U.PORTO

MESTRADO
GESTÃO COMERCIAL

Estratégias Comerciais e a Performance Económica: Evidência de Empresas Portuguesas

Carolina Macieira

M

2026



FEP FACULDADE DE **ECONOMIA E GESTÃO**

U.PORTO

FEP FACULDADE DE ECONOMIA E GESTÃO UNIVERSIDADE DO PORTO

Estratégias Comerciais e a Performance Económica: Evidência de Empresas Portuguesas

Carolina Delgado Macieira

Dissertação

Mestrado em Gestão Comercial

Orientado por

Professor Pedro José Ramos Moreira de Campos

Professor José António Cardoso Moreira

2026

Resumo

A presente dissertação analisa de que forma diferentes estratégias comerciais influenciam a performance económica das empresas portuguesas, distinguindo entre liderança em custos e diferenciação no quadro das estratégias genéricas de Porter (1980). O estudo centra-se nas empresas portuguesas, um segmento ainda pouco estudado na literatura.

Recorrendo a dados contabilísticos e financeiros extraídos da base SABI para os anos de 2022, 2023 e 2024, a análise incide sobre dois setores com estruturas competitivas distintas: a indústria transformadora (CAE C) e o comércio por grosso e a retalho (CAE G). A metodologia desenvolve-se em três fases: a identificação dos perfis estratégicos através de análise de *clusters (k-means)*, com base em dois rácios financeiros; a estimação de um modelo de regressão do *Return on Assets* (ROA) sobre a estratégia, controlando a dimensão e a idade das empresas; e a aplicação complementar de árvores de regressão, que permitem captar relações não lineares e efeitos de interação.

Os resultados são consistentes em ambos os setores: as empresas que seguem a liderança em custos apresentam, em média, um ROA superior em cerca de 1,9 a 2,0 pontos percentuais face às empresas que seguem uma estratégia de diferenciação, controlando a dimensão e a idade. As árvores de regressão confirmam estes resultados. Os resultados observados sustentam a hipótese de que a escolha da estratégia comercial está associada a efeitos distintos sobre a performance económica, sugerindo que, no contexto das empresas portuguesas, a liderança em custos constitui um caminho com retorno mais elevado do que a de uma estratégia de diferenciação.

Palavras-chave:

Estratégias comerciais; Performance económica; Liderança em custos; Diferenciação; Empresas portuguesas; Análise de *clusters*

Abstract

This dissertation analyses how different competitive strategies influence the economic performance of Portuguese firms, distinguishing between cost leadership and differentiation within the framework of Porter's (1980) generic strategies. The study focuses on Small and Medium-sized Enterprises, a segment that remains underexplored in literature.

Drawing on accounting and financial data extracted from the SABI database for the years 2022, 2023, and 2024, the analysis covers two sectors with distinct competitive structures: manufacturing (CAE C) and wholesale and retail trade (CAE G). The methodology unfolds in three stages: identification of strategic profiles through cluster analysis (k-means), based on two financial ratios; estimation of a regression model of Return on Assets (ROA) on strategy, controlling for firm size and age; and the complementary application of regression trees, which allow for the capture of non-linear relationships and interaction effects.

The results are consistent across both sectors: firms pursuing cost leadership exhibit, on average, a ROA approximately 1.9 to 2.0 percentage points higher than firms pursuing a differentiation strategy, controlling for size and age. The regression trees confirm these results. The findings support the hypothesis that the choice of competitive strategy is associated with distinct effects on economic performance, suggesting that, in the context of Portuguese firms, cost leadership constitutes a path with higher returns than differentiation.

Keywords:

Competitive strategies; Economic performance; Cost leadership; Differentiation; Portuguese SMEs; Cluster analysis

Índice

Resumo 2
Palavras-chave: 2
Abstract 3
Keywords: 3
I - Introdução 5
II – Revisão da Literatura 7
2.1 Estratégias Comerciais 7
2.2 Performance Económica 9
2.3 Relação entre as estratégias comerciais e a performance económica 10
2.4 Formulação da hipótese 11
III – Metodologia de investigação 14
3.1 Fase 1: análise de *clusters* 14
3.2 Fase 2: modelo econométrico 16
3.3 Fase 3: modelo de árvore de regressão 17
3.4 Utilização Complementar de Dois Modelos de Análise 18
3.5 Amostra 18
IV- Análises e resultados 23
4.1 Análise dos resultados da regressão linear 23
4.2 Análise dos resultados das regressões com variáveis multiplicativas 24
4.3 Análise dos resultados das árvores de regressão 26
V- Conclusões 31
Referências bibliográficas 34
Anexo 37
Caracterização da estratégia comercial 37

Índice de tabelas

Tabela 1: Limpeza da amostra 20
Tabela 2: Correlação das variáveis 21
Tabela 3: Coeficientes da regressão linear dos três anos 23
Tabela 4: Coeficientes dos modelos expandidos 25
Tabela 5: Resultados do impacto do tipo de estratégia na performance 27
Tabela 6: Resultados do impacto das variáveis de controlo na performance 28
Tabela 7: Segmentação de clusters 39
Tabela 8: Classificação de Clusters 39

I - Introdução

Nas últimas décadas, a forma como as empresas têm definido o seu posicionamento comercial estratégico tornou-se um tema bastante importante e debatido (Al-Shaer et al., 2023; Jermias, 2008; Valipour et al., 2012). A tensão entre competir pelo preço ou competir pelo valor está no centro das decisões que determinam a capacidade das empresas de sobreviver e crescer em mercados cada vez mais concorrenciais. Esta questão, formulada por Porter (1980) através das suas estratégias comerciais, continua a criar investigações relevantes para o mundo empresarial, particularmente em contextos onde as condições estruturais das empresas são diferentes dos mercados de grande escala onde a grande maioria dos estudos é feita (Jermias, 2008).

A presente dissertação insere-se neste debate e tem como objetivo principal analisar de que forma diferentes estratégias comerciais influenciam a performance económica das empresas portuguesas, distinguindo entre empresas orientadas para a liderança em custos e empresas orientadas para a diferenciação. Em concreto, pretende-se compreender se as empresas do primeiro desses grupos, que privilegiam a eficiência operacional e o crescimento do volume de vendas, apresentam uma performance económica distinta das do segundo grupo, que apostam na criação de valor percebido e na capacidade de praticar preços *premium*. O problema de investigação parte, assim, de uma tensão clássica da gestão comercial: crescer em vendas ou crescer em valor, sendo que as implicações desta escolha para a performance económica ainda são uma questão em aberto na literatura, com resultados que apontam em direções opostas consoante o contexto, o setor e as métricas de performance utilizadas (Dess & Davis, 1984; Hansen et al., 2015; Jermias, 2008; Valipour et al., 2012).

A relevância desta investigação assenta, no facto da relação entre estratégias comerciais e a performance económica continuar pouco estudada em contextos de pequena e média dimensão empresarial. A maior parte dos estudos já realizados assenta em economias de maior escala, com empresas cotadas e setores altamente industrializados, cuja realidade difere da estrutura empresarial portuguesa (Abdelqader et al., 2024; Jermias, 2008). Em Portugal, o tecido empresarial é composto maioritariamente por pequenas e médias empresas (PME), caracterizadas por recursos limitados, menor poder de mercado e maior sensibilidade a variações de custos e de procura. Neste contexto, a forma como estas empresas definem a sua estratégia comercial torna-se uma das questões mais importantes, pois influencia diretamente a sua performance, a sua capacidade de sobrevivência e o seu crescimento. A relevância do tema é também económica, pois as empresas competem num ambiente de crescente pressão e de margens operacionais cada vez mais reduzidas. Dou et al. (2021)

demonstram que a intensificação da concorrência tende a reduzir a performance económica das empresas, sobretudo em mercados com baixa diferenciação, o que torna mais pertinente compreender quais os posicionamentos estratégicos que trazem uma melhor performance. Por fim, esta investigação possui também relevância prática para gestores, investidores e decisores de política empresarial, ao oferecer evidência empírica sobre quais as estratégias comerciais associadas a melhor performance no contexto económico português.

Do ponto de vista teórico, este estudo contribui para uma melhor compreensão da relação entre as estratégias comerciais de Porter e a performance económica das empresas. Além disso, acrescenta evidência empírica sobre uma realidade ainda pouco estudada na literatura internacional, a das empresas portuguesas. Desta forma, ajuda a alargar o conhecimento existente sobre este tema para além dos contextos habitualmente analisados, que se focam sobretudo em economias de maior escala. A análise simultânea de dois setores com estruturas operacionais distintas, constitui um contributo adicional face a estudos anteriormente realizados como os de Jermias (2008) e Valipour et al. (2012). Do ponto de vista metodológico, a combinação de *clustering* com regressão linear e árvores de regressão representa uma abordagem mais abrangente do que a utilizada na maioria dos estudos de referência sobre este tema, permitindo captar tanto os efeitos médios da estratégia como padrões de performance não detetáveis por modelos puramente lineares. Do ponto de vista prático, os resultados fornecem conclusões para gestores de empresas portuguesas nos setores transformador e comercial, evidenciando que uma estratégia de liderança pelos custos, no contexto português, proporciona, em média, um retorno económico maior do que a utilização de uma estratégia de diferenciação.

O texto que se segue está organizado da seguinte forma. O Capítulo II apresenta a revisão da literatura, enquadrando teoricamente as estratégias comerciais, a performance económica e as relações entre ambas e a formulação de hipótese. O Capítulo III descreve a metodologia adotada e a amostra analisada. O Capítulo IV apresenta e discute os resultados obtidos. O Capítulo V encerra o trabalho com as considerações finais, onde sintetizei os contributos do estudo, as suas limitações e direções para investigações futuras.

II – Revisão da Literatura

2.1 Estratégias Comerciais

2.1.1. Contextualização

A estratégia comercial é um dos aspetos mais importantes para as empresas conseguirem gerar valor e ganhar vantagem face os seus concorrentes. Para Jermias (2008), o estudo de Porter (1980) desenvolveu um conjunto de referências que descreve como as empresas podem escolher uma estratégia de negócio de forma eficaz, dizendo que uma empresa deve escolher entre competir como um produtor de menor custo, a chamada estratégia de liderança em custos; ou competir oferecendo um produto singular com mais qualidade, a chamada estratégia de diferenciação. A estratégia comercial reside assim na capacidade da empresa conseguir escolher qual a melhor solução para conseguir alcançar mais valor (Jermias, 2008).

Porter (1980) para além destas duas estratégias que identificou, propôs ainda uma terceira estratégia, a estratégia do foco, que consiste em concentrar a atividade das empresas num segmento específico, aplicando a lógica da estratégia de liderança em custos e a estratégia de diferenciação, mas num contexto mais restrito. Contudo, para simplificar a presente análise, foi decido testar o modelo apenas com as estratégias de liderança em custos e de diferenciação, pois a estratégia de foco pode ser entendida como uma variante das anteriores e por isso podia confundir os resultados. Para que isso não aconteça, apenas foi aplicado estratégias que sejam percetíveis de serem distinguidas (Hansen et al., 2015).

