

Bruno Ricardo Barradas Santos

O IMPACTO DA LOCALIZAÇÃO E DA
DIMENSÃO NA RENDIBILIDADE
OPERACIONAL DOS HOTÉIS EM
PORTUGAL



2023

Bruno Ricardo Barradas Santos

O IMPACTO DA LOCALIZAÇÃO E DA DIMENSÃO NA RENDIBILIDADE OPERACIONAL DOS HOTÉIS EM PORTUGAL

Mestrado em Gestão de PME

Trabalho efetuado sob a orientação de:

Professora Adjunta Doutora Sandra Cristina Francisco Rebelo
Professor Coordenador Doutor Marco Paulo dos Santos Carrasco



2023

O IMPACTO DA LOCALIZAÇÃO E DA DIMENSÃO NA RENDIBILIDADE
OPERACIONAL DOS HOTÉIS EM PORTUGAL

Declaração de autoria de trabalho

Declaro ser o autor deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

Bruno Ricardo Barradas Santos

© Copyright: Bruno Ricardo Barradas Santos

A Universidade do Algarve reserva para si o direito, em conformidade com o disposto no Código do Direito de Autor e dos direitos Conexos, de arquivar, reproduzir e publicar a obra, independentemente do meio utilizado, bem como de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição para fins meramente educacionais ou de investigação e não comerciais, conquanto seja dado o devido crédito ao autor e editor respetivos.

À Ema, Sofia e Inês

Aos meus pais pela educação e pelos valores transmitidos

À minha família pelo apoio nesta caminhada

AGRADECIMENTOS

Este trabalho individual não seria possível sem o apoio de um conjunto de pessoas a quem estou muito grato.

Um especial agradecimento aos meus orientadores, Professora Doutora Sandra Cristina Francisco Rebelo e Professor Doutor Marco Paulo dos Santos Carrasco, pelo permanente apoio, disponibilidade e orientação.

Aos colegas hoteleiros que ajudaram, direta ou indiretamente, na recolha dos dados, essenciais ao estudo desenvolvido.

RESUMO

Este trabalho pretende estudar o impacto da localização e da dimensão na rendibilidade da hotelaria em Portugal. Atendendo à extensa orla costeira e à existência de diferentes níveis de concentração de unidades hoteleiras nas várias regiões do país, a localização foi analisada em termos de proximidade à costa e de aglomeração. Os restantes determinantes em análise são a dimensão, categoria e qualificação como PME. Como elemento inovador, é introduzido o GOPPAR como indicador de rendibilidade em alternativa ao RevPAR, amplamente utilizado em estudos anteriores. Com recurso à linguagem de programação *Python*, são desenvolvidos e aplicados modelos estatísticos regressivos a diferentes conjuntos de dados, obtidos através de inquérito e junto do Registo Nacional de Empreendimentos Turísticos, referentes a 87 unidades hoteleiras, relativos aos anos de 2017 a 2021. Os resultados apontam para uma forte relação na proximidade costeira e níveis de aglomeração com a rendibilidade operacional, quando medida por via do GOPPAR, relativamente ao período 2017-19. A dimensão, categoria ou qualificação como PME não se apresentaram estatisticamente relevantes. Valida-se a pertinência do estudo do GOPPAR na rendibilidade operacional e confirma-se a excecionalidade do período COVID nos estabelecimentos hoteleiros em Portugal.

Palavras-chave: Rendibilidade, hotel, localização, dimensão, GOPPAR

ABSTRACT

This study aims to examine the impact of location and size on the profitability of the hotel industry in Portugal. Given the extensive coastline and the presence of varying levels of concentration of hotels across different regions of the country, location is analysed in terms of proximity to the coast and agglomeration. Other determinants under study include size, category, and qualification as an SME (Small and Medium-sized Enterprise). As an innovative element, this study introduces GOPPAR (Gross Operating Profit Per Available Room) as an alternative profitability indicator to RevPAR (Revenue Per Available Room), which has been widely used in previous studies. Using Python programming language, regressive statistical models are developed and applied to various datasets, obtained through surveys and from the *Registo Nacional de Empreendimentos Turísticos*, pertaining to 87 hotels for the years 2017 to 2021. The findings indicate a strong relationship between coastal proximity and levels of agglomeration with operational profitability, as measured by GOPPAR, particularly for the period 2017-2019. Size, category, or qualification as an SME did not prove to be statistically significant. The study validates the relevance of examining GOPPAR in operational profitability and confirms the exceptional nature of the COVID period for hotel establishments in Portugal.

Keywords: profitability, hotel, location, dimension, GOPPAR

ÍNDICE GERAL

Agradecimentos	vi
Resumo	vii
Abstract.....	viii
Índice geral	ix
Índice de figuras	xii
Índice de tabelas	xiv
Lista de abreviaturas e siglas	xv
1. Introdução.....	1
1.1. Enquadramento	1
1.2. Abordagem teórica.....	3
1.3. Objetivos	6
1.4. Estrutura do trabalho.....	7
2. Caracterização do setor turístico português.....	9
2.1. Caracterização do alojamento turístico em Portugal	12
2.2. Caracterização dos estabelecimentos hoteleiros em Portugal.....	15
2.2.1 Oferta hoteleira.....	15
2.2.2 Dormidas	15
2.2.3 Proveitos	16
2.2.4 Investimentos.....	16
2.2.5 Distribuição territorial – a proximidade costeira.....	17
2.2.6 Enquadramento 2022.....	17
2.3. Síntese conclusiva.....	19
3. Revisão da literatura	20
3.1. Performance	20
3.1.1 A importância do USALI na normalização	22

3.1.2	Receitas de alojamento vs. receitas totais vs. resultado operacional	23
3.1.3	GOPPAR vs RevPAR.....	25
3.2.	Localização	27
3.2.1	Proximidade às zonas costeiras	28
3.2.2	Aglomeração do mercado.....	31
3.2.3	Outros fatores da localização.....	32
3.3.	Dimensão	33
3.4.	Categoria.....	36
3.5.	As pequenas e médias empresas na hotelaria nacional.....	38
3.6.	Síntese conclusiva.....	39
4.	Metodologia da investigação	41
4.1.	Formulação das hipóteses da investigação	42
4.2.	Operacionalização das variáveis	44
4.3.	Definição da população em estudo	46
4.4.	Instrumento de recolha.....	47
4.5.	Caracterização da população em estudo	50
4.6.	Processo de tratamento dos dados	56
4.7.	Síntese conclusiva.....	58
5.	Resultados e discussão	59
5.1.	Caracterização da amostra	60
5.2.	Pressupostos dos modelos de regressão linear múltipla	71
5.2.1	Linearidade	71
5.2.2	Autocorrelação.....	71
5.2.3	Homocedasticidade.....	71
5.2.4	Normalidade	72
5.2.5	Multicolinearidade.....	72
5.3.	Resultados dos modelos de regressão linear múltipla	73

5.3.1	Análise inicial e validação dos pressupostos.....	73
5.3.2	Resultados - Amostra GOPPAR Pré COVID.....	76
5.3.3	Resultados - Amostra GOPPAR durante COVID.....	77
5.3.4	Resultados - Amostra RevPAR pré COVID.....	78
5.3.5	Resultados - Amostra RevPAR durante COVID.....	79
5.3.6	Resultados - Comparativo.....	81
5.4.	Discussão dos resultados.....	83
5.5.	Síntese conclusiva.....	87
6.	Conclusão.....	88
6.1.	Síntese e conclusão.....	88
6.2.	Limitações da investigação.....	89
6.3.	Sugestão para futuras investigações.....	90
	Referências bibliográficas.....	91
	Apêndice A: Notebook_Prep_Questionario_Final.ipynb.....	97
	Apêndice B: Questionário.....	98
	Apêndice C: Notebook_AnaliseUniverso_Final.ipynb.....	101
	Apêndice D: Notebook_TransformacaoDados_Final.ipynb.....	107
	Apêndice E: Notebook_AnaliseAmostra_Final.ipynb.....	113

ÍNDICE DE FIGURAS

Figura 2.1 - Chegadas internacionais 2021	9
Figura 2.2 - Fluxo dos aeroportos nacionais 2021	10
Figura 2.3 - Balança turística portuguesa, viagens e turismo, 2009-2011.....	11
Figura 2.4 - Tipo de estabelecimentos por tipologia 2021	12
Figura 2.5 - Dormidas do mercado internacional em 2021	13
Figura 2.6 - Taxa de ocupação por quarto (%) em 2021	14
Figura 2.7 - Capacidade em camas 2021 (estabelecimentos hoteleiros)	15
Figura 3.1 - Preferências no turismo mundial 2021	29
Figura 3.2 - Dormidas na União Europeia.....	29
Figura 4.1 - Nº de unidades hoteleiras por tipologia	51
Figura 4.2- Distribuição da categoria (estrelas).....	51
Figura 4.3 - Distribuição do nº de UA	52
Figura 4.4 - Unidades hoteleiras por NUTSIII	52
Figura 4.5 - Unidades hoteleiras por distrito	53
Figura 4.6 - <i>Heatmap</i> do número de unidades hoteleiras	54
Figura 4.7- <i>Heatmap</i> do número de unidades de alojamento	55
Figura 5.1 - Distribuição das unidades hoteleiras por tipologia.....	61
Figura 5.2 - Distribuição das unidades hoteleiras por categoria	61
Figura 5.3 - Número de unidades hoteleiras por região (NUTSIII)	62
Figura 5.4 - Número de unidades hoteleiras por distrito	63
Figura 5.5 - <i>Heatmap</i> do número de unidades hoteleiras por distrito	64
Figura 5.6 - <i>Heatmap</i> do número de unidades de alojamento por distrito	65
Figura 5.7 - Histograma do número de UA	66
Figura 5.8 - Distribuição % das variáveis de proximidade costeira	66
Figura 5.9 - Utilização do USALI	67
Figura 5.10 - Dimensão das empresas - PME	67
Figura 5.11 - Quadro comparativo por distrito.....	68
Figura 5.12 - Quadro comparativo por categoria (estrelas).....	69
Figura 5.13 - Distribuição por nível de aglomeração	70
Figura 5.14 - Mapas de dispersão das variáveis independentes (amostra RevPAR).....	73
Figura 5.15 - Mapas de dispersão das variáveis independentes (Amostra GOPPAR)...	73
Figura 5.16 - Resultados GOPPAR pré COVID	76

Figura 5.17 - Resultados GOPPAR durante COVID	78
Figura 5.18 - Resultados RevPAR pré COVID.....	79
Figura 5.19 - Resultados RevPAR durante COVID	80
Figura 5.20 - Quadro sumário das hipóteses	87

ÍNDICE DE TABELAS

Tabela 3.1 - Escala dimensional de uma unidade hoteleira.....	33
Tabela 4.1 - Estatísticas por distrito (Unidades hoteleiras e unidades de alojamento) ..	53
Tabela 5.1 - Estatísticas descritivas dos modelos.....	75
Tabela 5.2 - Médias de performance - costeiras vs não costeiras.....	75
Tabela 5.3 - Médias de performance - aglomeração	76
Tabela 5.4 - Quadro comparativo dos modelos	81

LISTA DE ABREVIATURAS E SIGLAS

- ADR – Average daily rate
- AH&LA – American hotel & lodging association
- AL – Alojamento local
- CBD – Central business district
- COO – Chief operating officer
- COVID-19 – Coronavirus disease 2019
- EBIT – Earnings before interest, tax
- EBITDA – Earnings before interest, tax, depreciation and amortization
- ET2027 – Estratégia turismo 2027
- F&B – Food & beverage
- GOP – Gross operating profit
- GOPPAR – Gross operating profit per available room
- HC3 – Heteroscedasticity consistent 3rd variation
- HFTP – Hospitality financial technology professionals
- INE – Instituto nacional de estatística
- ISO – International standardization organization
- KPI – Key performance indicator
- NUTSIII – Nomenclatura das Unidades Territoriais para fins Estatísticos
- OLS – Ordinary least squares
- PIB – Produto interno bruto
- PME – Pequenas e médias empresas
- PMS – Property management system
- RevPAR – Revenue per available room
- RevPAS – Revenue per available seat
- RJET – Regime jurídico empreendimentos turísticos
- RNET – Registo nacional de empreendimentos turísticos
- TDP – Transformação de dados em painel
- TER – Turismo em espaço rural
- TO – Tour operator

TP – Turismo de Portugal

TRevPAR – Total revenue per available room

UA – Unidades de alojamento

UNWTO – United nations world travel organization

USALI - Uniform system of accounts for the lodging industry

VAB – Valor acrescentado bruto

VABGT - Valor acrescentado bruto gerado pelo turismo

VIF – Variance inflation factor

1. Introdução

1.1. Enquadramento

É hoje inquestionável o potencial turístico que Portugal apresenta, sustentado na sua riqueza histórica, cultural e paisagística (Daniel, 2010; Moreira, 2018; Sousa, 2021). O país tem um clima agradável durante todo o ano, praias de areia branca e águas cristalinas, encantadoras cidades históricas e arquitetura pitoresca, e uma gastronomia e enologia ímpares (Daniel, 2010; Moreira, 2018). A hotelaria portuguesa é reconhecidamente de elevada qualidade e considerada financeiramente acessível (Sousa, 2021).

