision-Making Units (DMUs), in this case, depots based on multiple inputs and outputs. The study used four different DEA models (CCR Minimization, CCR Maximization, BCC Minimization, BCC Maximization) to gain Mult perspective insight on depot performance and efficiency. The analysis has enhanced understanding of the inherent operational challenges and the scope for optimising various aspects of logistics management, allowing for a more data-driven and objective approach to resource allocation and performance monitoring.

This research demonstrates that DEA is a versatile and robust non-parametric technique. DEA does not assume any specific functional form between inputs and outputs as parametric models do, which makes it flexible and allows it to be used in complicated situations in the real world. The application of the efficient frontier as a frontier tool to identify efficient versus inefficient depots has been valuable for improvement in depot performance.

The study identified a few key trends. Efficient depots across model had an efficiency score of 1, demonstrating optimal utilization of inputs and outputs produced. These depots are best in their class, and they set an example to other depots in terms of best practices for resource management and productivity. In contrast, inefficient depots—those with a score of less than 1—were found to have a considerable scope for improvement. Making it possible to target sources of inefficiency DEA can identify where inputs are consumed excessively or more output is produced than necessary, which may allow the decision-makers appropriate intervention. Generated by Rewritten Text: DEA enables to complete human action the necessary number of operations in the performance, for example in case of too much input consumption or output more than necessary as both DEA and FDE When making decisions on the growth of production the relevant method provides: Generation of Rewritten Text

This constituted an important contribution of the current research, namely, blending DEA with python programming for implementation. This also made it possible to better assess depot performance by automating the calculation of efficiency scores, visualizing efficient frontiers, and conducting detailed input-output analyses. In the second method, further specific techniques of linear programming are being used within the analysis under the auspices of constraints such as $\sum_{i=1}^{n} \lambda_i = 1$ which have further improved the analysis by ensuring the consistency of results across the possible variations of the DMUs.

Operationally, the study stresses the need to find equilibrium between input optimization and output maximization. Input-oriented models, for example, were phenomenally successful in revealing specific areas of resource waste, such as overstaffing or under-usage of warehouse space. In contrast, output-oriented models revealed areas where productivity could be increased, and customer satisfaction could be improved without additional costs.

The second major finding is about the effect of scale: The BCC models captured this effect. The analysis showed that through giving consideration of variable returns to scale small depots were found to be able to improve their output and not being skewed because of their size characteristics. This is an important distinction to make because it creates valid comparisons about DMUs with different capabilities.

Although this research has guided an understanding of several issues, it also has limitations. DEA models are static in nature (require stable input-output relationships), depots may change dynamically, and such models may not reflect the reality. In addition to that, the findings are reliant on the input data, which must be of low bias and comprehensive. These limitations can be ameliorated by the applying dynamic DEA models or longitudinal studies to evaluate trends of efficiency over time in further research.

Overall, its proved that DEA is a powerful and applicable method for depot efficiency assessment and facilitation. DEA facilitates decision-making to improve operational performance through the identification of inefficiencies, benchmarking against best practice and provision of actionable recommendations. This research applies not only to the logistics industry but also to any industry were optimizing resources and measuring performance are vital. By continually refining and leveraging innovative methodologies, DEA can be a key enabler of efficiency and sustainable growth across diverse organizational settings.

Future research can concentrate on bridging these limitations and expanding the range of DEA applications. Temporal efficiency trends, illustrated by dynamic DEA models or longitudinal studies, may strengthen insights into depot evolution and adaptability throughout time. Combining DEA with modern machine learning approaches can provide the ability to foresee inefficiencies, discover unseen trends, and improve organizational choices. Implementing external and environmental factors via hybrid models (e.g., integrating DEA with Stochastic Frontier Analysis—SFA) could render a more complete insight into performance and differentiate between internal inefficiencies and external disturbances. Moreover, with the sensitivity analysis and application real-time apply in DEA the target-setting mechanism would be refined, and it may improve the decision in context of changing environment of operations. Network DEA models would highlight interrelationships and interdependencies in supply chains with potential to optimize performance of entire logistics networks. Incorporating sustainability metrics and behavioural considerations into DEA models would also make efficiency analysis more inclusive and responsive to the larger imperatives of environmental care and shaping organizational ethos. The performance with DEA can also be used for cross differences and adapting DEA combined with uncertain data improves its applicability. In combination, these innovations will ensure that DEA retains its position as an adaptive, powerful operational efficiency and strategic decision-making tool across industries.

REFRENCES:

Agarwala, N., Maity, S. and Sahu, T.N. 2023. Efficiency of Indian banks in fostering financial inclusion: an emerging economy perspective. *Journal of Financial Services Marketing*. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9817463/.

Banker, R., Chames, A., Cooper, W., Swarts, J. and Thomas, D. 1989. AN INTRODUCTION TO DATA ENVELOPMENT ANALYSIS WITH SOME OF ITS MODELS AND THEIR USES. *Research in Governmental and Nonprofit Accounting* 5, pp. 125–163. Available at: https://iiif.library.cmu.edu/file/Cooper_box0010c_fld00033_bdl0001_doc0001/Cooper_box0010c_fld00033_bdl0001_doc0001.pdf.

Cooper, W.W., Seiford, L.M. and Tone, K. 2007. The Basic CCR Model. *Data Envelopment Analysis*, pp. 21–39. doi: https://doi.org/10.1007/978-0-387-45283-8_2.

González-de-Julián, S., Vivas-Consuelo, D. and Barrachina-Martínez, I. 2024. Modelling efficiency in primary healthcare using the DEA methodology: an empirical analysis in a healthcare district. *BMC Health Services Research* 24(1). doi: https://doi.org/10.1186/s12913-024-11420-2.

Hoque, Md.R. and Rayhan, Md.I. 2012. DATA ENVELOPMENT ANALYSIS OF BANKING SECTOR IN BANGLADESH. *Russian Journal of Agricultural and Socio-Economic Sciences* 5(5), pp. 17–22. doi: https://doi.org/10.18551/rjoas.2012-05.02.

Kuah, C.T. and Wong, K.Y. 2011. Efficiency assessment of universities through data envelopment analysis. *Procedia Computer Science* 3, pp. 499–506. doi: https://doi.org/10.1016/j.procs.2010.12.084.

Martić, M., Novaković, M. and Baggia, A. 2009. Data Envelopment Analysis - Basic Models and their Utilization. *Organizacija* 42(2), pp. 37–43. doi: https://doi.org/10.2478/v10051-009-0001-6.

Sharma, A., Rastogi, S. and Gupta, N. 2020. Financial Efficiency of Non – Banking Financial Companies-Microfinance Institutions: A Data Envelopment Analysis. *TEST ENGINEERING AND MANAGEMENT* 83(May-June), pp. 9080–9091. Available at: https://www.researchgate.net/publication/341593163_Financial_Efficiency_of_Non_-_Banking_Financial_Companies-Microfinance_Institutions_A_Data_Envelopment_Analysis.