Relativamente às estratégias comerciais importa clarificar uma distinção que frequentemente gera mal-entendidos: tanto a estratégia de liderança em custos como a diferenciação têm como objetivo final maximizar a performance da empresa. O que as distingue não é o objetivo, mas sim o mecanismo competitivo que cada uma delas usa para chegar a esse objetivo.

2.1.2. A liderança em custos

A estratégia de liderança em custos assenta na capacidade de uma empresa produzir bens ou prestar serviços a um custo inferior ao dos seus concorrentes, com o propósito de obter uma vantagem competitiva no mercado (Agustia et al., 2020; Al-Shaer et al., 2023). Porter (1980) afirma que para seguir uma estratégia de liderança em custos as empresas devem possuir as seguintes características: ter o processo de fabrico o mais eficiente possível; ter um controlo dos custos para minimizar os gastos; minimizar os custos destas áreas: investigação, desenvolvimento, serviços e publicidade. As empresas que possuam estas

características competem a um nível de preço e de eficiência produtiva, dirigindo-se a clientes mais sensíveis aos preços. Al-Shaer et al. (2023) refere-se que as empresas que seguem esta estratégia tendem a focar-se mais na eficiente utilização de ativos, no controlo das despesas e na maximização da produtividade para conseguir ser o mais produtivas possível. Farida and Setiawan (2022) afirmam também que com esta estratégia é possível impossibilitar os concorrentes de atingir uma maior quota de mercado ao reduzir o custo do processo produtivo, o que permite oferecer preços mais baixos, criando assim uma barreira à entrada de novas empresas no mercado e reforçando a lealdade dos clientes. As empresas que seguem esta estratégia conquistam quota de mercado através dos preços agressivos que realizam, conseguindo produzir produtos de fabrico fácil e em massa, apoiando-se em tecnologias e equipamentos de ponta para maximizar a eficiência produtiva (Panwar et al., 2016). Sendo que, as empresas ao produzir produtos a um custo tão reduzido, devem de ter em atenção para não comprometer as suas bases concorrenciais, para não haver a probabilidade de o produto ficar com pouca qualidade em relação à concorrência e perder quota de mercado por causa disso (Ortega, 2010). Anthony and K. (1992) referem que a curva de aprendizagem de uma empresa cria barreiras à entrada no mercado e cria uma proteção face à concorrência ao gerar vantagens de custo para quem alcança grandes quotas de mercado nas fases iniciais do ciclo de vida. Embora estas vantagens possam ser temporárias, têm um impacto significativo na quota de mercado e na performance da empresa.

2.1.3. A estratégia de diferenciação

A estratégia de diferenciação consiste na oferta de produtos ou serviços que o mercado reconhece como únicos, singulares, permitindo à empresa praticar preços acima da média comparado com os seus concorrentes (Al-Shaer et al., 2023; Farida & Setiawan, 2022). Ortega (2010) define que numa estratégia de diferenciação "a empresa procura ser única na sua indústria ao longo de algumas dimensões amplamente valorizadas pelos compradores"(p.1275). Esta singularidade pode assumir diversas formas, incluindo a imagem de marca, o serviço prestado, os canais de distribuição, a qualidade ou os atributos específicos do produto (Chen et al., 2018). Al-Shaer et al. (2023) afirma que esta estratégia pretende criar fidelização dos clientes à marca através da oferta de produtos diferenciados, permitindo à empresa praticar preços *premium* que superam os custos associados à produção. Esta estratégia é caracterizada também pelo facto de ser necessário investir em investigação e desenvolvimento com o objetivo de aumentar a capacidade de inovação e de acompanhamento das inovações dos concorrentes (Jermias, 2008). Sendo que as empresas que seguem uma estratégia de diferenciação concentram-se mais na construção de uma

reputação e na criação de uma imagem para conseguir ganhar mais fidelidade dos clientes e por isso incorrem em mais gastos em publicidade e promoções (Abdelqader et al., 2024). Abdelqader et al. (2024) conclui ainda que as empresas que seguem uma estratégia de diferenciação lidam com uma maior incerteza, pois os grandes investimentos em inovação e desenvolvimento levam a que a atividade tenha um maior risco, pois a empresa está a apostar num produto que não sabe se vai gerar bons resultados. Esta característica tem implicações importantes para a estrutura financeira destas empresas. Em contrapartida, esta estratégia oferece vantagens competitivas de difícil replicação. Hansen et al. (2015) argumentam que uma estratégia de diferenciação multifacetada é mais difícil de imitar pelos concorrentes do que a estratégia unidimensional de liderança em custos, uma vez que combina múltiplas dimensões, como inovação, qualidade, serviço e marca.

2.2 Performance Económica

A performance económica pode ser definida como a capacidade de uma empresa alcançar os seus objetivos operacionais e estratégicos de forma eficiente, sendo avaliada com indicadores como a rentabilidade, o crescimento das vendas, a liquidez e a capacidade de investimento (Probohudono et al., 2025). Quando avaliada pela rentabilidade, a performance pode ser medida através de indicadores como o retorno sobre os ativos (ROA), o retorno sobre o capital próprio (ROE) e a margem de lucro. Anthony and K. (1992) definem o ROA como o resultado do EBIT dividido pelo total de ativos, utilizando-o como variável descritiva comparativa entre grupos de empresas em diferentes estágios do ciclo de vida. A escolha do ROA é justificada com o facto de se tratar de um indicador que tem capacidade de refletir a eficiência com que a empresa utiliza os seus ativos para gerar resultados, independentemente da respetiva estrutura financeira. Cho and Lee (2017) utilizam o ROA como variável de controlo nos seus modelos de performance financeira, validando a sua pertinência como indicador de performance contabilística. Existe um conjunto alargado de fatores que podem influenciar a performance económica, desde as condições do mercado em que a empresa está a operar como as características internas da empresa. Teixeira et al. (2019) afirmam que o modelo das cinco forças competitivas de Porter (1980) constitui uma referência fundamental para perceber de que forma as características do mercado condicionam a performance das empresas que nele operam. As condições estruturais do mercado como a rivalidade entre concorrentes, o poder de negociação de fornecedores e clientes, a ameaça de substitutos e as barreiras à entrada, definem o poder estratégico de cada empresa. Ao nível mais interno da empresa, os recursos que a empresa utiliza e as capacidades que tem são também um fator fundamental para alcançar uma boa performance. Ortega (2010), salienta

que tanto a estratégia comercial que a empresa utiliza como a gestão dos recursos que a empresa faz pressupõem a importância de uma posição estratégica vantajosa em relação aos seus concorrentes e que quanto mais recursos a empresa possuir, maior será a sua capacidade de desenvolver uma estratégia que lhe permita ter melhores vantagens competitivas. As capacidades de uma empresa para integrar, construir e reconfigurar competências internas e externas de forma a responder a ambientes em mudança rápida são vitais para conseguir manter a sua performance a longo prazo, incluindo a de gerar e ajustar os seus modelos de negócio em resposta às mudanças do mercado competitivo (Teece, 2018).

Existem ainda outros fatores que influenciam a performance, como é o caso do ciclo de vida da empresa. Anthony and K. (1992) demonstram que algumas determinantes contabilísticas da performance, como o crescimento das vendas e o investimento em capital, variam significativamente consoante o estágio do ciclo de vida em que a empresa se encontra. Empresas em fase de crescimento apresentam coeficientes de resposta muito mais elevados do que empresas em estagnação, isto é, a sua rapidez de resposta às mudanças do mercado é mais elevada do que empresas em estagnação, evidenciando que diferentes posicionamentos estratégicos ao longo do ciclo de vida correspondem a diferentes perfis de performance.

A dimensão da empresa, o setor de atividade e o contexto macroeconómico são também variáveis que podem influenciar a performance da empresa. Ortega (2010) controla, nos seus modelos, fatores como a dimensão da empresa, o momento de entrada no mercado, o potencial de mercado e o dinamismo, demonstrando que as empresas com maiores capacidades tecnológicas e de controlo de custos podem apresentar valores diferentes de performance em comparação com outras empresas.

2.3 Relação entre as estratégias comerciais e a performance económica

A premissa principal da relação entre as estratégias comerciais e a performance económica é que empresas com um posicionamento estratégico claramente definido, quer de liderança em custos, quer de diferenciação tendem a apresentar uma performance superior às que não seguem nenhuma orientação estratégica definida (Dess & Davis, 1984; Teixeira et al., 2019). A pertença a um grupo estratégico claramente definido está associada a melhores resultados do que as situações de *stuck in the middle*, onde a empresa não tem uma estratégia bem definida (Dess & Davis, 1984). Embora a literatura confirme que ambas as estratégias se associam positivamente à performance, a questão de qual delas produz melhores resultados permanece em aberto, com a evidência científica a apresentar resultados distintos. Hansen et al. (2015), num estudo com 441 empresas do setor florestal norte-americano,

concluíram que as empresas com estratégia de diferenciação pura obtêm uma performance superior ao das empresas com estratégia de liderança em custos pura, antecipando que o aumento das receitas decorrente da diferenciação é a principal razão para este diferencial. Em contraste, Dess and Davis (1984) num estudo realizado ao setor de produtos industriais, concluíram que o grupo de liderança em custos registava o retorno sobre os ativos mais elevado em contraste com o grupo de diferenciação. Jermias (2008) demonstrou que o impacto negativo da alavancagem financeira sobre a performance é mais visto nas empresas que seguem a estratégia de diferenciação do que as empresas que seguem a estratégia de liderança em custos, o que decorre das diferentes necessidades de flexibilidade financeira associadas a cada estratégia, enquanto as que seguem a de liderança em custos beneficiam do efeito disciplinador da dívida na eficiência de gestão, para as empresas que seguem uma estratégia de diferenciação um endividamento excessivo pode limitar a capacidade de inovação, o que pode influenciar negativamente a sua vantagem competitiva. Valipour et al. (2012) confirmou também esta distinção, encontrando uma relação positiva entre a estratégia de liderança em custos e a performance, e uma relação negativa entre a diferenciação e a performance, embora reconheçam a possibilidade de os resultados refletirem especificidades institucionais deste contexto.

A literatura tem mostrado que os efeitos da estratégia sobre a performance não são sempre diretos, podendo ser influenciados por várias variáveis. Chen et al. (2018) demonstram que tanto a liderança em custos como a diferenciação aumentam o efeito positivo da responsabilidade social empresarial sobre a performance económica: as empresas que seguem a estratégia de liderança em custos fazem-no através do controlo de custos explícitos, enquanto as empresas que seguem a estratégia de diferenciação o fazem através da gestão de custos implícitos e da inovação. Os autores concluem que, "níveis mais elevados de performance social e de liderança em custos ou diferenciação podem conduzir a uma melhor performance da empresa"(p.162).

2.4 Formulação da hipótese

A revisão de literatura apresentada anteriormente permitiu perceber um conjunto de aspetos sobre a relação entre a estratégia comercial adotada pelas empresas e a sua performance económica. Embora exista um consenso na literatura quanto ao facto de que tanto a estratégia de liderança em custos como a estratégia de diferenciação apresentam resultados positivos de performance, importa perceber se estas duas estratégias produzem efeitos semelhantes, ou se existe uma estratégia que apresente uma melhor performance económica em relação à outra.