A crescente procura por parte de investidores estrangeiros, fundos de investimentos e cadeias internacionais, são um importante indicador do momento em que se encontra o setor do turismo (Horwath HTL, 2022) A diversificação da origem geográfica destes investimentos resulta de um notório e meritório esforço desenvolvido, quer pelas entidades públicas como pelas privadas, na promoção do destino Portugal pelo mundo (Moreira, 2018).

A partir do presente século, assiste-se a uma transformação de Portugal enquanto destino turístico, do histórico turismo de sol e praia, ou “*sun, sea and sand*” (3S), para um destino com uma oferta diversificada e rica (Daniel, 2010). A gastronomia, as diferentes paisagens e culturas regionais e a riquíssima história portuguesa, trouxeram uma nova dimensão ao turismo em Portugal, alcançando novos mercados e visitantes (Daniel, 2010).

Para tal, contribuíram o aumento das ligações aéreas e na última década o desenvolvimento das *low cost*¹, particularmente nos destinos urbanos de Lisboa, Porto e nos destinos insulares da Madeira e dos Açores (Costa & Almeida, 2015), que não só colocaram estes no *top of mind*² dos viajantes europeus, como permitiram que os turistas chegassem por esta via a sub-destinos como o Douro Vinhateiro, o Interior Alentejano ou os Açores. Esta notoriedade veio acompanhada pelo crescente número de galardões atribuídos nos *World Travel Awards*³, que ano após ano, reconhecem Portugal e os seus

¹ Termo aplicado às companhias aéreas cujo modelo económico assenta no baixo custo operacional e consequentemente preços altamente competitivos

² Conceito aplicado à investigação do comportamento do cliente e utilizado em marketing. Representa a percentagem de respostas que mencionam um produto ou marca quando apresentadas a um consumidor uma categoria de produtos

³ Um dos mais antigos, reconhecidos e prestigiados programas de premiação da indústria turística mundial.

sub-destinos, como dos melhores do mundo (Moreira, 2018; Sousa, 2021; Travanca *et al.*, 2022).

Partindo da afirmação de Portugal como um destino turístico alavancado no turismo de sol e praia, ao longo das últimas décadas (Daniel, 2010; Travanca *et al.*, 2022), entende-se relevante investigar se a localização de novos investimentos fora da orla costeira, poderão influenciar a sua performance e conseqüentemente o seu sucesso. Melhor dizendo, em que medida a performance de uma unidade é influenciada pela sua localização, particularmente quanto à sua posição relativa face à costa.

1.2. Abordagem teórica

A localização é considerada desde sempre um dos mais importantes fatores de sucesso em hotelaria (Lado-Sestayo *et al.*, 2016, 2020; Song & Ko, 2017; Yang *et al.*, 2014). Provavelmente a mais famosa expressão do mundo hoteleiro, “*location, location, location*”, cuja verdadeira autoria apesar de desconhecida é muitas vezes atribuída a E.M.Statler ou Conrad Hilton, empresários pioneiros da indústria hoteleira das primeiras décadas do século 20, terá sido dada em resposta à seguinte questão: “Quais os três fatores mais importantes para o sucesso dos negócios?” De uma forma simplificada, e independentemente do seu autor, encerra a importância da localização no sucesso presente e futuro de uma unidade hoteleira (Hotel Travel News, 2009).

Depois de tomada, é das poucas escolhas que dificilmente podem ser alteradas, pelo que compreender a extensão dos seus impactos é absolutamente essencial ao sucesso da indústria (Song & Ko, 2017; Yang *et al.*, 2014).

Desde logo no custo associado ao investimento. Por um lado, o custo de aquisição do terreno, tradicionalmente mais elevado em zonas de maior pressão urbanística ou em destinos turísticos consolidados (Lado-Sestayo *et al.*, 2020; Song & Ko, 2017). Por outro lado, os custos de contexto, nomeadamente o licenciamento, regulamentos, burocracia, e o próprio custo da construção, muitas vezes influenciado pela disponibilidade de empresas especializadas, mão de obra e disponibilidade de recursos materiais.

A localização determina igualmente a disponibilidade de recursos humanos especializados e da oferta de fornecedores e prestadores de serviços, todos essenciais à prestação do serviço hoteleiro (Yang *et al.*, 2014). A existência de escolas técnicas, universidades, agências de emprego, e uma elevada concorrência podem contribuir para influenciar a disponibilidade de mão-de-obra na localização (Lado-Sestayo *et al.*, 2016). Noutra perspetiva, a proximidade de serviços adicionais, como restaurantes, lojas, atrações turísticas e outros serviços complementares, pode ser um fator determinante para os hóspedes escolherem um hotel em detrimento de outro (Yang *et al.*, 2014). Quando há uma oferta diversificada de serviços e atrações próximos ao hotel, isso pode aumentar a procura e a probabilidade de um hóspede regressar (Lado-Sestayo *et al.*, 2020; Lado-Sestayo & Vivel-Búa, 2019).

Com efeito, investigar a localização implica compreender a complexidade deste determinante. Similarmente, avaliar a importância da localização na performance implica compreender a performance. Como conceito abrangente e que pode ser analisado sob

diferentes perspectivas, pode-se equacionar qual o mais apropriado e que melhor permite auxiliar futuros investidores. Na indústria hoteleira a performance tem associada, indicadores como o rendimento por quarto disponível, em inglês *revenue per available room* (RevPAR), taxa de ocupação ou preço médio por quarto ocupado, em inglês *average daily rate* (ADR), que serão abordados em detalhe nesta dissertação.

No entanto, estes indicadores incidem particularmente na componente do rendimento, ignorando a componente dos gastos, pelo que se as vendas são o *input* e o resultado o *output*, seria aparentemente defensável a pertinência do último. A capacidade de converter vendas em resultados não é igual entre hotéis, mas a indústria hoteleira continua a privilegiar o RevPAR (Lee *et al.*, 2019; Schwartz *et al.*, 2017) em detrimento por exemplo do *Gross Operating Profit per Available Room*⁴ (GOPPAR), que incorpora por exemplo, a maior ou menor competência do operador hoteleiro na otimização dos recursos.

As indústrias de serviços, como a hoteleira, eram até muito recentemente, altamente focadas nas vendas, melhor expressado pelo desenvolvimento do *yield management*⁵ ou *revenue management*⁶ nas duas décadas, que procuram, aplicando fórmulas matemáticas, maximizar o volume de vendas de um serviço altamente perecível. O desenvolvimento destas funções, contribuíram por sua vez para o desenvolvimento de *software* específico como *Revenue Management Systems*, *Channel Managers*, entre tantos outros, que combinados produziram resultados visíveis e impactantes nos *stakeholders*⁷ (Kimes, 2003). Neste contexto e por estas razões, os indicadores baseados em *room revenue* eram, portanto, o principal *driver* para os *stakeholders*.

Em todas as reuniões, feiras ou eventos do *trade*⁸, os hoteleiros partilham ocupações, RevPAR e ADR entre si, na expectativa de perceberem o mercado e eventuais tendências (Diário imobiliário, 2023; Hilton Worldwide Holdings Inc., 2023; INE, 2023a, 2023b; Publituris, 2023). Com efeito, este movimento está em total contraponto com a evolução dos *cost controllers* (controladores de custo), que apesar de existirem há muito mais tempo, têm tido, até recentemente, pouca evolução tecnológica ou preferência dos *stakeholders*.

⁴ Resultado operacional dividido pelo número de quartos disponíveis num determinado período

⁵ Conjunto de estratégias que visam maximizar o rendimento por quarto disponível, através da utilização ótima da capacidade disponível de uma unidade hoteleira, utilizando modelos matemáticos previsionais.

⁶ Conjunto de estratégias que visam maximizar o rendimento total, através da gestão do preço e capacidade, mas também da segmentação do mercado, otimização da oferta e gestão da distribuição e respetivos canais.

⁷ Indivíduos, grupos ou entidades que possuem interesse ou são afetados pelas atividades, operações ou decisões de uma organização

⁸ Denominação aplicável a todos os profissionais que trabalham no sector hoteleiro

Adicionalmente ao que já foi referido, existe ainda elevada dificuldade em obter informação financeira completa da indústria, um fenómeno que é mundial e que impede a criação de amostras relevantes, e da criação da necessária comparabilidade essencial aos *stakeholders*. No entanto, este fenómeno de proteção dos dados tem vindo a perder força, com os hoteleiros a sobreporem a necessidade de comparabilidade ao receio de facultar dados financeiros da sua operação.

Partindo da complexidade da localização e da operacionalização da performance, considerou-se também a possibilidade de a performance estar relacionada não só com a localização, mas também com as características intrínsecas da unidade, que não podem ser descuradas, como a dimensão, a categoria e o facto de o hotel estar ou não integrado numa PME (Lado-Sestayo *et al.*, 2016, 2020; Yang *et al.*, 2014).

A dimensão de uma unidade condiciona os segmentos de mercado onde atua e o volume que induz às operações. As economias de escala são um fator relevante a considerar. A capacidade de uma unidade de grandes dimensões captar operações ou segmentos específicos, como congressos ou contratos de garantia, não pode ser desprezada.

A categoria, tradicionalmente medida sob a forma de estrelas, poderá impactar de forma diferenciada na performance de uma mesma unidade em localizações distintas. Por outras palavras, uma categoria superior pode ser mais relevante para a rendibilidade numa localização costeira, do que próximo de um aeroporto. Associada à categoria estão normalmente preços de venda mais elevados (INE, 2022), o que em linha com a avaliação da performance via RevPAR, poderá indiciar uma relação.

Considerando a relevância das pequenas e médias empresas (PME) no tecido empresarial português (Serrasqueiro & Nunes, 2008), representativas de aproximadamente 99,7% do universo empresarial, entende-se igualmente pertinente avaliar se esta distribuição encontra paralelo no setor turístico, e se as eventuais relações anteriormente apresentadas, poderão ter um comportamento independente ou não, da dimensão da empresa.

1.3. Objetivos

O presente estudo pretende atingir os seguintes objetivos:

- Intensificar a investigação empírica sobre a rendibilidade operacional dos hotéis em Portugal.
- Mostrar de que modo é que a localização das unidades hoteleiras portuguesas afeta a sua rendibilidade operacional, considerando a proximidade das unidades hoteleiras a zonas costeiras e a aglomeração do destino.
- Verificar o impacto da dimensão sobre a rendibilidade operacional dos hotéis portugueses.
- Analisar a *performance* dos hotéis em Portugal.

1.4. Estrutura do trabalho

O presente estudo foi estruturado em seis capítulos. O primeiro capítulo faz um enquadramento do problema em análise, apresenta a base teórica e os objetivos propostos, e termina com uma breve descrição da estrutura do trabalho.

De seguida, caracteriza-se o setor turístico português, em particular a hotelaria. Descreve-se a oferta hoteleira, a evolução do número de dormidas, proveitos e investimento efetuados. É abordada a distribuição territorial, no que à proximidade à costa respeita. Por fim, é feito um breve enquadramento do ano 2022, que encerra o segundo capítulo.

O terceiro capítulo corresponde à revisão bibliográfica. Abordar-se-á a complexidade da performance, a começar pela importância da normalização, nas duas principais correntes, focadas no rendimento e no resultado, com uma revisão bibliográfica aos trabalhos das últimas décadas. Depois, no âmbito da localização, analisar-se-á a investigação desenvolvida associada à localização na indústria hoteleira, em particular quanto à proximidade costeira, níveis de aglomeração existentes e outros fatores de localização, em particular quanto a eventuais relações entre estes. Seguidamente abordar-se-á a dimensão enquanto determinante de performance hoteleira, identificando os principais estudos e conclusões, bem como a forma como estas relações podem variar em função da localização. Similarmente, apresentar-se-ão as principais conclusões retiradas da revisão bibliográfica, relativamente à categoria, seguidas da perspetiva dimensional de empresa, enquanto PME ou não PME, e suas características. Por fim, far-se-á uma pequena sùmula das conclusões que servirão de base à formulação das hipóteses em estudo.

No quarto capítulo abordar-se-á a metodologia de investigação adotada. Após uma breve introdução, explicitar-se-ão as hipóteses de investigação e a operacionalização das respetivas variáveis. Far-se-á uma descrição do processo de seleção do universo em estudo bem como do instrumento utilizado na recolha dos dados. Após uma caracterização do universo adotado, será descrito em detalhe o processo de tratamento dos dados e finalizado o capítulo com um resumo do mesmo.

Seguir-se-á o quinto capítulo, dedicado à apresentação dos resultados obtidos com recurso aos modelos estatísticos utilizados. Este capítulo iniciar-se-á com uma introdução, seguida pela caracterização da amostra e pela descrição dos pressupostos dos

modelos de regressão linear. Depois serão apresentados os resultados dos modelos de regressão linear múltipla adotados, e far-se-á uma síntese deste capítulo.

Por fim, este estudo será concluído com um último capítulo, reservado às conclusões deste trabalho, incluindo referência a limitações desta investigação e propostas para investigações futuras.

2. Caracterização do setor turístico português

De acordo com o INE (2022) o Valor Acrescentado Bruto Gerado pelo Turismo (VABGT) representou 5,8% do Valor Acrescentado Bruto (VAB) nacional do ano de 2021, valor ainda assim, abaixo do registado em 2019 quando atingiu 8,1% do VAB nacional. Este setor, contribuiu em 2021 direta e indiretamente, com 16,8 mil milhões de euros para o Produto Interno Bruto (PIB) nacional, equivalente a 8,0% do PIB, e contribuiu para a recuperação do PIB num valor ligeiramente acima dos 33% (INE, 2022).

Com efeito, segundo a Organização Mundial do Turismo das Nações Unidas (UNWTO) Portugal encontrava-se em 2021 na 13ª posição mundial de chegadas de turistas internacionais, com cerca de 9,6 milhões de chegadas que representaram cerca de 13% das exportações nacionais, conforme quadro abaixo.

Figura 2.1 - Chegadas internacionais 2021

Country	Arrivals (million)	% share	Receipts (USD bn)	% share	Receipts per arrival (USD)	Tourism as % of Exports
France	48,4	11 %	40,8	6 %	843	5,0 %
Mexico	31,9	7 %	19,8	3 %	620	3,9 %
Spain	31,2	7 %	34,5	5 %	1.108	6,9 %
Türkiye	29,9	7 %	26,6	4 %	890	60,4 %
Italy	26,9	6 %	25,2	4 %	935	3,6 %
United States	22,1	5 %	70,2	11 %	3.177	3,3 %
Denmark	15,7	3 %	4,5	1 %	287	2,1 %
Greece	14,7	3 %	12,4	2 %	845	15,5 %
Utd Arab Emirates	11,5	3 %	34,4	5 %	3.001	6,5 %
Croatia	10,6	2 %	10,8	2 %	1.011	27,5 %
Germany	10,3	2 %	22,2	3 %	2.155	1,1 %
Poland	9,7	2 %	9,1	1 %	934	2,4 %
Portugal	9,6	2 %	11,9	2 %	1.238	13,0 %

Fonte: UNWTO tourism dashboard (2023)

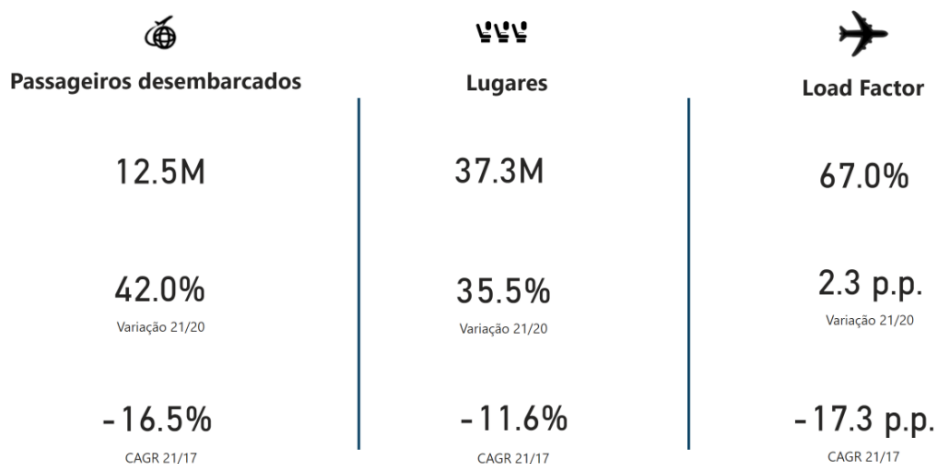
Neste ano, o principal mercado emissor permanece Espanha, com uma quota de 30,2%, seguido do mercado francês com 16,1% e do mercado britânico com 10,6%. É de assinalar o significativo crescimento das chegadas internacionais provenientes dos Estados Unidos da América, com um crescimento 122,2% (INE, 2022).

Contudo, considerando o volume de receitas do turismo em 2021, verificou-se que os primeiros três lugares são ocupados pela França, Reino Unido e Espanha, com 1,957, 1,420 e 1,405 milhões de euros respetivamente (Turismo de Portugal, 2022).

O fluxo de passageiros nos aeroportos nacionais, atingiu perto dos 12,49 milhões de passageiros aerotransportados em 2021, dos quais 9,8 milhões, ou seja 78,4% são passageiros internacionais (Centro de Competências em Turismo Moneris, 2022; Turismo

de Portugal, 2022). O número de lugares disponíveis atingiu as 37,3 milhões, contribuindo para uma relação entre lugares disponíveis e lugares ocupados (*Load Factor*) de 67,0 %.

Figura 2.2 - Fluxo dos aeroportos nacionais 2021

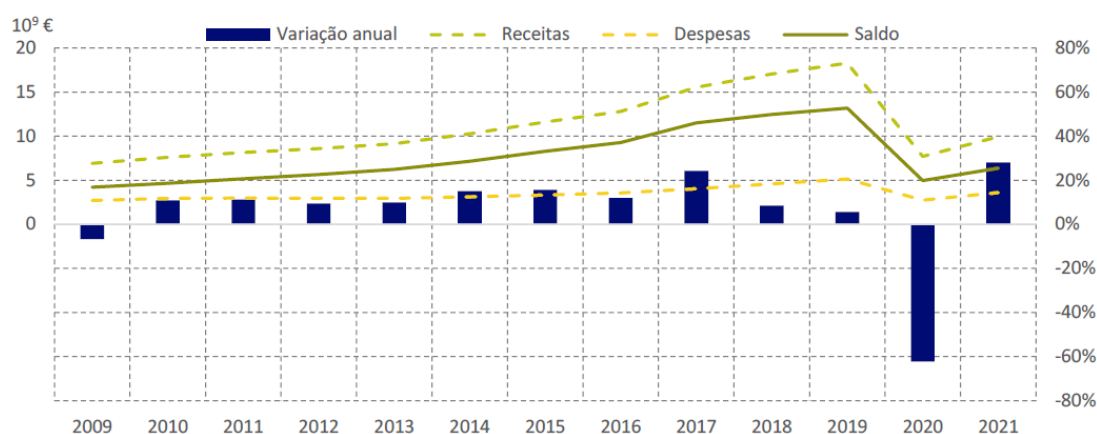


Fonte: Turismo de Portugal (2022)

Globalmente, Portugal atingiu em 2021 a fasquia dos 16,0 milhões de hóspedes e 42,6 milhões de dormidas, valores que se encontram numa trajetória de recuperação face à pandemia, mas ainda assim abaixo de 2019. Destes, o mercado interno representou 52,8% do total, posicionando-se abaixo de 2019 em 13,9%, ainda assim, com uma recuperação mais expressiva quando comparada com o mercado internacional que fechou em 2021 com menos 61,1% de dormidas quando comparado a 2019 (INE, 2022).

Relativamente aos proveitos, *opus citatum*, os proveitos totais atingiram a marca dos 2,3 mil milhões de euros, dos quais 1,8 mil milhões de euros respeitam a alojamento, valores abaixo de 2019 em 45,8% e 45,7% respetivamente. O saldo da balança de pagamentos no setor de Viagens e Turismo apresentou um crescimento de 30,5% face ao período homólogo, atingindo os 6.489 milhões de euros e recuperando face à significativa queda de 2020, mas ainda assim, muito abaixo de 2019 (Turismo de Portugal, 2022).

Figura 2.3 - Balança turística portuguesa, viagens e turismo, 2009-2011



Fonte: INE Estatísticas do turismo 2021 (2023, p.20)

Com efeito, o turismo é uma atividade considerada estratégica para a criação de riqueza, criação de emprego, contributo para a coesão territorial e para o desenvolvimento económico, social e ambiental (Resolução do Conselho de Ministros N° 134/2017 da Presidência do Conselho de Ministros).

É hoje um dos principais motores económicos do país, reconhecido mundialmente pela sua gastronomia, cultura, praias, paisagens (Horwath HTL, 2022) e pelo seu impacto positivo no desenvolvimento das regiões (Centro de Competências em Turismo Moneris, 2022).

O turismo enquanto setor, é absolutamente indissociável da oferta de alojamento turístico, que inclui dois grandes grupos: os empreendimentos turísticos e o alojamento local. Incluído no primeiro, um estabelecimento hoteleiro compreende “os empreendimentos turísticos destinados a proporcionar alojamento temporário e outros serviços acessórios ou de apoio, com ou sem fornecimento de refeições, e vocacionados a uma locação diária” (Decreto-Lei N° 39/2008 de 7 de março do Ministério da Economia e da Inovação).

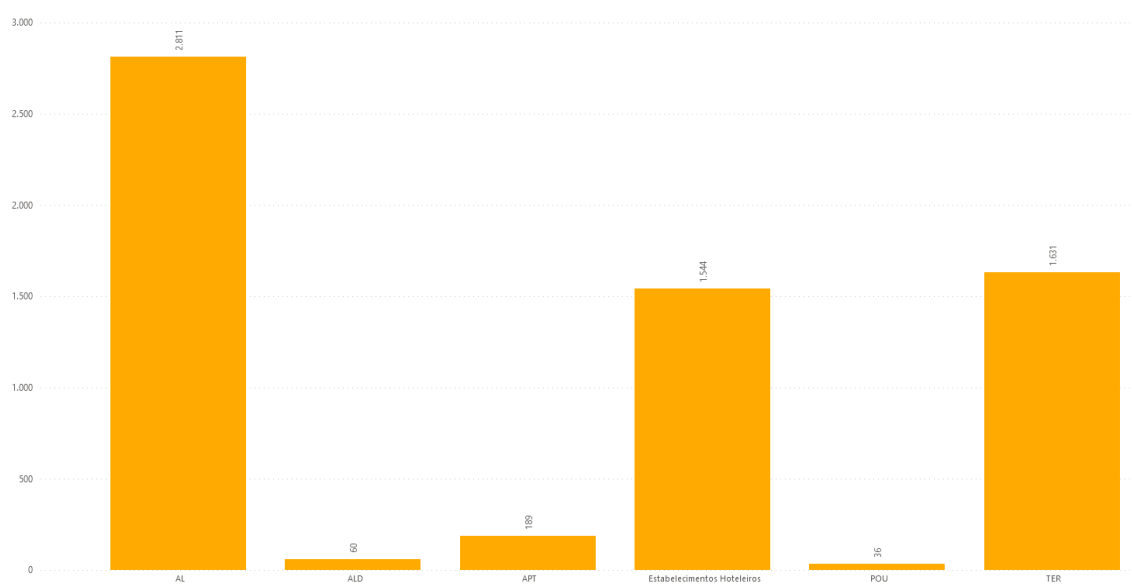
De acordo com este decreto, que define o Regime Jurídico da Instalação, Exploração e Funcionamento dos Empreendimentos Turísticos (RJET), os estabelecimentos hoteleiros podem ser classificados como hotéis, hotéis-apartamentos ou pousadas, devendo ainda possuir uma capacidade mínima de 10 unidades de alojamento.

2.1. Caracterização do alojamento turístico em Portugal

Conforme referido na secção anterior, de acordo com o RJET são considerados estabelecimentos hoteleiros os hotéis, hotéis-apartamentos e pousadas. No entanto, grande parte da estatística disponibilizada pelo Instituto Nacional de Estatística (INE) ou pelo Turismo de Portugal (TP), seguem conceitos adaptados (INE, 2022; Turismo de Portugal, 2022).

De um modo global e de acordo com o Turismo de Portugal (2022), a oferta de alojamento turístico em Portugal totalizava em 2021, 6.271 estabelecimentos turísticos, dos quais 44,8% correspondem a estabelecimentos de alojamento local (AL), 24,6% estabelecimentos hoteleiros e 26% estabelecimentos de Turismo em Espaço Rural (TER). De referir que estes números não incluem parques de campismo, colónias de férias e pousadas de juventude.