Como visto anteriormente, os resultados disponíveis apontam em direções diversas consoante o contexto, o setor e a medida de performance utilizada. Dess and Davis (1984) afirmam que o grupo de empresas orientado para a liderança em custos registou o retorno sobre os ativos mais elevado, enquanto o grupo de diferenciação se destacou no crescimento das vendas. Em sentido contrário, Hansen et al. (2015) concluíram que as empresas com estratégia de diferenciação pura superam as líderes em custos em termos de performance. Jermias (2008) demonstrou, por sua vez, que o impacto da alavancagem financeira sobre a performance é menos negativo para as empresas que seguem a estratégia de liderança em custos do que as que seguem a estratégia de diferenciação, mostrando assim que as duas estratégias implicam resultados de performance distintos. Também Abdelqader et al. (2024) mostram que a liderança em custos e a diferenciação produzem efeitos opostos sobre determinadas dimensões do valor da empresa, reforçando a ideia de que não são estrategicamente equivalentes em termos de resultados da performance.

Este leque de resultados, reflete a complexidade inerente à relação entre a estratégia e a performance e a necessidade de estudos adicionais em contextos geográficos e institucionais específicos. O contexto português é, neste sentido, particularmente relevante por se tratar de uma economia caracterizada por um conjunto empresarial com forte presença de empresas portuguesas, inseridas em setores com elevada concorrência, onde as opções estratégicas de posicionamento têm implicações fortes na performance das empresas (Teixeira et al., 2019).

Face a este enquadramento, em que a literatura não fornece uma ordem de grandeza relativa de cada estratégia sobre a performance, a presente dissertação propõe a seguinte hipótese de investigação:

H1: A escolha de uma estratégia comercial, nomeadamente liderança em custos ou diferenciação, está associada a um impacto específico na performance económica da empresa, sendo espectável que tal impacto seja distinto consoante o tipo de estratégia.

Esta hipótese é sustentada pela evidência demonstrada anteriormente de que empresas com posicionamentos estratégicos diferentes apresentam perfis de performance diferentes (Abdelqader et al., 2024; Dess & Davis, 1984; Hansen et al., 2015; Jermias, 2008), e pela premissa de Porter (1980) de que a escolha de uma estratégia genérica clara é uma condição necessária para alcançar performance acima da média (Teixeira et al., 2019). Importa, portanto, verificar, com base em dados de empresas portuguesas, se esta distinção se mantém e qual das duas estratégias se associa a melhores resultados ao nível do retorno

sobre os ativos.

III – Metodologia de investigação

A metodologia desta investigação foi desenhada para assegurar uma ligação clara entre o problema de investigação, a hipótese formulada, e os procedimentos adotados na análise dos resultados. Foi escolhida uma abordagem quantitativa pela necessidade de analisar relações objetivas entre variáveis financeiras, permitindo testar o impacto das estratégias comerciais na performance económica das empresas. Assim, todo o processo metodológico, desde a recolha, tratamento dos dados, a classificação estratégica das empresas e subsequente modelação estatística, foi delineado de modo a garantir consistência, transparência e robustez nos resultados obtidos.

A metodologia é composta assim por 3 fases, a primeira fase aborda a análise de *clusters*, a segunda análise aborda o modelo econométrico e a terceira fase aborda o modelo de árvore de regressão.

3.1 Fase 1: análise de *clusters*

Nesta primeira fase procede-se à classificação das empresas de acordo com a sua estratégia comercial, recorrendo à análise de *clusters*[1]. Esta técnica estatística multivariada permite agrupar empresas com características semelhantes, com base num conjunto de variáveis previamente selecionadas, formando grupos internamente homogéneos e externamente heterogéneos (Ikotun et al., 2025). O objetivo é identificar perfis estratégicos distintos de forma endógena, sem impor critérios arbitrários de classificação.

A utilização da análise de *clusters* nesta etapa inicial permite reduzir a subjetividade associada à classificação exógena das estratégica, assegurando que a distinção entre estas estratégias resulta de padrões observados nos dados. Desta forma, garante-se maior consistência entre a identificação dos perfis estratégicos e a análise econométrica subsequente.

A análise de *clusters* foi baseada no estudo de Jermias (2008) onde foram usadas as seguintes variáveis financeiras para diferenciar as empresas entre as duas estratégias anteriormente referidas: variável 1 – eficiência na utilização de ativos (vendas/ativo); variável 2 – capacidade de captar preços *premium* (margem bruta/vendas). Estas variáveis têm subjacentes diferenças na estrutura de custos, na intensidade operacional e na orientação estratégica, permitindo distinguir empresas com estratégia de liderança em custos de empresas com estratégia de diferenciação.

O estudo de Jermias (2008) apresentava ainda uma terceira variável: variável 3 – a

[1] A explicação do processo da classificação dos clusters encontra-se em anexo.

intensidade de I&D (despesas de I&D / vendas). Esta variável associada à intensidade de investigação e desenvolvimento (I&D), frequentemente utilizada na literatura como proxy da estratégia de diferenciação, não foi incluída no modelo devido à indisponibilidade de dados na base SABI. Embora a literatura utilize tipicamente medidas diretas, como o rácio de despesas de I&D sobre vendas para captar esta dimensão estratégica, a ausência desta informação para a maioria das empresas da amostra inviabiliza a sua operacionalização consistente. Neste contexto, a utilização de *proxies* alternativas baseadas em rubricas contabilísticas agregadas não garantiria uma correspondência adequada com o construto teórico de inovação, podendo introduzir os dados em erro de mensuração.

Assim, optou-se por excluir esta variável do modelo, privilegiando a utilização de indicadores diretamente observáveis e comparáveis entre empresas, em linha com a abordagem adotada na literatura, que mostra a necessidade de assegurar a validade das medidas utilizadas na análise.

Assim, espera-se que empresas diferenciadoras apresentem uma maior intensidade de investimento comercial e de inovação e uma maior capacidade de praticar preços premium, enquanto empresas orientadas para a liderança em custos revelem níveis mais elevados de eficiência operacional.

Com base nos valores assumidos por estas variáveis, as empresas são agrupadas em dois *clusters*, sendo cada *cluster* interpretado como representativo de um perfil estratégico específico.

Após a realização da análise de *clusters*, cada empresa é atribuída a um único grupo estratégico, de acordo com o *cluster* ao qual pertence. A literatura evidencia que o algoritmo do *clustering* é sensível à presença de observações isoladas, podendo estas comprometer a qualidade dos agrupamentos, sendo por isso recomendada a sua exclusão para obter soluções mais robustas (Prastyabudi et al., 2024). Neste contexto, após a aplicação do *clustering*, optou-se por considerar apenas os dois grupos de *clusters* de maior dimensão, por representarem os padrões mais consistentes e dominantes na amostra, interpretando-os como as estratégias relevantes em análise. Os restantes *clusters*, de dimensão reduzida, foram classificados como agregados de observações atípicas (*outliers*), por não apresentarem características suficientemente consistentes que permitam a sua associação clara a nenhuma das estratégias identificadas.

Com base nessa classificação, é construída uma variável binária indicadora da estratégia comercial, que irá assumir valores de 1 para a estratégia de liderança nos custos e valores de 0 para a estratégia de diferenciação. Esta variável é posteriormente utilizada como

variável explicativa no modelo de regressão linear, permitindo analisar o impacto das diferentes estratégias comerciais na performance económica das empresas.

3.2 Fase 2: modelo econométrico

Na segunda parte o estudo utiliza o seguinte modelo de regressão linear baseado no trabalho de Jermias (2008) e de Valipour et al. (2012) para testar as hipóteses.

$$ROAit = B0 + B1ESTit + B2DIMit + B3\,IDAit + u,$$

onde ROA é o EBIT dividido pelo ativo total e representa a performance económica da empresa i no ano t; EST é a variável binária do modelo e representa a estratégia que as empresas adotam, atingindo o valor de 1 para empresas com uma estratégia de liderança em custos e o valor de 0 para empresas com uma estratégia de diferenciação. DIM representa a dimensão da empresa, medida pelo logaritmo do ativo total no ano t; IDA representa a idade da empresa i no ano t; u representa a variável de erro. DIM e IDA são variáveis de controlo.

A escolha do ROA justifica-se ainda pela sua estabilidade relativa face a flutuações de curto prazo e pela sua adequação em análises comparativas entre empresas pertencentes a diferentes setores, desde que acompanhada de variáveis de controlo apropriadas (Dess & Davis, 1984; Park et al., 2019). Assim, esta variável permite avaliar de forma consistente o impacto das estratégias comerciais, liderança em custos ou diferenciação, sobre a performance económica das empresas portuguesas incluídas na amostra.

Relativamente a DIM, a evidência constante da literatura aponta para um efeito não muito claro. Por um lado, empresas de maior dimensão tendem a beneficiar de economias de escala, maior acesso a recursos e melhor posicionamento competitivo, o que pode traduzir-se num impacto positivo na performance. Por outro lado, estruturas organizacionais mais complexas podem gerar ineficiências, rigidez e maiores custos fixos, conduzindo potencialmente a um efeito negativo. Neste sentido, alguns estudos incluem a dimensão como variável de controlo precisamente para isolar estes efeitos contrastantes sobre a performance. Assim, embora o sinal esperado possa ser positivo, é possível considerar o efeito indeterminado à priori (Jermias, 2008; Park et al., 2019; Valipour et al., 2012).

Relativamente à IDA, a literatura sugere que empresas mais antigas (ou que entraram mais cedo no mercado) poderão beneficiar de vantagens acumuladas, como experiência, reputação e aprendizagem organizacional, apontando para um efeito positivo na performance. No entanto, tal como no caso da dimensão, empresas mais antigas podem também apresentar maior estagnação estratégica e menor capacidade de adaptação, o que pode limitar a sua performance. Assim, o sinal esperado para o coeficiente de IDA tende a ser positivo na literatura, mas com possíveis limitações associadas à rigidez organizacional

(Ortega, 2010).

3.3 Fase 3: modelo de árvore de regressão

Na terceira parte do estudo é aplicado um algoritmo de árvores de regressão para complementar os resultados obtidos da regressão linear. As árvores de regressão são uma metodologia de aprendizagem automática (*machine learning*) muito usada em vários estudos, sendo especialmente adequada para analisar relações complexas e não lineares (Sutton, 2005). Testa-se, pois, nesta fase, a possibilidade da relação entre estratégia e performance não ser linear. Contrariamente à regressão da Fase 2, que considera uma relação linear entre as variáveis explicativas e a variável dependente, as árvores de regressão dividem as variáveis explicativas sucessivamente em subconjuntos cada vez mais homogéneos, criando previsões com base na média da variável dependente em cada subconjunto terminal chamado de "folha". O resultado desta repartição pode ser representado graficamente como uma árvore de decisão, tornando a interpretação dos resultados mais percetível.