Figura 2.4 - Tipo de estabelecimentos por tipologia 2021



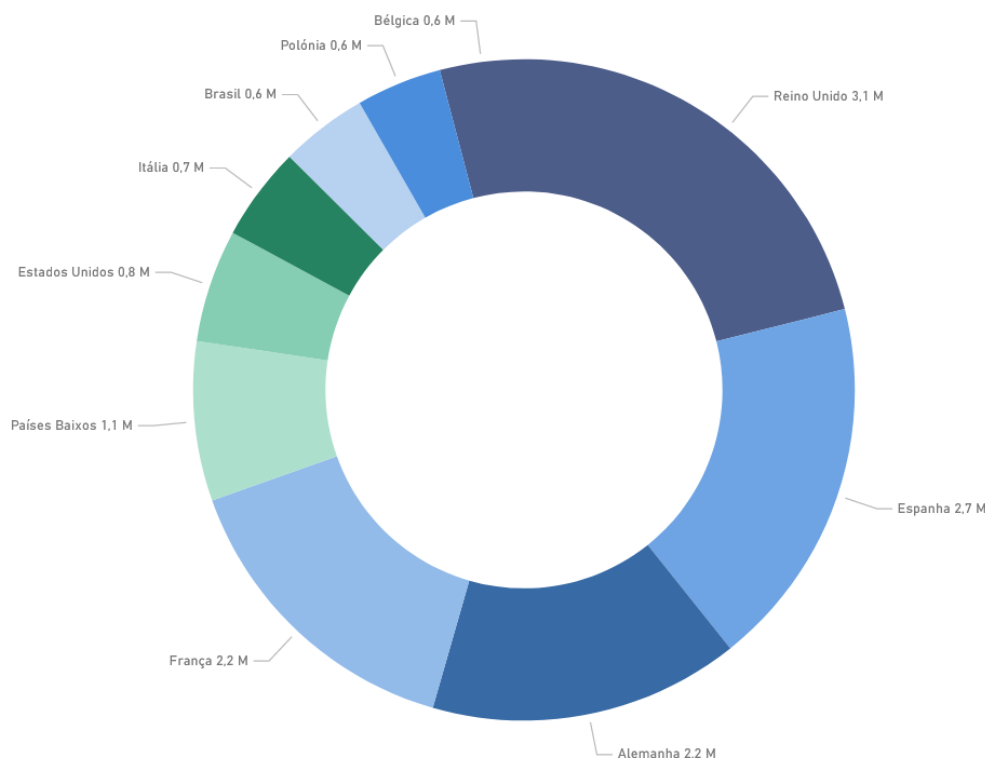
Fonte: Turismo de Portugal (2023)

No que respeita ao número de quartos, estes possuíam um total de 179,5 mil quartos e 404,9 mil camas, valores abaixo da oferta disponibilizada em 2019, em qualquer dos indicadores. (Turismo de Portugal, 2022).

Abordando a procura, o número de dormidas em 2021 atingiu de acordo com o Turismo de Portugal (2022), 37,3 milhões de dormidas, distribuídas equitativamente entre mercado nacional (50,01%) e internacional (49,99%), em linha com 2020, mas ainda longe da relação registada nos anos anteriores à pandemia, com valores na ordem dos

30% e 70% respetivamente. Em recuperação, o turismo internacional continua a ser liderado pelo Reino Unido, seguido de Espanha e da Alemanha.

Figura 2.5 - Dormidas do mercado internacional em 2021



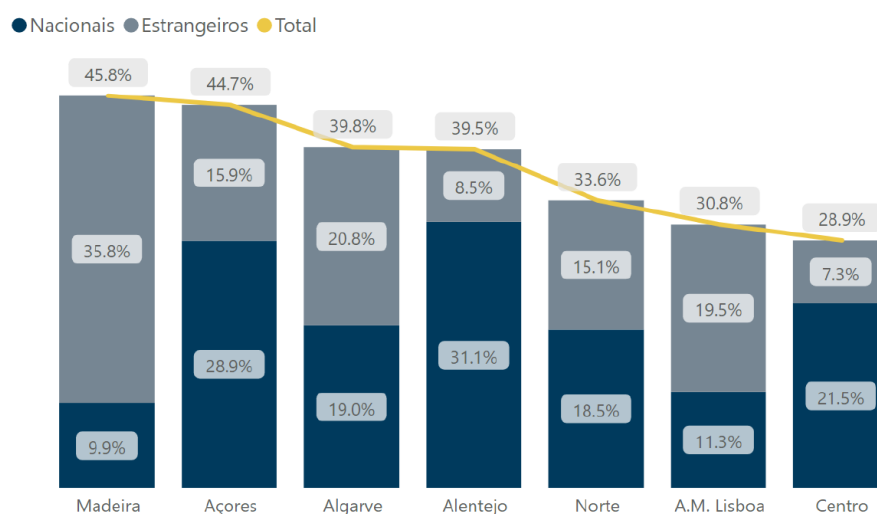
Fonte: Turismo de Portugal (2023)

Esta procura continua a evidenciar um elevado nível de sazonalidade, com julho, agosto e setembro a representarem 47,2% das dormidas nacionais, valores acima dos 36,3% registados em 2019 (Turismo de Portugal, 2022).

Afastando-se deste modo, de um dos objetivos da Estratégia Turismo 2020-2027 (ET2027), nomeadamente, o combate à sazonalidade (Resolução nº 134/2017 de 27 de setembro da Presidência do Conselho de Ministros).

O Algarve continua a região líder em termos de dormidas, destacando-se com um total de 10,9 milhões, seguida da Área Metropolitana de Lisboa e Madeira com 7,7 milhões e 4,4 milhões respetivamente. No entanto, no que respeita à taxa média de ocupação por quarto, as três primeiras posições do ranking pertencem à Madeira, seguida dos Açores e só depois o Algarve, com valores de 45.8%, 44.7% e 39.8% respetivamente conforme figura 2.6.

Figura 2.6 - Taxa de ocupação por quarto (%) em 2021



Fonte: Turismo de Portugal (2022)

Da análise aos proveitos registados em 2021 e de acordo com o Turismo de Portugal (2022), Portugal alcançou 2,3 mil milhões de euros de proveitos globais, dos quais 1,8 mil milhões, ou seja, perto de 75%, corresponderam a proveitos de alojamento. Estes números estão ainda muito abaixo dos 4,2 mil milhões de proveitos globais registados em 2019, mas revelam uma forte recuperação com um crescimento de 61,2%, o que permitiu atingir um RevPAR, na ordem dos 32,55 euros em 2021 e um ADR na ordem 88,24 euros, com este último valor muito próximo do registado em 2019.

Considerando as dinâmicas de mercado, verifica-se que a atividade encontra-se concentrada num pequeno número de empresas gestoras, uma vez que 4,1% das empresas concentraram 50% das dormidas, ao passo que 28,2% das empresas concentravam 90% das dormidas (INE, 2022). Em termos de rendimentos, a mesma fonte revela que 20,1% das empresas concentraram 90% dos proveitos totais gerados pelos estabelecimentos de alojamento turístico em 2021.

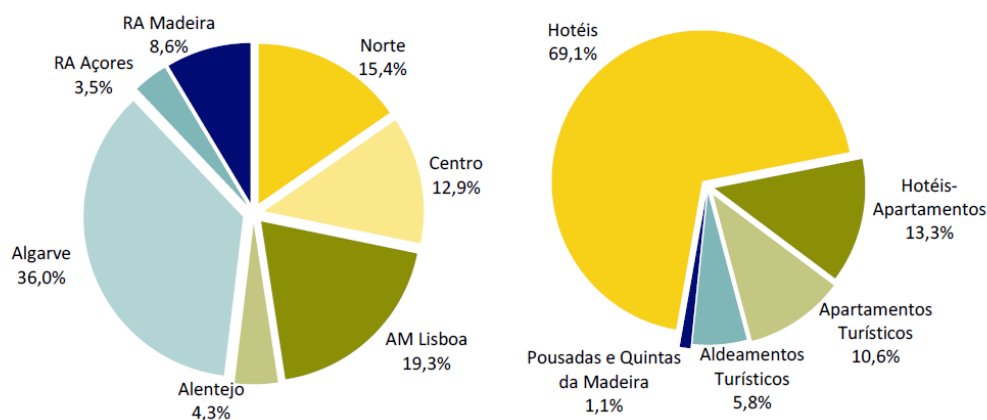
2.2. Caracterização dos estabelecimentos hoteleiros em Portugal

A oferta de estabelecimentos hoteleiros, remete-nos para diferentes abordagens por parte do Turismo de Portugal e do INE. As estatísticas publicadas pelo INE no seu relatório anual (INE, 2022), analisam a hotelaria (ou estabelecimentos hoteleiros), adicionando à definição definida do RJET, os aldeamentos turísticos e as Quintas da Madeira perfazendo de acordo com este relatório, 1.829 estabelecimentos de hotelaria.

2.2.1 Oferta hoteleira

Os estabelecimentos hoteleiros representam apenas 29,3% do total de estabelecimentos de alojamento turístico, mas 75,3% das camas da capacidade global de todos os estabelecimentos de alojamento turístico (INE, 2022). Em 2021 o Norte representava 24,1% de todos os estabelecimentos hoteleiros com o Algarve a representar 21,1%. No entanto, no que respeita a camas, a maioria destas localizavam-se no Algarve (36,0%), e a maioria dos estabelecimentos correspondiam a hotéis (69,1%) (INE, 2022), conforme figura abaixo.

Figura 2.7 - Capacidade em camas 2021 (estabelecimentos hoteleiros)



Fonte: INE Estatísticas do turismo 2021 (2022, p.29)

2.2.2 Dormidas

Em 2021 registaram-se 30,2 milhões de dormidas e 11,4 milhões de hóspedes, sendo que a maioria destas dormidas (72,6%) ocorreram em hotéis e maioritariamente no Algarve (32,7%). No que respeita a mercados, o turista nacional continua a colocar o Algarve no topo das suas preferências, região que atinge um total de 32,3% do total de

dormidas de residentes em linha com a primeira posição no topo das preferências dos visitantes internacionais, onde atinge 33,1% do total das dormidas (INE, 2022). De acordo com a mesma publicação, os mercados do Reino Unido, Espanha, França e Alemanha representaram em 2021 55,2% do total das dormidas internacionais.

A estada média, ou seja, o número de dias que um hóspede pernoita numa determinada unidade hoteleira por estadia, atingiu em 2021, 2,64 noites, sendo este valor mais elevado na região do Algarve (4,13) e na região da Madeira (4,89). Os mercados turísticos do Reino Unido, Irlanda, Alemanha e Suécia destacam-se pelas estadias mais elevadas, na ordem dos 4,58 e 4,20 dias (INE, 2022).

2.2.3 Proveitos

Os proveitos totais em hotelaria atingiram em 2021, 2 mil milhões de euros, sendo que destes 1,5 mil milhões, aproximadamente 75%, correspondem a proveitos de alojamento, permitindo atingir um RevPAR de 35,4 euros e um ADR de 91,8 euros. Das regiões que mais contribuem para estes valores, destaca-se o Algarve com 695 milhões e a área metropolitana de Lisboa com 443 milhões (INE, 2022).

2.2.4 Investimentos

Associada ao crescimento positivo do mercado hoteleiro, a atratividade é objeto de análise no relatório publicado anualmente sobre Portugal pela Horwath HTL (2022), relatório esse que identifica um forte crescimento no número de cadeias internacionais presentes no panorama hoteleiro nacional. Em 2021, Portugal assistiu a 330 milhões de euros em transações hoteleiras, um número recorde que foi largamente ultrapassado em 2022⁹ onde o volume atingiu 900 milhões de euros. Não obstante, os anos recentes têm sido de relativa estabilidade no número de quartos disponíveis, pois de acordo com este mesmo estudo, o número de quartos em hotéis aumentou 5% face a 2021, com as cadeias hoteleiras nacionais a ocupar o topo do ranking por número de quartos. No que respeita ao número de unidades de alojamento (UA), o grupo Pestana Hotel Group lidera com 7.472 UA, seguido do Vila Galé Hotéis com 4.474 UA e em 3º lugar o grupo Accor com 3.956 UA.

⁹ Este facto deve ser considerado extraordinário, pois decorre da venda do projecto Crow a um fundo gerido pela Davidson Kempner, por um valor anunciado próximos dos 650 milhões de euros para os ativos hoteleiros e que englobava um conjunto de ativos geridos pela ECS Capital.

Quanto a investimentos em novos projetos hoteleiros, 2023 apresentava um elevado número de novos hotéis em pipeline¹⁰, que acrescentarão mais de 3.600 UA à capacidade existente, distribuídas por mais de 70 hotéis (Horwath HTL, 2022).

2.2.5 Distribuição territorial – a proximidade costeira

A proximidade costeira é um fator competitivo que contribui positivamente para a rendibilidade (Menicucci, 2018). Em Portugal, a publicação “Estatísticas do Turismo 2021” (INE, 2022) indica que no ano de 2021, 54,9% dos estabelecimentos turísticos localizavam-se em zonas costeiras, conforme definido pelo Eurostat, nomeadamente a) apresentar fronteira marítima e b) 50% da sua área encontrar-se a menos de 10 km do mar. De salientar, de acordo com a mesma publicação, que as áreas costeiras registaram em 2021 crescimentos superiores no número de dormidas que as áreas não costeiras.

No que respeita à hotelaria, a percentagem de estabelecimentos em zona costeira era de 67,3% em 2021, o número de camas representava 81,4%, evidenciando uma dimensão média superior nas zonas costeiras. Em linha com esta assunção, o nº de dormidas, proveitos totais e proveitos de alojamento são particularmente expressivos, com 83,4%, 85,5% e 86,0% respetivamente nesta zona (INE, 2022).

Considerando as várias regiões, em 2021 as dormidas das zonas costeiras da área metropolitana de Lisboa e região do Algarve, representaram 99,2% e 98,4% (INE, 2022).

2.2.6 Enquadramento 2022

Considerando-se que apesar do tema desta dissertação incidir sobre o período de 2017 a 2021, seria importante assinalar a evolução da hotelaria em particular, e do turismo no geral no ano de 2022.

2020 e 2021 ficam marcados pela crise provocada pela COVID-19 (*coronavirus disease 2019*), que assolou a economia nacional, com particular incidência sobre o turismo (Centro de Competências em Turismo Moneris, 2022). A este período, que colocou em dúvida o cumprimento de alguns dos objetivos fundamentais contidos na ET2027 (Resolução do Conselho de Ministros No 134/2017, 2017), veio suceder-se o ano de 2022, em que os proveitos do setor do alojamento turístico assinalaram novo máximo, impulsionado por um crescimento acentuado nos preços (INE, 2023a).

¹⁰ Fluxo futuro de aberturas de hotéis

De acordo com os dados preliminares do INE para a atividade turística (INE, 2023b), o ano de 2022 registou 26,5 milhões de hóspedes e 69,5 milhões de dormidas, número ligeiramente abaixo de 2019, ano referência para o turismo no pré COVID. De acordo com a mesma publicação, e no que respeita à distribuição de dormidas, a distribuição entre mercado externo e interno normalizou, com o mercado externo a subir até aos 67% das dormidas totais, dos quais 19,3% são do reino Unido (INE, 2023a). Neste comparativo, a hotelaria¹¹, que representou 82,2% das dormidas totais em 2022, ficou 1,4% abaixo de 2019, contudo, os rendimentos de alojamento e globais, aumentaram 16,5% e 15,2% respetivamente (Turismo de Portugal, 2023).

De acordo com a mesma publicação, a contração das taxas de ocupação em 2,2% foi significativamente compensada pelo crescimento no ADR, possibilitando um crescimento de 11,5% do RevPAR, o que permitiu que o saldo da balança de pagamento no turismo, tenha aumentado 18,3% face a 2019.

Por fim, importa referir que a evolução da remuneração bruta mensal nominal média do setor da hotelaria e restauração¹², relativa a dezembro de 2022, atingiu os 8,0%, assinalando a maior variação homóloga. Dos setores avaliados e comparando o mesmo período, o referido indicador para o total de todos os setores de atividade, atingiu os 4,2% (INE, 2023c).

¹¹ Hotéis, Hotéis-Apartamentos, apartamentos turísticos, aldeamentos turísticos, pousadas e quintas da Madeira

¹² “Alojamento, restauração e similares”

2.3. Síntese conclusiva

Neste capítulo caracterizou-se o setor turístico português, com maior incidência no ano de 2021, último em estudo. Estabeleceu-se comparativos com o ano de 2019, considerado o último ano de referência antes do COVID.

Quantificou-se a importância do turismo para a economia nacional, através do seu contributo direto e indireto, e apresentaram-se dados sobre os mercados emissores e sobre as chegadas de turistas a território nacional, por forma a melhor compreender a procura.

Seguidamente, abordou-se a oferta de alojamento, em particular a distribuição por tipologia, o contributo em número de camas e o volume de dormidas que induzem. Abordou-se a sazonalidade, a concentração geográfica e salientou-se a elevada concentração dos rendimentos num reduzido número de empresas gestoras.

Depois, fez-se uma breve descrição da oferta por geografias e detalhou-se as dormidas e os proveitos da hotelaria. Abordou-se a distribuição por proximidade costeira e por fim apresentou-se dados preliminares de 2022 para evidenciar a tendência positiva.

Em suma, confirmou-se uma tendência de recuperação do setor, mas, em muitos indicadores, longe dos dados de 2019. Como referido anteriormente, os dados preliminares de 2022 indicam uma retoma da normalidade, especialmente no equilíbrio entre o turismo nacional e internacional, e um crescimento nos rendimentos do setor. O investimento em novos projetos hoteleiros é também um indicador positivo para os anos futuros, com um número considerável de novos estabelecimentos em desenvolvimento previstos para 2023.

3. Revisão da literatura

3.1. Performance

A performance em hotelaria é um conceito abrangente e multidimensional (Sainaghi, 2010), mas de extrema importância na monitorização e execução da estratégia do negócio (Lado-Sestayo *et al.*, 2017; Pnevmatikoudi & Stavrinoudis, 2016). Existem múltiplos indicadores de performance ou *key performance indicators* (KPI), demonstrativos da complexidade e das características multidimensionais da indústria hoteleira (Pnevmatikoudi & Stavrinoudis, 2016).

Nos últimos 30 anos, a avaliação da performance tem evoluído de um sistema assente em indicadores financeiros para um conjunto de indicadores financeiros e não financeiros, como a satisfação do cliente e a qualidade, o que permitiu agrupar os indicadores em dois grandes grupos: financeiros e estratégicos (Bresciani *et al.*, 2015). Sainaghi (2010), por sua vez, denominou estes dois grandes grupos por financeiros e operacionais.

Numa perspetiva financeira, a performance em hotelaria pode ser medida na dimensão relativa à unidade utilizando o GOPPAR ou o *Total Revenue Per Available Room*¹³ (TRevPAR) ou na dimensão relativa a qualquer dos departamentos que a compõem, utilizando o *departmental profit* ou o *gross departmental profit*. Pode-se avaliar o *Revenue Per Available Seat* (RevPAS) e obter um indicador de performance da operação de comidas e bebidas, ou *food & beverage* (F&B), da mesma forma que se pode avaliar a taxa de ocupação e obter um indicador classificado como relativo à performance de vendas, ainda que limitativo.

Existem também indicadores não financeiros, como o nível de satisfação dos hóspedes, reputação da unidade, taxas de rotação do pessoal, satisfação do pessoal, entre outros (Pnevmatikoudi & Stavrinoudis, 2016; Santos *et al.*, 2016; Singh & Schmidgall, 2002).

Esta multiplicidade de perspetivas tem contribuído para que a performance hoteleira tenha sido objeto de inúmeros estudos académicos e investigações (Assaf & Tsionas, 2018; Bresciani *et al.*, 2015; Harris & Mongiello, 2001; Kala & Bagri, 2014; Menicucci, 2018; Sainaghi, 2010; Sainaghi *et al.*, 2018, 2020; Singh & Schmidgall, 2002;

¹³ Rendimentos totais divididos pelo número de quartos disponíveis num determinado período

Tan *et al.*, 2022). No entanto, esta tendência, salvo algumas exceções, é relativamente recente e em rápido desenvolvimento. (Sainaghi *et al.*, 2018).

Este interesse não é exclusivo dos acadêmicos (Sainaghi *et al.*, 2020). A própria indústria, alavancada no aparecimento de consultoras especializadas no setor turístico, na disseminação mundial das cadeias internacionais e respetivos padrões de *reporting* baseados em *Uniform System of Accounts for the Lodging Industry* (USALI) e na criação de empresas especializadas em *benchmarking* hoteleiro, foram obrigadas a refletir profundamente sobre como medir e comparar, a performance hoteleira (Hotstats, 2021b, 2021a, 2022; Pnevmatikoudi & Stavrinoudis, 2016; Singh & Schmidgall, 2002). Segundo estes autores, esta informação pode ser usada para identificar boas práticas e áreas de oportunidade, permitindo aos diretores hoteleiros decisões informadas sobre como permanecer competitivos.

O acompanhamento da performance hoteleira é da maior importância para um gestor hoteleiro (Schwartz *et al.*, 2017; Singh & Schmidgall, 2002). Enquanto negócio de capital intensivo (Tan *et al.*, 2022), em que os serviços vendidos são altamente perecíveis, a importância da medição de desempenho na indústria hoteleira não pode ser menosprezada. Ponderando a efemeridade do quarto, representado no famoso dito hoteleiro, em que “um quarto não vendido é um quarto perdido”¹⁴ Lima Santos *et al.* (2016), defendem que o negócio hoteleiro é considerado um negócio de volume.

Este modelo de negócio, depende de um elevado volume de transações, neste caso de bens e serviços perecíveis, a preços relativamente baixos quando comparados a outras indústrias, para gerar um volume de receitas significativo.

A avaliação do desempenho permite monitorizar o progresso em séries temporais, garantindo que os hotéis estão no caminho certo para alcançar seus objetivos e fazer ajustes necessários sempre que exigível (Schwartz *et al.*, 2017; Singh & Schmidgall, 2002).

A medição da performance também permite a identificação de oportunidades de melhoria. Através da análise de dados de desempenho, os diretores de hotéis podem identificar áreas onde o negócio pode ser tornado mais eficiente e efetivo, resultando em maior performance, seja na perspectiva financeira, operacional, reputacional ou na satisfação do cliente (Kala & Bagri, 2014; Singh & Schmidgall, 2002). Outra vantagem

¹⁴ Este adágio hoteleiro expressa a oportunidade perdida de venda de um quarto. Uma unidade hoteleira apresenta tantas oportunidades de venda quantos dias do ano aberta vezes o número de quartos. Quando num determinado dia um quarto não é vendido, a receita inerente está perdida para sempre e não poderá ser recuperada.

da medição da performance é que esta fornece aos diretores hoteleiros os dados necessários a uma tomada de decisão sustentada. As percepções e tendências reveladas através dos dados de desempenho podem suportar futuras estratégias e planos, ajudando a garantir o sucesso contínuo do negócio (Singh & Schmidgall, 2002).

Resumindo, este é um tema que atrai simultaneamente os práticos e os teóricos (Sainaghi, 2010) e cujas conclusões não são unânimes, conforme abordar-se-á nas próximas subsecções.

3.1.1 A importância do USALI na normalização

O USALI representa um conjunto de normas de registo contabilístico e financeiro, desenvolvido pela *American Hotel & Lodging Association* (AH&LA) e pela *Hospitality Financial and Technology Professionals* (HFTP) para o setor da hotelaria. O seu objetivo principal é a normalização da informação financeira e de gestão (Santos *et al.*, 2016). A sua primeira versão foi publicada em 1926, sendo atualizada regularmente para refletir as tendências e mudanças no setor hoteleiro (Santos *et al.*, 2016). Atualmente está em vigor a 11ª versão, publicada em 2011, estando previsto o lançamento da 12ª versão no outono de 2023 (HFTP, 2023), isto depois de vários adiamentos, em consequência de disputas relativas aos direitos de autor, desde 2018 detidos pela HFTP.

A adoção do USALI pelas grandes cadeias hoteleiras¹⁵, em cujos contratos de gestão a sua utilização está definida como obrigatória, tem suportado o interesse e uma utilização crescente por parte dos hoteleiros (Santos *et al.*, 2016). Pela mesma razão, a expansão destas marcas a múltiplas regiões do globo, facilitou a sua aceitação junto de investidores e financiadores, sendo atualmente de utilização obrigatória para empresas de *benchmark* como a HotStats ou a STR (Faria *et al.*, 2018).

No entanto, esta associação às grandes cadeias, consultoras ou investidores não coloca em causa a sua aplicabilidade a hotéis independentes e de pequena dimensão. Aliás, a sua flexibilidade é outra razão do seu sucesso pois permite adaptar-se a qualquer tipo de hotel, independentemente da dimensão, segmento ou classificação (Correia *et al.*, 2016). A sua adoção é, pois, de acordo com estes autores, altamente recomendada para hotéis que pretendam monitorizar a sua performance e compará-la com os seus concorrentes.

¹⁵ Hyatt Hotels Corp., Marriot International Inc., Viceroy Hotel Group, Hilton Hotels, Rosewood Hotel Management, IHG, são algumas das cadeias que utilizam o USALI

Em Portugal, a adoção do USALI tem vindo a crescer, muito por força da entrada de cadeias hoteleiras internacionais por via de contratos de gestão ou de franchising que assim o exigem. No entanto, a sua integração é ainda limitada, conforme confirmou o trabalho de Faria *et al.* (2018), onde de uma amostra de 66 unidades hoteleiras do Algarve, verificou que apenas cerca de metade adota o USALI.