O processo de construção de uma árvore de regressão começa com a totalidade da amostra no "nó" raiz. Em cada etapa, o algoritmo seleciona a variável explicativa e o valor de corte que melhor divide a amostra em dois subgrupos, minimizando a variância intragrupo da variável dependente. Este processo é repetido de forma sucessiva até que um critério de paragem seja atingido, normalmente associado à dimensão mínima dos nós ou à ausência de um melhor resultado na repartição (Sutton, 2005). Para evitar o sobreajustamento (*overfitting*), recorre-se a técnicas de "poda" (*pruning*), que simplificam a estrutura da árvore eliminando divisões que não contribuem para a explicação da variância total da variável dependente (Sutton, 2005).

A principal vantagem deste modelo face à regressão linear está na sua capacidade de captar efeitos de limiar (*threshold effects*), ou seja, situações em que o impacto de uma variável explicativa sobre a performance não é constante ao longo de toda a amostra, mas varia consoante o valor dessa mesma variável ou de outras variáveis presentes no modelo. Do mesmo modo, as árvores de regressão permitem identificar interações entre variáveis, sem que estas precisem de ser especificadas *a priori*, ao contrário do que acontece na regressão linear. Tornando-as mais úteis em contextos como o presente, em que a relação entre estratégia comercial e performance possa ser diferente em função da dimensão ou da idade das empresas (Delen et al., 2013). Adicionalmente, a representação gráfica em forma de árvore facilita a interpretação dos resultados, mostrando de forma clara quais as combinações de estratégias comerciais estão associadas a diferentes níveis de performance económica.

3.4 Utilização Complementar de Dois Modelos de Análise

A regressão linear é o modelo principal para testar a hipótese formulada no estudo, seguindo a metodologia de vários estudos realizados complementares ao presente estudo (Jermias, 2008; Valipour et al., 2012). No entanto, os modelos de regressão linear seguem um conjunto de pressupostos, entre os quais a linearidade da relação entre as variáveis explicativas e a variável dependente, a ausência de interações não especificadas e a homogeneidade dos efeitos ao longo de toda a amostra, que podem não ser os melhores para analisar certos conjuntos de dados com uma amostra muito grande e complexa. Com isto, é possível que a relação entre as estratégias comercial e a performance económica não seja linear, criando efeitos de interação que um modelo de regressão linear muito provavelmente não consegue alcançar da forma mais adequada. Como observam Delen et al. (2013), os estudos que recorrem a técnicas estatísticas como a regressão linear podem não conseguir alcançar adequadamente as relações entre variáveis quando estas assumem uma relação não linear ou quando existem interações entre elas, uma limitação que estes autores identificaram.

Por esse motivo, no que se pode considerar uma novidade neste tipo de estudos, foi desenvolvido um segundo modelo baseado em árvores de regressão, com o objetivo de complementar os resultados obtidos através da regressão linear. A utilização de outro método distinto para além da regressão linear não tem como objetivo substituir a análise que foi realizada inicialmente, mas sim fornecer uma perspetiva adicional sobre a estrutura dos dados e as relações entre as variáveis.

3.5 Amostra

A amostra é composta por empresas portuguesas pertencentes aos setores industriais CAE C (Indústrias Transformadoras) e comerciais CAE G (Comércio por grosso e a retalho), extraídas da base de dados SABI (*Sistema de Análise de Balanços Ibéricos*).

A seleção dos setores de atividade, indústrias transformadoras (CAE C) e comércio por grosso e a retalho (CAE G), assenta na relevância para o estudo das estratégias de Porter (1980). O CAE C é dos setores mais frequentemente analisado em estudos empíricos sobre liderança em custos e diferenciação. A investigação de Jermias (2008), que serve de base metodológica ao meu estudo, foi desenvolvida a partir de uma amostra de empresas manufatureiras, o que contribui para a adequação desta abordagem ao contexto do tecido industrial português. Dess and Davis (1984) confirmam também que é no setor transformador que a distinção entre orientação para a eficiência operacional e orientação para a diferenciação de produto se identifica de forma mais nítida.

O CAE G representa as empresas do setor comercial, estas operam em mercados com elevada pressão concorrencial sobre preços e tem uma necessidade de criar valor focada na diferenciação do serviço e do posicionamento, tornando-o igualmente importante para a análise das estratégias genéricas de Porter (1980). A inclusão de dois setores com componentes estruturais diferentes permite ainda verificar a robustez das relações entre estratégia e performance em contextos económicos distintos (Farida & Setiawan, 2022; Teixeira et al., 2019)

Tal como em Jermias (2008), será utilizado um painel temporal mínimo de três anos, permitindo suavizar volatilidades anuais e aumentar a robustez estatística dos resultados. Os dados serão tratados em Excel e posteriormente analisados no software estatístico (SPSS).

Os dados foram recolhidos no período compreendido entre novembro de 2025 e maio de 2026, pelo que os anos em análise correspondem aos anos fiscais de 2022, 2023 e 2024. Para a definição da estratégia de extração da amostra foram estabelecidos critérios específicos. Em primeiro lugar, consideraram-se apenas empresas ativas à data da recolha da informação. Em segundo lugar, restringiu-se a análise a empresas de nacionalidade portuguesa. Adicionalmente, foram selecionadas exclusivamente empresas das divisões 10 a 33, correspondentes ao setor da indústria transformadora (CAE Setor C); e das divisões 45 a 47, correspondentes ao setor do comércio (CAE Setor G). Por fim, constituiu requisito a existência de informação financeira disponível para os três anos em análise.

No que respeita às variáveis recolhidas, foram extraídos os seguintes indicadores económico-financeiros: vendas e serviços prestados, total do ativo, EBIT, capital próprio e margem bruta. Para além destes elementos, foi igualmente recolhida a data de constituição de cada empresa, permitindo a posterior consideração da idade como variável de controlo.

Com o objetivo de garantir a qualidade dos dados utilizados na análise, foi aplicado um conjunto de procedimentos de limpeza à base de dados inicial.

Em primeiro lugar, foram eliminadas todas as empresas com capital próprio negativo. Empresas nesta situação encontram-se tipicamente em situação de dificuldades financeiras, apresentando um perfil financeiro que não é representativo de uma orientação estratégica estável, o que poderia distorcer os resultados da análise. A sua inclusão introduziria na amostra observações economicamente não comparáveis com as demais, comprometendo a validade dos grupos estratégicos.

Em segundo lugar, foram excluídas todas as empresas com ativo total inferior a 100.00 euros. Empresas com um volume de ativos tão reduzido apresentam rácios financeiros potencialmente instáveis e sem significado económico interpretável, uma vez que

pequenas variações absolutas em numeradores ou denominadores produzem valores extremos nos rácios calculados.

Em terceiro lugar, foi aplicado um filtro ao ROA, retendo apenas as observações com valores compreendidos no intervalo (−0,8; 0,8). Este procedimento visa eliminar valores extremos que, embora possam corresponder a situações reais, representam comportamentos económicos atípicos, como perdas ou lucros extraordinários, que correspondem a *outliers* relativamente à distribuição normal da performance e que poderiam distorcer os resultados das regressões.

Por último, foram eliminadas todas as observações classificadas como "0" ou como erro de divisão ("#DIV/0!"), correspondentes a situações em que os denominadores das variáveis calculadas são nulos, tornando os rácios indefinidos ou matematicamente inválidos.

A seguinte tabela carateriza a amostra depois da limpeza aos dados.

Tabela 1: Limpeza da amostra[2]

		Amostra retirada do SABI	Capital Próprio negativo	Ativo < 100	ROA fora de [-0,8;0,8]	"0" ou "#DIV/0!"	Amostra final
CAE C	2024	28 880	4 338	6 688	720	174	20 372
	2023	28 880	4 249	6 972	741	556	20 142
	2022	28 880	4 155	7 351	759	241	19 795
CAE G	2024	69 606	12 598	20 237	2 391	1 228	43 243
	2023	69 606	12 367	20 976	2 283	1 353	44 362
	2022	69 606	12 371	22 142	2 513	1 897	43 203

Após a recolha dos dados, procedeu-se à construção das variáveis que integram o modelo de análise, com base nas fórmulas previamente apresentadas acima, no presente capítulo.

Antes de avançar para a estimação dos modelos de regressão, procedeu-se a uma análise preliminar das associações bivariadas entre as variáveis contínuas do estudo. Esta análise tem como objetivo caracterizar o sentido e a intensidade das relações entre as variáveis em cada subamostra.

Os coeficientes de correlação de Pearson foram calculados separadamente para cada subsecção resultante do cruzamento entre setor de atividade (CAE C e CAE G), ano (2022, 2023 e 2024) e estratégia comercial adotada (liderança em custos e diferenciação), totalizando doze subamostras. A tabela seguinte resume os resultados obtidos.

[2] Note de que a amostra final não corresponde à amostra retirada do SABI menos as restantes colunas pois as empresas podem estar em mais do que uma coluna ao mesmo tempo, por exemplo uma empresa pode ter o ativo inferior a 100 e estar fora do intervalo do ROA.

Tabela 2: Correlação das variáveis

Setor (CAE)	Ano	Estratégia	N	ROA–DIM	ROA–IDA	DIM–IDA
CAE C	2024	Liderança em custos	5 062	0,060	-0,106	**0,344**
		Diferenciação	15 302	0,076	-0,068	**0,318**
	2023	Liderança em custos	4 562	0,076	-0,051	**0,321**
		Diferenciação	15 577	0,052	-0,117	**0,371**
	2022	Liderança em custos	4 949	-0,007 (n.s.)	-0,134	**0,382**
		Diferenciação	14 840	0,042	-0,062	**0,331**
CAE G	2024	Liderança em custos	12 357	0,034	-0,108	**0,293**
		Diferenciação	30 879	0,113	-0,111	**0,195**
	2023	Liderança em custos	12 811	0,012 (n.s.)	-0,117	**0,311**
		Diferenciação	29 359	0,092	-0,108	**0,201**
	2022	Liderança em custos	11 015	0,017 (a)	-0,121	**0,325**
		Diferenciação	30 755	0,092	-0,113	**0,211**

Nota: coeficientes de correlação de Pearson. Salvo indicação em contrário, todas as correlações são significativas a p < 0,001 (unilateral). (n.s.) = não significativa (ROA–DIM: p = 0,312 em CAE C/2022/liderança; p = 0,096 em CAE G/2023/liderança). (a) significância marginal (ROA–DIM: p = 0,040 em CAE G/2022/liderança).

A correlação entre dimensão e idade (DIM-IDA) é positiva em todas as variáveis, variando entre 0,195 e 0,382, portanto, magnitudes baixas a moderadas. Observa-se ainda que a associação é sistematicamente mais forte no setor C (0,318–0,382) do que no setor G (0,195–0,325), trata-se de uma inferência a partir dos dados, não testada formalmente, que sugere que na indústria transformadora dimensão e antiguidade tendem a crescer de forma mais conjunta do que no comércio.

A correlação entre idade e performance (ROA-IDA) é negativa em todas as doze variáveis, entre −0,051 e −0,134. A direção é, portanto, consistente, empresas mais antigas apresentam, em média, ROA ligeiramente inferior, mas a magnitude é fraca em todos os casos, o que significa que não é possível retirar conclusões conclusivas a partir destes coeficientes isolados.