Em suma, a importância da implementação do USALI é potenciada pela necessidade de normalização dos indicadores hoteleiros, válida para a presente, passadas e futuras investigações.

3.1.2 Receitas de alojamento vs. receitas totais vs. resultado operacional

Apesar do forte desenvolvimento que a indústria do turismo assistiu nas últimas décadas, é quase tradição em Portugal, ano após ano, anunciar resultados sob diferentes dimensões (região, marca hoteleira, cadeia hoteleira, observatório) em dois, mais recentemente, três KPI: Taxa de ocupação, ADR e RevPAR (Centro de Competências em Turismo Moneris, 2022; Hotstats, 2021b, 2022; INE, 2022, 2023b, 2023a; Turismo de Portugal, 2022, 2023).

A taxa de ocupação, o ADR e o RevPAR, são há largas décadas utilizados tradicionalmente para avaliar a performance dos hotéis (Lee *et al.*, 2019). Não é por isso invulgar que constem, a par de outros indicadores, como as preferências dos diretores hoteleiros (Kala & Bagri, 2014) e dos executivos financeiros (Singh & Schmidgall, 2002), bem como dos académicos que estudam a performance deste setor (Assaf & Tsionas, 2018; Kim *et al.*, 2013).

As razões para esta preferência são várias, destacando-se a facilidade de cálculo e a sua elevada comparabilidade, o que os qualifica como indicadores concorrenciais (Sainaghi *et al.*, 2020; Schwartz *et al.*, 2017). Complementarmente, a receita gerada no aluguer de quartos, isto é, o *Room Revenue*, representa na grande maioria das unidades, nunca menos de 65% da receita gerada (Hotstats, 2019; Turismo de Portugal, 2022, 2023), funcionando ainda como catalisador de todas as restantes categorias de rendimentos, sendo por isso também um KPI representativo de respetiva performance.

Por todas estas razões, analistas e hoteleiros continuam, até aos dias de hoje, a apresentar o RevPAR (Lee *et al.*, 2019) nas apresentações de resultados do setor ou das próprias cadeias ou marcas.

Não obstante, alguns autores (Dimitrić *et al.*, 2019; Lado-Sestayo *et al.*, 2016, 2020; Lado-Sestayo & Vivel-Búa, 2018a, 2018b; Lee, 2009; Serrasqueiro & Nunes,

2008), apontando a dificuldade em obter informação financeira operacional das unidades e a inexistência de uma base de dados pública com estes dados, têm optado por indicadores associados ao *Earnings Before Interest and Taxes* (EBIT), *Earnings Before Interest, Taxes, Depreciation and Amortization* (EBITDA), entre outros, tido por financeiros. Schwartz *et al.* (2017), defendem que esta opção também se deve à relevância destes indicadores para investidores e acionistas.

Por sua vez e considerando as limitações do RevPAR, existe uma crescente discussão sobre a utilização de indicadores baseados apenas na receita de alojamento versus indicadores que incorporem outros fluxos de receitas, como o F&B, spas, salas de conferência, casinos, parques de estacionamento entre outros (Kim *et al.*, 2013; Lee *et al.*, 2019; Sainaghi *et al.*, 2020).

Nesta temática, alguns investigadores incorporaram o impacto do F&B (Kim *et al.*, 2013), enquanto fonte de receita relevante, particularmente em hotéis *full service*, ou seja, que incorporam outros serviços aos clientes para além do quarto, nomeadamente SPA, restaurante, bar, centro de congressos, sala de reuniões, salões de banquetes, parque de estacionamento, ou até mesmo casino. Nesse sentido, a globalidade das receitas pode ser avaliada pelo TRevPAR (*Total Revenue per Available Room*), indicador que engloba toda a receita por quarto disponível, e cuja relevância é tanto maior quanto a dimensão ou a multiplicidade de serviços que a unidade incorpora e disponibiliza (Kim *et al.*, 2013; Lado-Sestayo *et al.*, 2017).

O facto de o TRevPAR e o RevPAR não considerarem gastos, tem originado debate entre investigadores (Schwartz *et al.*, 2017).

Este debate também tem sido objeto de destaque em alguns artigos recentes de algumas das mais relevantes consultoras internacionais em matéria de benchmarking. No *International Hospitality Forum*, ocorrido em maio de 2022 em Berlim, o debate sobre a melhor métrica voltou à discussão, com o COO da Hotstats, Michael Groove, a defender a utilização do GOPPAR pelas insuficiências já apontadas do RevPAR (Hotstats, 2021b).

Recentemente, o relatório anual de 2021 sobre a performance hoteleira da Hotstats (2022) veio colocar a rendibilidade como o grande tema para os próximos anos. Também da autoria da Hotstats (2021a) uma publicação do seu blog fazia uma curiosa analogia a duas modalidades do atletismo, a maratona e os 100 metros. Nesta analogia, em que o GOPPAR representava o primeiro e o RevPAR o segundo, o autor defendia que vender é naturalmente positivo e um bom indicador de performance, mas o que realmente conta é a capacidade de converter essas vendas em resultado, dia após dia, mês após mês,

ano após ano, de forma consistente e duradoura. E durante esse processo, ou “maratona”, manter sob monitorização os concorrentes, as tendências, o ambiente envolvente e demais fatores em consideração.

3.1.3 GOPPAR vs RevPAR

O RevPAR e o GOPPAR têm sido utilizados de forma consistente tanto pela indústria hoteleira como pelos investigadores para relatar ou estudar a performance de uma unidade hoteleira (Bresciani *et al.*, 2015; Claver-Cortés *et al.*, 2007; Kim *et al.*, 2013; Lee *et al.*, 2019; Schwartz *et al.*, 2017; Singh & Schmidgall, 2002; Tan *et al.*, 2022).

O RevPAR combina a taxa de ocupação com o ADR pelo que compensa as limitações individuais de ambos. No entanto, o não cumprimento estrito do USALI, pode conduzir a manipulação ou distorção deste KPI, em particular em unidades em que os quartos disponíveis podem ser ajustados – como em unidades sazonais – ou em que as componentes extra alojamento podem alavancar o ADR, como casinos ou centros de conferências (Schwartz *et al.*, 2017).

Nesse sentido, o GOPPAR tem ganho preponderância nos últimos anos, sendo a sua principal vantagem face a indicadores baseados em proveitos, a integração da componente de custo, o que permite uma completa avaliação de eficiência das operações (Lee *et al.*, 2019; Schwartz *et al.*, 2017).

O GOPPAR integra a globalidade das receitas, nomeadamente, alojamento, F&B, Spa, congressos, reuniões, entre outras, e a globalidade dos custos, sejam diretos como de alojamento, F&B ou Spa, ou indiretos, como de energias, manutenção, administração ou vendas e marketing, o que permite reportar os lucros operacionais (Santos *et al.*, 2016).

Considerando a utilização massificada do RevPAR, Schwartz *et al.*, (2017) analisaram uma eventual correlação entre o RevPAR e o GOPPAR, que poderia, por simplicidade e disponibilidade, privilegiar a utilização suportada cientificamente do primeiro em detrimento do último. Para o efeito investigaram dados financeiros de mais de 1000 hotéis localizados nos Estados Unidos ao longo dos anos de 2006 a 2011. Nos seus resultados, os autores concluíram de um modo geral, pela existência de fortes correlações entre o RevPAR e o GOPPAR, o que suportaria a utilização do RevPAR como indexante do GOPPAR permitindo a sua utilização para efeitos preditivos. Porém, o seu comportamento não era linear, ou seja, em unidades em que o volume de vendas extra alojamento, ou seja, F&B, alugueres de salas, conferências, SPA, fosse elevado, o

GOPPAR apresentava-se como indicador mais apropriado, pois apesar de existente, a relação entre ambos era visivelmente mais fraca.

Outra conclusão apontada por este estudo, refere uma tendência de decréscimo, ainda que ligeira, desta correlação.

Mais recentemente Lee *et al.*, (2019), numa análise ao comportamento destes dois indicadores na perspetiva da unidade e da empresa, apontaram resultados mistos ao nível da unidade, mas concluíram pela existência de uma relação superior entre a empresa e o RevPAR comparativamente à empresa e o GOPPAR, particularmente quanto às características intrínsecas da unidade, como a sua localização, dimensão ou categoria.

Lado-Sestayo *et al.*, (2017), de modo a considerar o total das receitas e evitar distorções decorrentes da incorporação de gastos, utilizaram o TREVPAR como indicador para analisar a performance hoteleira.

Na impossibilidade de obter estes dados financeiros operacionais, alguns investigadores optam por aproximações ao GOP considerando os dados disponíveis (Claver-Cortés *et al.*, 2007).

Em suma, considerando que a investigação académica sobre o indicador que melhor traduz a rendibilidade e performance hoteleira é significativamente reduzida (Lee *et al.*, 2019), não é possível uma conclusão clara sobre qual o indicador mais apropriado.

3.2. Localização

A escolha do local adequado para a implementação de uma unidade hoteleira pode ter um impacto significativo no número de reservas e na receita gerada, considerando que uma localização inadequada pode levar a uma baixa ocupação e a uma menor rentabilidade (Lado-Sestayo *et al.*, 2016).

A localização é um dos fatores mais importantes e amplamente estudado na investigação dos determinantes de performance de uma unidade hoteleira, bem como a sua eficiência (Assaf & Tsionas, 2018; Lado-Sestayo *et al.*, 2016, 2017, 2020; Lado-Sestayo & Vivel-Búa, 2018b, 2018a, 2019; Lee *et al.*, 2019; Menicucci, 2018; Oliveira *et al.*, 2013; Romero *et al.*, 2020; Song & Ko, 2017).

Uma das características da variável localização é a sua relatividade. Pode-se avaliar em diferentes escalas, como sejam um país, uma região, um destino turístico, uma cidade ou até mesmo uma zona (Assaf & Tsionas, 2018; Barros *et al.*, 2011; Lado-Sestayo *et al.*, 2016, 2020; Lado-Sestayo & Vivel-Búa, 2018a; Menicucci, 2018; Oliveira *et al.*, 2013; Romero *et al.*, 2020; Tan *et al.*, 2022).

Além disso, uma mesma localização pode ser influenciada por um conjunto de variáveis secundárias, nem todas com a mesma preponderância. Neste âmbito, Lado-Sestayo *et al.* (2020), concluíram que cerca de 50% da rendibilidade poderia ser explicada por variáveis secundárias que influenciavam a determinante principal. Por essa razão, a investigação do impacto da localização deve ter em consideração não só as características da unidade, como as características do próprio destino em análise (Lado-Sestayo *et al.*, 2016; Lado-Sestayo & Vivel-Búa, 2018a).

Avaliar a proximidade de uma unidade hoteleira a um aeroporto, pode ter uma relevância diferente quando comparado um destino citadino, a um destino de sol e praia. Neste último, a proximidade à costa poderá ser um critério mais relevante, considerando as motivações deste tipo de turismo (Lado-Sestayo *et al.*, 2020; Lado-Sestayo & Vivel-Búa, 2018a). Nesta linha, Lado-Sestayo *et al.* (2017), avaliaram a performance procurando conjugar três diferentes abordagens: o modelo geográfico, o modelo focado na competição e o modelo focado na aglomeração. O primeiro centrado na proximidade ao *Central Business District* (CBD), ou zona/centro de negócios enquanto determinante, o segundo atendendo ao nível de competição e fatores diferenciadores entre hotéis no

destino e, finalmente, o terceiro avaliando a aglomeração¹⁶ enquanto produto de externalidades do destino.

Apesar do exposto anteriormente, a localização geográfica, quando analisada de um modo generalizado, não tem que ser impeditiva. Antes pelo contrário. No seu estudo sobre a eficiência das pousadas de Portugal, Barros (2005) concluiu que as Pousadas situadas em cidades ou próximas destas eram mais eficientes do que as situadas em localizações remotas, que podia recorrer, da elevada procura inerente aos destinos urbanos.

Em linha com este estudo estão as conclusões de Assaf e Tsionas (2018) que ao analisarem a eficiência, concluíram que alguns tipos de hotéis, em função da sua localização, beneficiavam de uma maior eficiência, nomeadamente os urbanos e suburbanos. Este estudo concluiu ainda que os *resorts*, normalmente associados a destinos turísticos de sol e praia, indicavam de igual modo, níveis de eficiência mais elevados.