A correlação entre dimensão e performance (ROA-DIM) é também fraca, com coeficientes entre −0,007 e 0,113. Apesar da maioria dos coeficientes atingir significância estatística, esse

resultado é largamente atribuível à dimensão das subamostras. Observa-se, contudo, um padrão consistente no setor G, onde a associação é sistematicamente mais forte nas subamostras de diferenciação do que nas de liderança em custos, sendo nestas últimas indistinguível de zero. Trata-se de uma regularidade descritiva, cuja interpretação substantiva exigiria testes adicionais.

Analisando os resultados, no setor G, a diferença entre estratégias aponta sempre na mesma direção, nas três correlações e nos três anos. Nas subamostras de diferenciação, a associação entre dimensão e performance é consistentemente superior (0,092 a 0,113) à observada nas subamostras de liderança em custos (0,012 a 0,034), onde é indistinguível de zero. Inversamente, a associação entre dimensão e idade é sempre mais forte nas empresas com orientação para custos (0,293 a 0,325) do que nas orientadas para a diferenciação (0,195 a 0,211). Já a relação entre idade e performance é praticamente idêntica nos dois grupos, situando-se em torno de −0,11 em todas as células, o que sugere que a antiguidade da empresa se associa negativamente à rendibilidade de forma independente da estratégia adotada.

No setor C não se observa uma separação equivalente entre estratégias, os coeficientes dos dois grupos sobrepõem-se em todas as dimensões e não se identifica uma ordenação estável ao longo dos três anos. A única regularidade relevante é a maior intensidade da associação entre dimensão e idade neste setor (0,318 a 0,382) face ao setor G, transversal a ambas as estratégias.

Concluída esta fase, os dados foram segmentados por setor de atividade (CAE), sendo que para fazer a análise foi realizada uma média dos valores dos 3 anos para conseguir simplificar a leitura dos resultados.

IV- Análises e resultados

4.1 Análise dos resultados da regressão linear

Tabela 3: Coeficientes da regressão linear dos três anos

	CAE C		CAE G	
	β	Sig.	β	Sig.
Constante	,019	<,001	,001	,675
Estratégia (EST)	,019	<,001	,020	<,001
Dimensão (DIM)	,016	<,001	,026	<,001
Idade (IDA)	,000	<,001	,000	<,001
R^2 ajustado	,038		,065	

Nota: β = coeficiente não padronizado; Sig. = nível de significância

A variável estratégia (EST) apresenta um coeficiente positivo e estatisticamente significativo em ambos os setores (β = 0,019 no CAE C; β = 0,020 no CAE G). Este resultado indica que, em média e controlando para a dimensão e a idade, as empresas classificadas como líderes em custos apresentam um ROA superior em aproximadamente 1,9 a 2,0 pontos percentuais face às empresas com estratégia de diferenciação. A consistência deste efeito nos dois setores reforça a seguinte conclusão: a liderança em custos está positivamente associada à performance económica. A dimensão da empresa (DIM) revela igualmente um efeito positivo e significativo sobre o ROA em ambos os setores (β = 0,016 no CAE C; β = 0,026 no CAE G), o que associa a maior escala a economias de produção, maior poder negocial e acesso mais eficiente a recursos financeiros. A idade (IDA) apresenta um coeficiente positivo, mas muito próximo de zero, sendo, contudo, estatisticamente significativo em ambas as estimações. Mas a conclusão retirada é que a idade tem impacto muito reduzido no ROA e por isso não produz qualquer efeito no modelo.

Os modelos apresentam um R^2 ajustado reduzido (0,038 no CAE C e 0,065 no CAE G), o que significa que a estratégia, a dimensão e a idade explicam apenas uma fração da variabilidade do ROA. Este resultado é expectável pois a rendibilidade dos ativos é determinada por um conjunto muito alargado de fatores, estrutura de custos, posicionamento no mercado, qualidade da gestão, condições setoriais e conjunturais, dos quais o modelo capta deliberadamente apenas três. O objetivo da estimação não é a previsão do ROA de cada empresa, mas a identificação de um efeito médio associado ao posicionamento estratégico, controlando dois fatores estruturais. Nesse plano, o que releva é a magnitude, o sinal e a estabilidade do coeficiente de EST entre setores, e não a capacidade explicativa global do modelo.

Em síntese, os resultados obtidos confirmam a H1. A hipótese estabelecia que a utilização de uma estratégia comercial estaria associada a um impacto específico na performance económica, esperando-se que esse impacto fosse distinto consoante o tipo de estratégia escolhido. O coeficiente da variável EST é positivo e estatisticamente significativo em ambos os setores, o que indica que as empresas orientadas para a liderança em custos e as empresas orientadas para a diferenciação não apresentam níveis equivalentes de performance, mesmo após controlo da dimensão e da idade. A H1 é, portanto, corroborada. Para além de corroborar a hipótese conseguimos verificar qual das estratégias tem um maior impacto na performance. Neste caso, irá ser a liderança em custos, com um ROA superior em cerca de 1,9 a 2,0 pontos percentuais. O facto de este resultado se repetir nos dois setores analisados, que têm estruturas competitivas diferentes, sugere que não se trata de uma particularidade de um setor em concreto.

4.2 Análise dos resultados das regressões com variáveis multiplicativas

Com o intuito de complementar a análise precedente, foi procurado aferir a eventual não linearidade das variáveis do modelo, por isso o modelo foi expandido para criar 2 novos modelos: no primeiro, inclui uma nova variável, EST*DIM, que designei por dimensão multiplicativa; no segundo, inclui a variável EST*IDA, que designei por idade multiplicativa. Seguem os novos modelos criados:

Modelo 1: $ROAit = B0 + B1ESTit + B2DIMit + B3\,IDAit + B4\,EST * DIMit + u$

Modelo 2: $ROAit = B0 + B1ESTit + B2DIMit + B3\,IDAit + B4\,EST * IDAit + u$

As variáveis EST*DIM e EST*IDA são termos de interação, obtidos pelo produto entre a variável binária da estratégia (EST) e cada uma das variáveis de controlo, respetivamente a dimensão (DIM) e a idade (IDA) da empresa.

O objetivo da sua inclusão é de testar se o efeito da estratégia comercial sobre a performance económica é constante em todas as empresas, ou se depende das características estruturais de cada uma. No modelo base, o coeficiente de EST mede um efeito médio único, aplicado de igual forma a todas as empresas da amostra, independentemente da sua dimensão ou antiguidade. Esta especificação assume, portanto, que a vantagem associada a uma estratégia é a mesma numa empresa de pequena dimensão e numa empresa de grande dimensão.

No Modelo 1, o coeficiente B4 indica em quanto o efeito da dimensão sobre o ROA difere entre empresas com estratégia de liderança em custos e empresas com estratégia de

diferenciação. Um coeficiente estatisticamente significativo revela que a dimensão não se traduz em rendibilidade da mesma forma nos dois grupos estratégicos, isto é, que o benefício de crescer depende do posicionamento adotado. Um coeficiente nulo ou não significativo indica que o efeito da dimensão é idêntico nos dois grupos, sendo a relação, nesse aspeto, linear. O Modelo 2 replica esta lógica para a idade da empresa.

Estas duas variáveis têm como objetivo, assim, verificar a não linearidade na relação das variáveis do estudo: não uma não linearidade na própria variável dependente, mas a possibilidade de o efeito da estratégia variar consoante o perfil estrutural da empresa.

Os modelos foram regredidos para ambos os setores da CAE.

Estas regressões permitiram analisar se o efeito da estratégia competitiva sobre a performance económica é não linear em função das características das empresas. Em particular, procuram verificar se a dimensão e a idade das empresas reforçam ou atenuam a influência da estratégia competitiva na performance económica.

Com base nos resultados foi verificado que no modelo 2 a variável idade multiplicativa não apresentava impacto no ROA pois apresentava $\beta = 0$ em ambos os CAE. Como já tinha sido concluído anteriormente esta também não apresentava impacto no modelo base principal. Já o modelo 1 apresenta resultados diferentes para a variável dimensão multiplicativa. Quando controlamos o modelo para a estratégia e para a dimensão, esta apresenta um impacto maior nas empresas que adotam uma estratégia da diferenciação como podemos verificar na tabela seguinte.

Tabela 4: Coeficientes dos modelos expandidos

	CAE C — EST*DIM		CAE G — EST*DIM	
	β	Sig.	β	Sig.
Estratégia de diferenciação (EST= 0)	,018	<,001	,034	<,001
Estratégia de custos (EST= 1)	,010	<,001	,012	<,001

Assim a conclusão é que a dimensão empresarial contribui mais significativamente para a performance económica das empresas que seguem a estratégia de diferenciação. Por sua vez, nas empresas que seguem a estratégia de liderança em custos, o benefício adicional decorrente do aumento da dimensão revela-se menos expressivo.

Portanto os resultados indicam que a dimensão contribui de forma mais expressiva para a performance económica das empresas que seguem uma estratégia de diferenciação do que para as que seguem uma estratégia de liderança em custos. Pois nas primeiras, um aumento

da dimensão traduz-se num acréscimo de ROA quase duplo no CAE C e cerca de três vezes superior no CAE G, face ao observado nas segundas.

Esta diferença decorre da natureza dos recursos que cada estratégia utiliza. A diferenciação assenta em investimentos cujo retorno depende criticamente da escala a que são explorados, gastos em investigação e desenvolvimento, em publicidade e na construção de marca constituem, na sua maior parte, custos fixos, que não variam proporcionalmente com o volume de vendas. Uma empresa de maior dimensão pode diluir estes custos por uma base de atividade mais ampla, extraindo retorno de um investimento que uma empresa pequena teria dificuldade em suportar, ou suportaria em condições de rentabilidade desfavoráveis. Acresce que estas empresas enfrentam maior incerteza, uma vez que os investimentos em inovação envolvem produtos cujo desempenho comercial é desconhecido à partida (Abdelqader et al., 2024), pelo que a dimensão funciona também como capacidade de absorver insucessos pontuais sem comprometer a atividade. Neste sentido, a evidência é consistente com a proposição de Ortega (2010) de que quanto maiores os recursos ao dispor da empresa, maior a sua capacidade de sustentar uma estratégia que gere vantagem competitiva.

Na liderança em custos, o mecanismo competitivo é diferente. A vantagem resulta sobretudo da eficiência dos processos, do controlo de gastos e da rotação do ativo, dimensões em que o ganho marginal associado ao crescimento tende a esgotar-se mais cedo. As economias de escala são relevantes, mas estão em larga medida já incorporadas no próprio posicionamento estratégico destas empresas, o que ajuda a explicar que o benefício adicional decorrente do aumento da dimensão se revele menos expressivo.

Daqui resulta que a desvantagem média da diferenciação face à liderança em custos, identificada no modelo base, não é uniforme ao longo da amostra. Concentra-se nas empresas de menor dimensão, onde os investimentos exigidos pela diferenciação não encontram escala suficiente para gerar retorno. É um resultado com relevância particular no contexto português, marcado pela predominância de empresas de reduzida dimensão.

4.3 Análise dos resultados das árvores de regressão

A análise dos resultados obtidos através das árvores de regressão exige uma interpretação diferente da realizada para as regressões lineares. As árvores de regressão que utilizam o algoritmo *CHAID* produzem estruturas ramificadas com alguma extensão, resultantes de sucessivas repartições da amostra com base nos valores das variáveis explicativas, o que, embora enriqueça a capacidade analítica do modelo, dificulta a leitura direta dos resultados

(Delen et al., 2013). Para evitar esta limitação interpretativa, e sem comprometer o rigor analítico, optou-se por apresentar os resultados de maneira simplificada, inicialmente com os resultados da variável estratégia e posteriormente das variáveis de controlo.

Dado tratar-se de uma metodologia pouco frequente neste tipo de estudos, importa clarificar previamente o significado dos indicadores apresentados nas tabelas seguintes. A média da performance corresponde ao valor médio do ROA das empresas incluídas em cada nó terminal da árvore, ou seja, ao valor que o modelo prevê para todas as empresas que partilham aquele conjunto de características. Não se trata de um coeficiente, como nas regressões, mas de uma média condicional, o ROA esperado dado o perfil da empresa. O risco da estimativa constitui a medida de erro de previsão do modelo, correspondendo ao erro quadrático médio calculado sobre os dados de treino, isto é, ao desvio médio entre os valores de ROA observados e os valores previstos em cada nó terminal (Delen et al., 2013).

Tabela 5: Resultados do impacto do tipo de estratégia na performance

	CAE C		CAE G	
	Liderança em Custos	Diferenciação	Liderança em Custos	Diferenciação
Média da Performance	,081	,061	,090	,065
Nº de observações	5 062	15 302	12 357	30 879
% Nº obs.	24,9	75,1	28,6	71,4
Risco da estimativa	,004		,004	

Com base nos resultados observados na tabela foi possível verificar que para ambos os setores da CAE o valor da média da performance (ROA) é maior para a estratégia de liderança em custos. Apesar da diferença entre os valores não ser elevada em termos absolutos, importa considerar a escala em que o ROA se distribui. Tratando-se de um rácio que assume valores num intervalo reduzido, uma diferença de dois pontos percentuais corresponde a um acréscimo de cerca de um terço sobre a média do grupo com menor performance, o que lhe confere relevância. Assim foi concluído, como já tinha sido verificado a partir dos resultados das regressões anteriores, que as empresas que optam por seguir a liderança em custos apresentam valores de ROA mais elevados e por isso uma performance económica melhor.

Estes resultados permitem também corroborar a H1. As médias de ROA obtidas nos nós terminais das árvores são efetivamente distintas entre os dois grupos estratégicos, e essa distinção verifica-se nos dois setores da CAE, pelo que a H1 é confirmada. O sentido da diferença, favorável à estratégia liderança em custos, constitui informação adicional que a hipótese não antecipava e que vem no mesmo sentido do já observado nas regressões. Importa ainda salientar que este resultado é obtido por uma metodologia que não impõe

qualquer forma funcional à relação entre as variáveis, ao contrário do que sucede na regressão linear. O facto de duas técnicas com pressupostos tão distintos conduzirem à mesma conclusão reforça a confiança no suporte dado à H1.

Para avaliar a qualidade do modelo utilizamos o risco de estimativa que constitui uma medida de erro de previsão do modelo de árvore de regressão. Quanto menor o valor do risco estimado, maior a capacidade preditiva do modelo, ou seja, menor a dispersão dos valores de ROA em torno da média prevista em cada folha da árvore. Um risco estimado reduzido indica que o modelo consegue prever com superior precisão o nível de performance económica das empresas com base nas variáveis incluídas, o que reforça a validade das conclusões retiradas da análise. Nas arvores feitas anteriormente, o valor do risco estimado é igual a 0,004 em todos os modelos estimados, o que sugere uma capacidade preditiva consistente entre as variáveis.

Tabela 6: Resultados do impacto das variáveis de controlo na performance

	CAE C				CAE G			
	Menor idade	Maior idade	Menor dimensão	Maior dimensão	Menor idade	Maior idade	Menor dimensão	Maior dimensão
Média da Performance	,066	,079	,053	,084	,055	,087	,050	,091
Nº de observações	2 988	1 927	2 036	4 073	3 466	5 556	6 660	4 324
% Nº obs.	14,7	9,5	10	20	8	12,9	15,4	10
Risco Estimativa	,004				,004			

No sentido de facilitar a análise dos resultados, optou-se por tabelar apenas os extremos das distribuições das variáveis DIM e IDA, denominadas como "menor", o extremo inferior e "maior", o extremo superior.

Utilizou-se esta abordagem pois numa primeira abordagem, procurou-se reduzir a dimensão das árvores através de técnicas de poda (*pruning*), de modo a obter estruturas mais compactas e de leitura mais fácil. Contudo, este procedimento não obteve os resultados esperados, uma vez que as árvores resultantes ou perdiam as repartições relevantes para a análise, ou mantinham uma extensão que continuava a dificultar a leitura e a análise da árvore. Optou-se, assim, por analisar as árvores completas e apresentar apenas os valores dos nós extremos, assim esta solução mantém toda a informação necessária à interpretação sem perder a legibilidade dos resultados.

Os coeficientes da média da performance representam o impacto de cada variável e intervalo na performance. A partir dos valores tabelados, constata-se que para ambos os setores as

empresas que apresentam maior idade e maior dimensão, independente da estratégia, apresentam superior performance.

Nestas árvores o valor do risco estimado é igual às árvores apresentadas anteriormente, o que sugere uma capacidade preditiva consistente entre as variáveis.

Comparando estes resultados com os obtidos através das regressões lineares, verifica-se que existe convergência quanto à dimensão empresarial, mas divergência quanto à idade. No caso da dimensão, ambas as metodologias apontam no mesmo sentido, nas regressões o coeficiente da variável DIM era positivo e significativo em ambos os setores e nas árvores de regressão a diferença de ROA entre as empresas de menor e de maior dimensão é de 3,1 pontos percentuais no CAE C e de 4,1 pontos percentuais no CAE G. A dimensão surge, assim, como um determinante consistente da performance económica, independentemente da técnica utilizada.

Já quanto à idade, os resultados não são coincidentes. Nas regressões, a variável IDA apresentava um coeficiente de magnitude praticamente nula, tendo sido considerada como não tendo relevância económica para explicar o ROA. Nas árvores de regressão, pelo contrário, a idade aparece como uma variável relevante, com uma diferença de 1,3 pontos percentuais no CAE C e de 3,2 pontos percentuais no CAE G entre os extremos da sua distribuição. No caso do CAE G, esta diferença é inclusivamente superior à que se verifica entre as duas estratégias.

Esta divergência não significa que os métodos se contradigam, mas resulta da forma diferente como cada um trata as variáveis. A regressão linear estima um efeito constante ao longo de toda a distribuição da idade, assumindo que cada ano adicional produz sempre o mesmo acréscimo de ROA. Se o efeito da idade se fizer sentir apenas nos valores extremos, ou se for de sinal contrário consoante o intervalo, esses efeitos anulam-se mutuamente e o coeficiente aproxima-se de zero. As árvores de regressão não impõem essa restrição, uma vez que repartem a amostra em intervalos e estimam uma média para cada um deles, conseguindo por isso captar efeitos que se manifestam apenas em determinados segmentos. O que estes resultados dizem é que a idade influencia a performance de forma não linear, o que explica a sua aparente irrelevância no modelo de regressão e confirma a utilidade da aplicação conjunta das duas metodologias, tal como fundamentado na secção 3.4.

Concluindo, ambas as metodologias convergem quanto ao resultado principal do estudo, ainda que difiram na forma como captam o efeito da idade: as empresas que seguem a liderança em custos apresentam ROA superior às que seguem a diferenciação, em ambos os setores, com um diferencial que não é, contudo, uniforme, uma vez que varia com a dimensão

e a idade das empresas. A H1 fica assim confirmada, quer quanto à existência de um impacto distinto consoante a estratégia, quer quanto à consistência desse impacto entre as diferentes metodologias.

V- Conclusões

A presente dissertação teve como objetivo central analisar de que forma diferentes estratégias comerciais influenciam a performance económica das empresas portuguesas, testando empiricamente se a escolha entre a liderança em custos e a diferenciação produz efeitos distintos sobre a rentabilidade, aferida pelo rácio *Return on Assets* (ROA).

Os resultados da regressão linear revelam que a variável estratégia (EST) apresenta um coeficiente positivo e estatisticamente significativo em ambos os setores analisados. Este resultado indica que as empresas classificadas adotantes de uma estratégia de custos apresentam, em média, um ROA superior em cerca de 2,0 pontos percentuais face às empresas que seguem a estratégia de diferenciação, controlando os efeitos da dimensão e da idade. Este resultado corrobora a hipótese subjacente ao estudo.

Os resultados das árvores de regressão confirmam esta conclusão, as empresas que seguem a estratégia de liderança em custos apresentam uma média de ROA superior em ambos os setores da CAE, comparativamente às empresas que seguem a estratégia de diferenciação.

A consideração de variáveis multiplicativas na análise de regressão adicionou uma perspetiva complementar: o efeito moderador da dimensão na performance não é linear, sendo mais expressivo nas empresas que seguem a estratégia de diferenciação do que nas que seguem a estratégia de liderança em custos. Uma potencial razão é que as empresas de maior dimensão, com maiores recursos, são as que melhor podem adotar este tipo de estratégia e refletir isso na performance. Os resultados das árvores de regressão confirmam esta conclusão da não linearidade da dimensão, apresentando também evidência empírica de que o mesmo acontece com a idade das empresas. Este último efeito não era percecionado a partir da análise de regressão.

Em suma, a hipótese formulada, que está subjacente ao presente trabalho de investigação, de que a escolha entre estratégias comerciais distintas está associada a efeitos diferenciados na performance económica, é confirmada pela evidência empírica obtida. A estratégia de liderança em custos associa-se de forma consistente e estatisticamente significativa a níveis superiores de ROA face à estratégia de diferenciação, em ambos os setores analisados. Esta evidência é robusta, pois mantém-se estável nas duas metodologias utilizadas.

Esta dissertação posiciona-se numa linha de investigação que assenta na relação entre estratégias genéricas de Porter (1980; 1985) e performance empresarial. A maioria dos estudos que analisa esta relação, nomeadamente Abdelqader et al. (2024), Hansen et al.

(2015), Jermias (2008), Valipour et al. (2012) foi conduzida em economias de maior dimensão, com empresas cotadas ou em setores específicos com estruturas competitivas distintas das observadas no mercado português. Porém, os resultados desta dissertação confirmam, no contexto das empresas portuguesas, as conclusões de Dess and Davis (1984) e Valipour et al. (2012) de que a liderança em custos tende a produzir níveis de ROA superiores às de uma estratégia de diferenciação. Eles contrastam, no entanto, com os resultados de Hansen et al. (2015), que encontraram uma performance superior para as empresas com estratégias de diferenciação no setor florestal norte-americano, resultados estes que podem, potencialmente, ser justificados pelo contexto sectorial e geográfico que está subjacente ao estudo.