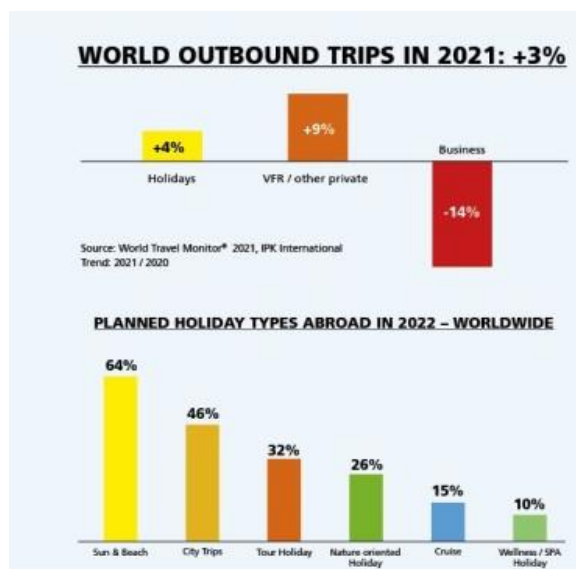
À escala regional, Oliveira *et al.* (2013), num estudo desenvolvido relativamente a unidades hoteleiras no localizadas no Algarve, identificaram diferenças na eficiência de unidades hoteleiras localizadas no Barlavento, quando comparadas com as unidades hoteleiras localizadas no Sotavento.

3.2.1 Proximidade às zonas costeiras

Um dos grandes impulsionadores do turismo mundial continua a ser o segmento sol e praias, conhecido pelo *Sun, Sea and Sand* (3 “S”), conforme reconfirmou o World Travel Monitor® de 2021 (figura 3.1) (ITB Berlin & IPK International, 2022). Por esta razão, a proximidade a zonas costeiras pode ser um fator de localização preponderante no estudo da performance.

¹⁶ Concentração de unidades hoteleiras numa determinada zona geográfica

Figura 3.1 - Preferências no turismo mundial 2021



Fonte: (ITB Berlin & IPK International, 2022)

Por esta razão, a sazonalidade é e continuará a ser uma característica intrínseca ao turismo (Daniel, 2010; Duro & Turrión-Prats, 2019; Marques, 2010). Tomando por exemplo o referido Algarve, janeiro nunca terá níveis de faturação de agosto. É uma limitação própria do clima, que conjugada com as férias escolares, origina os tradicionais fluxos turísticos sazonais no nosso país (Duro & Turrión-Prats, 2019; Travanca *et al.*, 2022).

Mas não é um fenómeno único do nosso país. Espanha, Turquia, Grécia e outros destinos mediterrâneos têm o mesmo problema (Duro & Turrión-Prats, 2019). A grande maioria dos destinos turístico possuem ciclos e épocas altas e baixas conforme figura abaixo:

Figura 3.2 - Dormidas na União Europeia



Fonte: Eurostat (2022)

Considerando a importância que o segmento dos 3S, tem para o turismo, o impacto da localização, na sua perspectiva de proximidade à costa, tem sido motivo de investigação. Tal é particularmente expressivo nos países do mediterrâneo. Lado-Sestayo *et al.* (2016), ao investigarem a rentabilidade dos hotéis localizados em Espanha, concluíram que os níveis de procura eram substancialmente mais elevados nas zonas costeiras, em parte justificado pela preponderância do segmento Sol e Praia no turismo espanhol.

Este dado foi reforçado por Romero *et al.* (2020), que observaram diferenças significativas de performance entre as PME hoteleiras de Espanha, país com longa tradição no segmento de sol e praia, decorrente da vasta costa que possui. Neste estudo, que visou analisar o contributo dos operadores turísticos (TO) para a performance de PME hoteleiras nas zonas costeiras e no interior de Espanha, concluiu que o impacto dos TO na performance era significativamente negativo nas zonas costeiras quando comparado com as zonas de interior. Apesar de não referido, considera-se que a atividade decorrente dos TO procura volume e escala.

Esta escala permite rentabilidade e a falta dela implica prejuízos. Quando um TO estuda operações charters para um destino, com ou sem frota aérea própria, tem que estudar a potencial rentabilidade desse destino. Porque montar uma operação para o destino Espanha com X aviões e para a Turquia ou Portugal com Z aviões? Destas avaliações, mais complexas do que acabámos de referir, depende a escala, porque os aviões fretados requerem um elevado *load factor* para alcançar rentabilidade (Mihailescu, 2010). O segmento de sol e praia é particularmente propício para garantir um elevado volume num espaço temporal tradicionalmente mais reduzido (Mihailescu, 2010).

Assinala-se um ponto que deve ser destacado destas conclusões. Se o impacto indicado anteriormente é relevante no volume de vendas das unidades hoteleiras localizadas nas zonas costeiras, a sua rentabilidade, pelo contrário é prejudicada, fruto da pressão nas tarifas que o TO tende a exercer e que as PME não conseguem combater (Romero *et al.*, 2020). Assim, reforça-se o que já foi dito na secção anterior, ou seja, a mais apropriada definição da medida de performance, não é indiferente ao objeto de estudo.

Num estudo desenvolvido em Itália, Menicucci (2018) concluiu no mesmo sentido, ao apontar evidências de que a localização, em particular a proximidade a zonas costeiras tem um significativo impacto positivo sobre a rentabilidade. A autora defende

que este impacto pode justificar-se pelo maior investimento público e privado na promoção e no desenvolvimento em infraestruturas de que estas zonas beneficiam.

3.2.2 Aglomeração do mercado

A elevada concorrência ou a ausência desta por via do nível de aglomeração do mercado, é outro fator que pode determinar significativamente o impacto da localização na performance de uma unidade hoteleira. A estrutura do mercado é, pois, uma característica cuja importância não pode ser negligenciada (Lado-Sestayo *et al.*, 2016). É essencial entender quais são os hotéis concorrentes na zona ou no destino, bem como a respetiva qualidade e preços. Uma localização que é altamente competitiva pode tornar mais difícil a captação de clientes, a menos que o hotel ofereça algo único que o diferencie dos concorrentes. Ainda segundo Lado-Sestayo *et al.*, (2016), a aglomeração do mercado pode ser provocada por externalidades específicas do destino ou pelo impacto decorrente de concorrentes de dimensão superior. Uma elevada aglomeração traduzir-se-á num número elevado de unidades hoteleiras num determinado espaço geográfico, como seja um destino, zona ou região.

Neste âmbito, Travanca *et al.* (2022) concluíram existir uma relação positiva significativa entre a rentabilidade de hotéis integrados em empresas PME e a sua localização no Algarve, apesar da elevada concorrência e aglomeração. As elevadas taxas de ocupação anuais registadas no Algarve, foram uma das justificações apontadas, sendo que se entende, que a procura em destinos turísticos consolidados, com boas ligações aéreas, como é o caso do Algarve, poderá contribuir para esta conclusão.

Porém, apesar do destino ser geograficamente pequeno, estes resultados poderão não ser iguais em todos os hotéis no Algarve. A própria localização dentro de um destino turístico, como uma unidade hoteleira na primeira linha da praia, também por ser um fator diferenciador como se viu anteriormente.

Lado-Sestayo *et al.* (2016), apontaram algumas conclusões interessantes a este propósito. Por um lado, identificaram uma forte relação entre os níveis de concentração do mercado e a performance, *id est*, quanto maior a concentração da quota de mercado, melhor a performance. No entanto, esta maior concentração de mercado é tradicionalmente encontrada em destino turísticos pouco desenvolvidos, com poucos fornecedores de camas (hotéis) e conseqüentemente com reduzidos níveis de aglomeração. Por outro lado, identificaram uma forte relação entre a procura e a

performance, em particular nos destinos turísticos onde a concentração é mais reduzida por força da elevada aglomeração.

Estas conclusões, aparentemente contraditórias, foram objeto de nova investigação por dois dos anteriores autores (Lado-Sestayo & Vivel-Búa, 2018a) e novamente foram suportadas no que respeita à forte relação entre a concentração do mercado e a rendibilidade, e quanto à relação positiva entre a procura e a rendibilidade.

Efetivamente, uma localização com pouca concorrência pode tornar mais fácil para um hotel atrair clientes, manter preços mais elevados e maiores performances (Lado-Sestayo & Vivel-Búa, 2018a), desde que, consiga fazê-los (aos clientes) chegar a esse destino, tarefa nem sempre fácil fruto das limitações das ligações aéreas ou das ligações rodoviárias de um destino em desenvolvimento.

3.2.3 Outros fatores da localização

A proximidade a zonas urbanas também é um fator determinante na performance hoteleira (Lado-Sestayo *et al.*, 2020). Estes autores concluíram que a proximidade ao CBD, ou zonas centrais de negócios, tinham impacto na rendibilidade dos hotéis, relativamente a hotéis de cidade. Do mesmo modo, a proximidade a vias de transportes, como aeroportos, são igualmente determinantes a considerar na análise da localização. Lado-Sestayo e Vivel-Búa (2019), desenvolveram um algoritmo, utilizando a inteligência artificial por via de redes neurais, que pretendia prever a rendibilidade de um hotel tendo em consideração múltiplos determinantes. Neste âmbito, concluíram que a proximidade de hotéis ao aeroporto e estações de comboios, contribuí significativamente para a sua performance. Aliás, anteriormente esta relação já havia sido identificada, ainda que parcialmente, num estudo desenvolvido no setor hoteleiro espanhol por Lado-Sestayo *et al.* (2017). Neste último, as conclusões apontavam para uma relação pouco significativa, ainda que existente, relativamente aos aeroportos. Porém, a proximidade geográfica a estações de comboios era significativamente relevante.

3.3. Dimensão

A dimensão de uma unidade hoteleira, a par da localização, é um fator de profunda análise aquando da decisão de investimento (Tan *et al.*, 2022).

Nesse sentido, vários estudos têm investigado o impacto da dimensão sobre a performance hoteleira (Assaf & Tsionas, 2018; Barros, 2005; Bresciani *et al.*, 2015; Claver-Cortés *et al.*, 2007; Dimitrić *et al.*, 2019; Kim *et al.*, 2013; Lado-Sestayo *et al.*, 2016, 2020; Lado-Sestayo & Vivel-Búa, 2018a, 2018b, 2019; Menicucci, 2018; Travanca *et al.*, 2022) ou na performance em geral (Lee, 2009; Serrasqueiro & Nunes, 2008).

Enquanto determinante de performance, a dimensão constitui um fator que pode ter um impacto significativo no desempenho e pode ser avaliado sob diversas perspetivas ou métricas. A dimensão de uma unidade pode ser avaliada, *exempli gratia*, pelo número de quartos ou vendas totais (Dimitrić *et al.*, 2019), pelo valor do ativo (Lado-Sestayo & Vivel-Búa, 2018a, 2019), número de camas ou empregados (Romero *et al.*, 2020), ou como hotéis grandes ou pequenos (Assaf & Tsionas, 2018; Menicucci, 2018).

Importa relevar que numa perspetiva empírica, a dimensão de um hotel é comumente avaliada em número de quartos (Assaf & Tsionas, 2018; Barros, 2005; Bresciani *et al.*, 2015; Claver-Cortés *et al.*, 2007; Kim *et al.*, 2013; Lado-Sestayo *et al.*, 2016, 2020; Lado-Sestayo & Vivel-Búa, 2018b; Schwartz *et al.*, 2017), podendo ser categorizada da seguinte forma:

Tabela 3.1 - Escala dimensional de uma unidade hoteleira

Nº de UA	Classificação
0 a 25	Pequeno
26 a 99	Médio
100 a 299	Grande
acima de 300	Muito grande

Fonte: Siteminder, (sem data)

Dos estudos analisados, a maioria evidencia uma relação positiva entre a dimensão e a performance (Claver-Cortés *et al.*, 2007; Dimitrić *et al.*, 2019; Kim *et al.*, 2013; Lado-Sestayo *et al.*, 2016, 2020; Lado-Sestayo & Vivel-Búa, 2018a, 2018b; Menicucci, 2018; Romero *et al.*, 2020; Travanca *et al.*, 2022), justificada maioritariamente por economias de escala.

No entanto, considerando a complexidade do tema e o efeito que este determinante pode ter sobre outras variáveis, encontrou-se diferentes relações entre a dimensão e a performance.

Por exemplo, o impacto da dimensão na performance é substancialmente diferente em hotéis próximos ou afastados de centros urbanos, com o impacto da dimensão a ser superior nos primeiros (Lado-Sestayo *et al.*, 2020).

Nesta perspetiva, também Romero *et al.*, (2020) encontraram evidências que o impacto do tamanho na performance era significativamente maior nos hotéis costeiros do que nos hotéis não costeiros, apontando maiores níveis de procura das zonas costeiras, em resultado do mercado dos 3S.

Similarmente, o efeito da dimensão na performance não é igual quando investigado em hotéis em Portugal, onde o efeito é positivo, ou na Croácia e Grécia onde não existe uma relação estatística significativa (Dimitrić *et al.*, 2019).