Podem referir-se alguns contributos principais do presente estudo. O tratar-se de um estudo inovador para o contexto empresarial português. O ter subjacente um desenho metodológico comparativo, tomando e confrontando os setores distintos da CAE, o que permitiu testar a hipótese em contextos operacionais diferentes dentro do mesmo quadro institucional. A maioria dos estudos de referência analisa amostras sectorialmente homogéneas, o que limita a generalização das suas conclusões. Em terceiro, a combinação metodológica de técnicas de *clustering*, regressão linear e árvores de regressão, que representa uma abordagem mais abrangente do que a utilizada em estudos anteriores e, simultaneamente, proporciona maior robustez aos resultados obtidos.

Porém, esta dissertação não é isenta de limitações. Em primeiro lugar, a operacionalização das estratégias comerciais assentou em dois rácios financeiros. Embora esta abordagem seja consistente com a literatura, nomeadamente com Jermias (2008), trata-se de *proxies* que captam apenas dimensões parciais dos posicionamentos estratégicos reais das empresas. A ausência de dados sobre gastos de investigação e desenvolvimento na base SABI impediu a inclusão de uma variável que a literatura reconhece como central na operacionalização da estratégia de diferenciação, os gastos de I&D, o que pode ter condicionado a precisão da classificação estratégica de algumas empresas. Outra limitação foi de não ter considerado que podem existir empresas que, simultaneamente, utilizam diferentes estratégias para distintas linhas de produtos. Por último, o horizonte temporal de três anos, embora adequado para analisar volatilidades anuais, pode não ser suficiente para captar os efeitos de estratégias que produzem resultados a médio e longo prazo, como é frequentemente o caso de uma estratégia de diferenciação, cujos investimentos em inovação e construção de marca têm um horizonte de retorno tipicamente mais alargado. No entanto, pesem tais limitações, considero que as mesmas, dada a robustez da evidência empírica

recolhida, não condicionam as conclusões acima referidas.

As limitações identificadas apontam naturalmente para várias direções de investigação que poderiam aprofundar e ampliar os contributos desta dissertação. Uma primeira sugestão prende-se com a extensão da análise a outros setores de atividade, nomeadamente serviços e tecnologia, onde a relação entre estratégia e performance pode apresentar dinâmicas substancialmente diferentes das observadas na indústria transformadora e no comércio.

Uma segunda linha de investigação futura consistiria na incorporação de variáveis que captem mais diretamente a intensidade de inovação, como despesas em I&D, número de patentes registadas ou indicadores de novos produtos lançados, utilizando fontes de dados que disponibilizem esta informação para um conjunto representativo de empresas portuguesas. Tal permitiria uma operacionalização mais precisa da estratégia de diferenciação e uma comparação mais informada com os resultados aqui obtidos.

Por fim, seria relevante alargar o horizonte temporal da análise de forma a incluir pelo menos um ciclo económico completo, permitindo avaliar se a vantagem da liderança em custos sobre a diferenciação se mantém em períodos de recessão e de expansão, ou se é sensível à fase do ciclo económico.

Referências bibliográficas

Abdelqader, M., Uyar, A., & Kuzey, C. (2024). CSR transparency and firm value: Is the connection differential for cost leaders and differentiators? *Journal of Cleaner Production*, *443*. https://doi.org/10.1016/j.jclepro.2024.141102

Agustia, D., Muhammad, N. P. A., & Permatasari, Y. (2020). Earnings management, business strategy, and bankruptcy risk: evidence from Indonesia. *Heliyon*, *6*(2), e03317. https://doi.org/10.1016/j.heliyon.2020.e03317

Al-Shaer, H., Kuzey, C., Uyar, A., & Karaman, A. S. (2023). Corporate strategy, board composition, and firm value. *International Journal of Finance & Economics*, *29*(3), 3177–3202. https://doi.org/10.1002/ijfe.2827

Anthony, J. H., & K., R. (1992). Association between accounting performance measures and stock prices. *Journal of Accounting and Economics*, *15*, 203–227.

Chen, C.-J., Guo, R.-S., Hsiao, Y.-C., & Chen, K.-L. (2018). How business strategy in non-financial firms moderates the curvilinear effects of corporate social responsibility and irresponsibility on corporate financial performance. *Journal of Business Research*, *92*, 154–167. https://doi.org/10.1016/j.jbusres.2018.07.030

Cho, S. Y., & Lee, C. (2017). Managerial Efficiency, Corporate Social Performance, and Corporate Financial Performance. *Journal of Business Ethics*, *158*(2), 467–486. https://doi.org/10.1007/s10551-017-3760-7

Cowen, S. S., & Hoffer, J. A. (1982). *Usefulness of financial ratios in a single industry* (Vol. 10). https://doi.org/10.1016/0148-2963(82)90020-0

Delen, D., Kuzey, C., & Uyar, A. (2013). Measuring firm performance using financial ratios: A decision tree approach. *Expert Systems with Applications*, *40*(10), 3970–3983. https://doi.org/10.1016/j.eswa.2013.01.012

Dess, G. G., & Davis, P. S. (1984). *Porter's (1980) Generic Strategies as Determinants of Strategic Group Membership and Organizational Performance. Academy of Management Journal* (Vol. 27). https://doi.org/http://dx.doi.org/10.2307/256040

Dou, W. W., Ji, Y., & Wu, W. (2021). Competition, profitability, and discount rates. *Journal of Financial Economics*, *140*(2), 582–620. https://doi.org/10.1016/j.jfineco.2020.12.013

Farida, I., & Setiawan, D. (2022). Business Strategies and Competitive Advantage: The Role of Performance and Innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, *8*(3). https://doi.org/10.3390/joitmc8030163

Hansen, E., Nybakk, E., & Panwar, R. (2015). Pure versus hybrid competitive strategies in the forest sector: Performance implications. *Forest Policy and Economics*, *54*, 51–57.

https://doi.org/10.1016/j.forpol.2015.02.001

Ikotun, A. M., Habyarimana, F., & Ezugwu, A. E. (2025). Cluster validity indices for automatic clustering: A comprehensive review. *Heliyon*, *11*(2), e41953. https://doi.org/10.1016/j.heliyon.2025.e41953

Jermias, J. (2008). The relative influence of competitive intensity and business strategy on the relationship between financial leverage and performance. *The British Accounting Review*, *40*(1), 71–86. https://doi.org/10.1016/j.bar.2007.11.001

Khan, I. K., Daud, H. B., Zainuddin, N. B., Sokkalingam, R., Farooq, M., Baig, M. E., Ayub, G., & Zafar, M. (2024). Determining the optimal number of clusters by Enhanced Gap Statistic in K-mean algorithm. *Egyptian Informatics Journal*, *27*. https://doi.org/10.1016/j.eij.2024.100504

Linares-Mustarós, S., Coenders, G., & Vives-Mestres, M. (2018). Financial performance and distress profiles. From classification according to financial ratios to compositional classification. *Advances in Accounting*, *40*, 1–10. https://doi.org/10.1016/j.adiac.2017.10.003

Ortega, M. J. R. (2010). Competitive strategies and firm performance: Technological capabilities' moderating roles. *Journal of Business Research*, *63*(12), 1273–1281. https://doi.org/10.1016/j.jbusres.2009.09.007

Panwar, R., Nybakk, E., Hansen, E., & Pinkse, J. (2016). The effect of small firms' competitive strategies on their community and environmental engagement. *Journal of Cleaner Production*, *129*, 578–585. https://doi.org/10.1016/j.jclepro.2016.03.141

Park, W., Sung, C. S., & Byun, C. G. (2019). Impact of Unlisted Small and Medium-Sized Enterprises' Business Strategies on Future Performance and Growth Sustainability. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(3). https://doi.org/10.3390/joitmc5030060

Porter, M. E. (1980). *Competitive strategy: Techniques for analyzing industries and competitors*. Free Press

Porter, M. E. (1985). *Competitive advantage: Creating and sustaining superior performance*. Free Press

Prastyabudi, W. A., Alifah, A. N., & Nurdin, A. (2024). Segmenting the Higher Education Market An Analysis of Admissions Data Using K-Means Clustering. *Procedia Computer Science*.

Probohudono, A. N., Suhardjanto, D., Aligarh, F., Chayati, N., & Putra, A. A. (2025). Navigating MSMEs' performance through innovation and digital IT capabilities in business strategy. *Social Sciences & Humanities Open*, *12*.

https://doi.org/10.1016/j.ssaho.2025.101810

Sutton, C. D. (2005). *Classification and Regression Trees, Bagging, and Boosting* (Vol. 24). https://doi.org/10.1016/S0169-7161(04)24011-1

Teece, D. J. (2018). Business models and dynamic capabilities. *Long Range Planning*, *51*(1), 40–49. https://doi.org/10.1016/j.lrp.2017.06.007

Teixeira, N., Lisboa, I., Costa, T., & Godinho, T. (2019). Business strategies and financial performance. *Handbook of Research on Entrepreneurship, Innovation, and Internationalization*

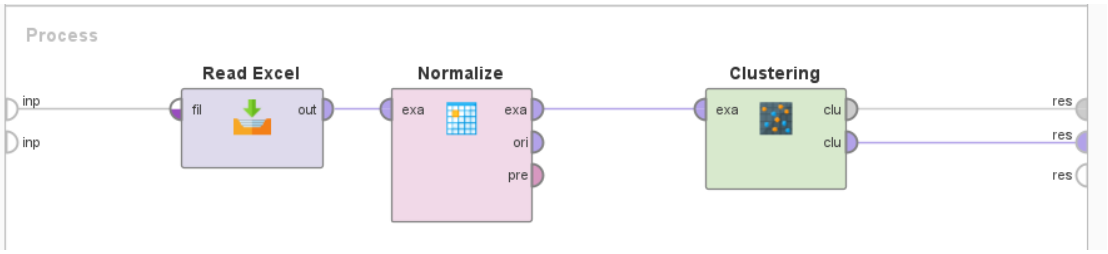
Valipour, H., Birjandi, H., & Honarbakhsh, S. (2012). The Effects of Cost Leadership Strategy and Product Differentiation Strategy. *Journal of Asian Business Strategy*, *2*, 14–23.

Anexo

Caracterização da estratégia comercial

No presente texto é apresentado o processo de análise realizada aos dados. Para a análise de clusters foi utilizado o *AI Studio RapidMiner* versão 2025.1.0 e foi desenvolvido o seguinte processo:

Figura 1: Processo do *RapidMiner*



Este processo tem como objetivo o de separar os dados em segmentos de acordo com a estratégia comercial adotada. Para se realizar este processo teve-se em primeiro lugar de selecionar o K (que representa o nº de clusters que vamos repartir). A seleção do K, número de clusters, trata-se como se fosse um problema de seleção em aprendizagem não supervisionada, uma vez que este parâmetro não é conhecido à priori (Khan et al., 2024). Neste sentido, foi adotado uma abordagem de "tentativa e erro", onde se atribuíram diferentes valores para K, até encontrar a melhor solução de dados para o modelo. Este processo é consistente com a prática iterativa dos métodos de *clustering* e com os procedimentos recomendados para a identificação de soluções estruturalmente estáveis (Khan et al., 2024). A comparação entre múltiplas soluções de *clustering* permite selecionar aquela que melhor representa a estrutura subjacente dos dados, sendo este processo interpretado na literatura como um procedimento de seleção de modelo suportado por critérios de validade interna (Ikotun et al., 2025).

A avaliação da qualidade dos diferentes segmentos baseia-se no equilíbrio entre a coesão *intra-cluster* e a separação *inter-cluster*, dimensões centrais na análise de validade de *clusters*. Os índices de validade interna avaliam precisamente estas duas propriedades, a homogeneidade dentro de cada grupo e o afastamento entre grupos distintos, constituindo instrumentos fundamentais para suportar a decisão sobre o número ótimo de *clusters* (Ikotun et al., 2025). Neste estudo, foi selecionado o índice de *Davies-Bouldin* para avaliar os *clusters*. Este índice mede os *clusters* com base na relação entre a dispersão *intra-cluster* e a separação *inter-cluster*. Especificamente, o índice calcula, para cada *cluster*, a razão entre a soma da dispersão desse *cluster* e do seu *cluster* mais próximo, e a distância entre os respetivos centroides, sendo depois calculada a média dessas razões máximas em todos os *clusters* o

índice de *Davies-Bouldin* encontra a similaridade média *inter-cluster* entre quaisquer dois *clusters* e o seu vizinho mais próximo. Valores mais baixos do índice indicam os *clusters* mais compactos internamente e mais distintos entre si (Ikotun et al., 2025). Este índice foi calculado para as soluções de *clustering* obtidas em cada setor (CAE C e CAE G) e em cada ano da amostra (2022-2024). Estes deram soluções entre (0,4 e 0,5) o que corresponde a índices satisfatórios quando combinados com o restante procedimento e com objetivo desta análise.

A escolha do número de *clusters* resulta, assim, da junção entre a qualidade dos índices de validade interna e a da interpretação das soluções obtidas, assegurando que os grupos identificados são simultaneamente consistentes do ponto de vista estatístico e relevantes do ponto de vista analítico (Khan et al., 2024).

Após a definição do número de *clusters*, foi introduzido um critério adicional relativo à dimensão mínima dos grupos, sendo que apenas considerei os dois maiores grupos de clusters pois o objetivo da análise foi de conseguir separar os dados entre as duas estratégias. Esta opção metodológica já foi realizada anteriormente na literatura sobre *clusters,* onde em situações que existam bases de dados de grande dimensão, a presença de grupos com reduzida expressão tende a refletir observações atípicas ou estruturas pouco representativas da distribuição global dos dados, comprometendo a capacidade de sumarização e interpretação do modelo (Linares-Mustarós et al., 2018). Assim, quando se utilizam rácios financeiros como variáveis de *clusters*, é reconhecida a probabilidade para a formação de *clusters* de muito pequena dimensão compostos por observações com valores extremos, seja muito elevados ou muito reduzidos, que não constituem grupos estrategicamente coerentes, mas sim casos atípicos que distorcem a solução final (Cowen & Hoffer, 1982). Assim, neste tipo de situações existe a necessidade de separar o subconjunto de observações que não se enquadra nas estruturas principais, permitindo melhorar a qualidade dos grupos obtidos e a estabilidade dos centroides dos clusters (Linares-Mustarós et al., 2018). Esta abordagem segue o princípio de que o *clustering* deve maximizar a homogeneidade intra-grupo e a heterogeneidade entre grupos, evitando a criação de segmentos residuais que não constituem estruturas naturais dos dados e que prejudicam a capacidade explicativa da segmentação (Ikotun et al., 2025).

A aplicação do algoritmo de *clusters* permitiu identificar diferentes grupos de empresas para cada setor de atividade (CAE C e CAE G) sendo que foi utilizado o ano de 2024 como referência para fazer a distinção das estratégias por ser o ano mais recente da pesquisa e consequentemente o ano mais representativo das condições competitivas atuais

das empresas analisadas. Para a separação do ano de 2024, depois de serem realizadas as tentativas para determinar o K, foi selecionado para o CAE C o K=3, sendo que foram excluídas 8 empresas como *outliers*, e para o CAE G foi selecionado K=4, sendo que foram excluídas 7 empresas como *outliers*. A tabela seguinte apresenta a distribuição do número de empresas por cluster, evidenciando a dimensão relativa de cada segmento resultante do processo de segmentação.

Tabela 7: Segmentação de *clusters*

	CAE C	
	2024	
	Cluster 0	Cluster 1
Nº de empresas	15 302	5 062

	CAE G	
	2024	
	Cluster 0	Cluster 1
Nº de empresas	30 879	12 357

Com base nos valores médios das variáveis estratégicas obtidos para cada *cluster*, procedeu-se à sua caracterização de acordo com as tipologias de estratégia competitiva. A operacionalização desta caracterização seguiu o enquadramento teórico das estratégias comerciais, segundo o qual diferentes combinações de indicadores económico-financeiros refletem posicionamentos estratégicos distintos ao nível da liderança em custos e da diferenciação (Porter, 1980, 1985). Neste sentido, os *clusters* que apresentam valores mais elevados na variável Vendas/Ativos foram associados a uma estratégia de liderança em custos, uma vez que este indicador traduz maior eficiência na utilização dos recursos e maior rotação do ativo, características associadas à estratégica orientada para a eficiência operacional. Por outro lado, os *clusters* que apresentam valores superiores na variável Margem/Vendas estão mais associados à estrutura de despesas e enquadram-se na estratégia de diferenciação. Isto porque tendem a gerar mais valor com os produtos diferenciadores, o que implica um investimento mais elevado (Jermias, 2008).

Contudo, nem todos os *clusters* apresentam um padrão claramente alinhado com uma única estratégia competitiva. Assim, como foi referido anteriormente foram retirados os grupos considerados como *outliers* para cada divisão de cada CAE.

A tabela seguinte mostra os resultados e a classificação de cada *cluster*:

Tabela 8: Classificação de *Clusters*

CAE C

	2024	
	Grupo 1	Grupo 2
Nº de empresas	5 062	15 302
	Média	Média
Variável 1: Vendas/Ativos	2,07	0,79
Variável 2: Margem/Vendas	0,59	0,77

CAE G		
	2024	
	Grupo 1	Grupo 2
Nº de empresas	12 357	30 879
	Média	Média
Variável 1: Vendas/Ativos	2,55	0,92
Variável 2: Margem/Vendas	0,30	0,82

Grupos 1	Liderança em custos
Grupos 2	Diferenciação

No CAE C, a análise dos dados de 2024 permite identificar os dois perfis estratégicos. O Grupo 1, constituído por 5 062 empresas, regista um valor médio da Variável 1 (Vendas/Ativos) de 2,07, valor bastante superior ao do Grupo 2 que tem apenas uma média de 0,79, e uma Variável 2 (Margem/Vendas) de 0,59, valor mais baixo que a Variável do Grupo 2, foi concluído assim que o Grupo 1 caracteriza-se por seguir uma estratégia de liderança de custos, visto que tem os valores da Variável 1 mais elevados e os valores da Variável 2 mais reduzidos em comparação com o Grupo 2. Pelo contrário o Grupo 2, constituído por 15 302 empresas, apresenta valores mais elevados da Variável 2 comparado como Grupo 1 e valores mais baixos de Variável 1 comparando com o outro Grupo por isso foi definido que o Grupo 2 segue uma estratégia de diferenciação.

No CAE G, a separação estratégica torna-se ainda mais visível. O Grupo 1, com 12 357 empresas, apresenta uma Variável 1 de 2,55, bastante superior à Variável 1 do Grupo 2 (de 0,92) e uma Variável 2 de apenas 0,30, traduzindo-se claramente num grupo que segue a estratégia de liderança em custos. O Grupo 2, composto por 30 879 empresas, distingue-se por uma Variável 2 de 0,82, claramente superior à Variável 2 do Grupo 1 de 0,30. Sendo este perfil identificado como um perfil que segue a estratégia de diferenciação.

Portanto foi concluído que em ambos os CAE o Grupo 1 representa o conjunto de empresas que seguem a estratégia de liderança em custos e o Grupo 2 representa o conjunto

de empresas que seguem a estratégia de diferenciação.

Após a identificação dos dois grupos de *clusters*, os resultados obtidos no *RapidMiner* foram exportados para o Microsoft Excel, procedeu-se à organização da informação por ano, sendo que para a posterior análise da regressão linear foi usado os dois grupos de clusters ao mesmo tempo através da variável binária, isto dentro de cada setor de atividade (CAE). Esta estruturação permitiu analisar as relações de como cada grupo influenciava a performance.

Numa fase subsequente, e já em Excel, foram incorporadas as variáveis de controlo definidas no modelo, possibilitando a construção de uma base de dados final específica para cada cluster. A introdução destas variáveis teve como finalidade controlar o efeito de características estruturais das empresas que a literatura identifica como potencialmente influentes na performance, permitindo assim uma estimação mais robusta das relações entre as variáveis principais do modelo (Jermias, 2008; Valipour et al., 2012).

Após a segmentação das empresas através do processo de *clustering* realizado para o cálculo da variável binária, calculou-se a média das variáveis ao longo dos três anos em análise. Esta opção metodológica teve como objetivo simplificar a análise e facilitar a interpretação dos resultados, permitindo obter uma medida mais representativa do comportamento das empresas durante o período considerado. Para além disso, a utilização de valores médios contribui para reduzir o impacto de flutuações ou circunstanciais verificadas em determinado ano, privilegiando uma perspetiva mais estável e consistente das características e da performance das empresas.

Concluída a preparação das bases de dados, procedeu-se à estimação dos modelos de regressão linear e das árvores de decisão recorrendo ao software *IBM SPSS Statistics*. Para o processo, foi utilizada a base de dados previamente construída em Microsoft Excel. À semelhança do procedimento usado na regressão linear, o cálculo das árvores de regressão foi realizado dentro de cada *cluster*, de forma a captar as especificidades das relações entre variáveis em função dos perfis estratégicos identificados na fase de *clustering*.

Para criação das árvores de regressão utilizou-se o algoritmo *CHAID (Chi-squared Automatic Interaction Detection)*, que procede à divisão da amostra através de testes de qui-quadrado, identificando as variáveis explicativas e os valores de corte que estabelecem partições estatisticamente significativas da variável dependente (Delen et al., 2013). Ao contrário do algoritmo *CART*, o *CHAID* permite divisões múltiplas em cada nó, gerando uma estrutura de árvore potencialmente mais ramificada e adequada à deteção de interações entre variáveis (Delen et al., 2013).

Os resultados foram obtidos sob a forma de representação gráfica em estrutura de

árvore de decisão, permitindo identificar, de forma visualmente intuitiva, as combinações de variáveis e os valores de corte associados a diferentes níveis de performance económica.