Kim *et al.*, (2013), operacionalizando a performance por via do RevPAR e do TRevPAR, identificaram uma relação positiva entre a dimensão e a performance, em linha com o exposto anteriormente. No entanto, visando aprofundar este impacto, em particular quanto ao tipo de hotel, verificaram que uma maior dimensão, podia no caso de hotéis resort ou de negócios, ter um efeito negativo na performance. Segundo os autores, menor dimensão implica menor capacidade de estruturas de F&B, *id est* um menor número de lugares sentados, o que limita a aceitação de grupos e conseqüentemente, induz venda de quartos a turistas a um ADR superior. Por esta razão o impacto da dimensão nos rendimentos de F&B eram estatisticamente insignificantes.

Apontando limitações ao RevPAR ou TRevPAR, outros autores optaram por indicadores alternativos de performance, como o GOP (Claver-Cortés *et al.*, 2007) ou o GOPPAR (Menicucci, 2018), sendo que as conclusões foram similares, apontando neste último, economias de escala, maior leque de serviços e conseqüentemente maior procura, como algumas das justificações.

Bresciani *et al.* (2015) não encontraram relação estatísticas significativas entre a dimensão e a performance, justificada pela localização das unidades em estudo. Segundo os autores, as unidades de maiores dimensões eram localizadas em zonas periféricas, e em alguns casos, localizações remotas.

Na mesma linha, ainda que relativamente à eficiência enquanto medida de performance, Assaf e Tsionas (2018) concluíram que nem sempre hotéis de maior dimensão apresentam melhores performances.

Em síntese, existe uma relação entre a dimensão, medida maioritariamente sob a forma de número de quartos, e a performance. No entanto, não pode ser negligenciado o

impacto que a conjugação de outros determinantes, produzem na dimensão, circunstância que induz a significativas e relevantes variações.

3.4. Categoria

A categorização de hotéis por intermédio do sistema de estrelas, é uma forma de avaliar a qualidade dos hotéis, sendo que as estrelas representam a qualidade geral dos serviços e das instalações do hotel (Bresciani *et al.*, 2015). Universalmente, quanto mais estrelas um hotel tiver, mais luxuoso e completo será, sendo as avaliações baseadas em critérios como a qualidade das instalações, a variedade de serviços oferecidos, a localização e a qualidade do atendimento ao cliente, áreas, entre outros (Oliveira *et al.*, 2013).

Em Portugal, esta categorização está definida no RJET, através do Decreto-Lei n.º 39/2008 de 7 de março, com as alterações produzidas por: Declaração de Retificação n.º 25/2008; Decreto-Lei n.º 228/2009; Decreto-Lei n.º 15/2014; Declaração de Retificação n.º 19/2014; Decreto-Lei n.º 128/2014; Decreto-Lei n.º 186/2015; Decreto-Lei n.º 80/2017; Decreto-Lei n.º 9/2021.

Este documento legislativo, remete nos termos do seu artigo 4º para a Portaria n.º 309/2015 retificada pela Declaração de Retificação n.º 49/2015 que detalha os requisitos que cada componente, sejam instalações, equipamentos e mobiliário ou serviços, entre outros, terão que ser cumpridos para pertencer a uma determinada categoria. A conjugação de 197 requisitos, tal como definidos neste regulamento, determinam o que será considerada uma unidade de cinco estrelas, a mais luxuosa e uma unidade de uma estrela, a mais económica, em processos de auditoria e atribuição desenvolvidos pelo Turismo de Portugal, no caso dos estabelecimentos hoteleiros. O objetivo, tal como definido no preâmbulo desta Portaria, visa estabelecer uma marca de qualidade junto do público geral, uma vez que a grande maioria não consegue detalhar especificamente em que consistem as diferenças. Dito de outra forma, procurou estabelecer uma relação crescente entre o nível de qualidade e o número de estrelas.

Apesar de ser possível encontrar este sistema e esta relação mundialmente, na realidade não existe um sistema único e universal adotado pelos vários países (Oliveira *et al.*, 2013). Não obstante, os estudos analisados apontam a utilização das estrelas como principal métrica na avaliação do impacto da categoria sobre a performance (Bresciani *et al.*, 2015; Claver-Cortés *et al.*, 2007; Lado-Sestayo *et al.*, 2020; Lado-Sestayo & Vivel-Búa, 2018a; Oliveira *et al.*, 2013).

A relação entre a categoria e a performance de uma unidade hoteleira, não encontra conclusões unânimes.

Algumas das investigações desenvolvidas e analisadas apontam para uma relação positiva entre o número de estrelas e a performance (Bresciani *et al.*, 2015; Claver-Cortés *et al.*, 2007). Bresciani *et al.*, (2015) concluíram que a categoria, avaliada na métrica do número de estrelas possuía uma elevada correlação com a performance, medida por via do RevPAR. Identificaram igualmente uma relação entre dimensão e categoria, nomeadamente quanto a hotéis maiores possuírem tendencialmente categorias superiores.

Esta conjugação de determinantes é aplicada em outros estudos, produzindo resultados distintos, consoante os determinantes considerados. Tendo por exemplo a localização, em particular a proximidade a um aeroporto, revelou que o impacto das estrelas num hotel de aeroporto comparativamente a um hotel de urbano é estatisticamente insignificante (Lado-Sestayo *et al.*, 2020).

Lado-Sestayo e Vivel-Búa (2018a) analisaram o impacto da categoria na rendibilidade dos hotéis espanhóis no período de 2005-11, não encontraram evidências no sentido de concluir pelo contributo positivo da categoria na rendibilidade. Oliveira *et al.* (2013) também não encontram, ainda que de forma limitada, evidências de uma relação positiva entre um maior número de estrelas e uma maior eficiência.

3.5. As pequenas e médias empresas na hotelaria nacional

De acordo com a recomendação da Comissão Europeia 2003/361/EC de 6 de maio, transposta para o normativo nacional pelo Decreto Lei nº 372/2007, uma PME (pequenas e médias empresas), é definida como uma empresa que emprega menos de 250 colaboradores e com um volume de negócio não superior a 50 milhões de euros e/ou um balanço não superior a 45 milhões de euros.

Tal como sucede no tecido empresarial nacional, também no turismo, em particular no setor hoteleiro, as PME representam uma percentagem muito significativa deste universo (Travanca *et al.*, 2022).

Romero *et al.* (2020) analisaram 17 447 empresas espanholas do setor turístico, que em 2018 possuíam pelo menos um colaborador tendo apurado que 99% eram PME. Este estudo evidenciou as dificuldades que as PME têm que enfrentar na negociação com organizações de maior dimensão, em particular os TO, dado o seu fraco poder negocial.

As PME hoteleiras possuem recursos limitados, como sejam os recursos humanos ou tecnológicos, o que pode dificultar a sua competição num mercado cada vez mais competitivo (Alvarez *et al.*, 2021).

Por outro lado, as PME hoteleiras também têm algumas vantagens, com destaque para a flexibilidade. As estruturas de recursos reduzidas são uma mais valia em períodos de crise e potencialmente na rapidez de tomada de decisão (Panno, 2019).

3.6. Síntese conclusiva

Da análise à informação disponível e investigação analisada, validou-se a importância do turismo enquanto atividade económica catalisadora do desenvolvimento do país. Neste sentido, é da maior relevância intensificar a investigação científica por forma a que os *stakeholders* do setor estejam melhor capacitados e suportados para o processo de decisão, com vista a um setor mais competitivo na competição global.

Regista-se, no que respeita aos estabelecimentos hoteleiros, a importância da localização na sua performance, mas também a sua complexidade e diferentes abordagens.

A performance tem sido abordada sob duas principais perspetivas, uma relativa, medida sob forma de eficiência e outra absoluta, medida sob forma de rendibilidade. Relativamente à rendibilidade, esta avaliação incide maioritariamente sobre os dados do hotel ou sobre os dados públicos da empresa. A opção por esta tem sido muitas vezes adotada pela facilidade no acesso à informação, ainda que condicionada pela possibilidade de uma empresa gerir várias unidades hoteleiras com características e localizações diferenciadas. A investigação tendo por base a unidade hoteleira, divide-se entre indicadores tidos por operacionais ou setoriais, tal como o RevPAR, TRevPAR, GOP ou GOPPAR e indicadores contabilísticos ou financeiros, como a rendibilidade do ativo, ou o resultado antes de juros, depreciações, amortizações e impostos, entre outros indicadores.

Neste âmbito, a adoção do USALI tem possibilitado uma maior comparabilidade e adoção de indicadores mais próximos da realidade operacional sendo incorporados nas investigações mais recentes e contribuindo para a discussão atual entre RevPAR, TRevPAR ou GOPPAR, enquanto melhor avaliador de performance operacional.

Quanto à localização, o tema é igualmente complexo e pode ser abordado em diferentes perspetivas. Existem correntes dedicadas à localização enquanto proximidade a centros urbanos, de negócios ou de conferências, outras correntes dedicadas à proximidade a infraestruturas de transportes como aeroportos ou estações de comboios e outras correntes dedicadas à proximidade costeira. Esta última tem particular ênfase nos países do mediterrâneo devido à sua elevada sazonalidade e predominância do turismo de sol e praia (3S).

As conclusões não são no geral consensuais. No entanto, relativamente à proximidade costeira, existem evidências de uma relação positiva.

Na mesma linha, a dimensão, medida maioritariamente em número de quartos, possui, de acordo com os estudos analisados, uma relação positiva. No entanto, estas conclusões enumeram fatores que podem condicionar estas conclusões como o tipo de hotel ou a proximidade a determinadas estruturas ou focos de atração (localização).

Relativamente à categoria, as conclusões não são unânimes e uma vez mais, é referida a conjugação com outros determinantes e a especificidade das características das unidades hoteleiras. Uma unidade com mais estrelas, disponibiliza um número superior de serviços, infraestruturas e qualidade, sendo expectável um preço de venda superior. A forma como esse preço de venda é convertido em resultado e avaliado enquanto performance é uma questão que não reúne consenso e que aparenta não explicar *per si* a performance.

Por fim, sendo a maioria das empresas nacionais PME, as evidências apontam para uma maior flexibilidade destas por contraponto a uma menor capacidade negocial junto de terceiros, não havendo evidências objetivas da influência desta na determinante localização.

4. Metodologia da investigação

Tendo por base a revisão da literatura sobre o tema em estudo, desenvolvida no capítulo anterior, o presente capítulo começará por explicar a formulação das hipóteses em estudo e sua relação com os estudos anteriores.

Abordar-se-á igualmente a operacionalização das variáveis, descrevendo e explicitando-se os indicadores e métricas utilizados, bem como as opções tomadas a este respeito.

Em seguida, far-se-á ainda uma descrição do método de investigação aplicado para testar as hipóteses formuladas. Neste âmbito, explicitar-se-á a população em estudo, a forma e as fontes de obtenção do mesmo. Abordar-se-á em detalhe o instrumento de recolha, em particular quanto à sua estrutura, conteúdo e objetivos das questões, mas também quanto ao processo de recolha, esforços desenvolvidos e dificuldades que se encontraram até à constituição da amostra. Abordar-se-ão também as ferramentas utilizadas na análise dos dados recolhidos, as variações desenvolvidas e a opção pelo modelo final.

Por fim, far-se-á uma descrição das técnicas estatísticas utilizadas na análise dos dados.

4.1. Formulação das hipóteses da investigação

Nesta secção formular-se-ão as hipóteses de investigação, tendo por base a revisão da literatura e os objetivos da presente investigação.

Uma das questões principais de estudo, a localização, compreende particular complexidade e um carácter multidimensional (Lado-Sestayo *et al.*, 2020; Lado-Sestayo & Vivel-Búa, 2018a). Assim, optou-se por analisá-las sob diferentes perspetivas, combinando o contributo individual dos determinantes. Nesse sentido, a primeira hipótese em estudo pretende confirmar a existência de uma relação positiva entre a proximidade costeira e a performance. Por outras palavras, quanto mais próximo da costa estiver uma unidade hoteleira, maior será a sua performance, em linha com o estudo de Menicucci (2018). Sendo assim:

Hipótese 1.1: A performance de uma unidade hoteleira é positivamente influenciada pela proximidade costeira, *ceteris paribus*.

Considerando a importância da estrutura do mercado, Lado-Sestayo *et al.* (2016), identificaram uma forte relação entre a procura e a performance, associada a destinos com elevada aglomeração. Em linha com este estudo, Travanca *et al.* (2022) apontaram evidências no mesmo sentido, relativamente ao destino Algarve, tradicionalmente caracterizado por uma forte aglomeração. Assumiu-se, portanto, que as localizações com maiores níveis de aglomeração produzem melhores performances e rendibilidades. Assim temos:

Hipótese 1.2: A performance de uma unidade hoteleira é positivamente influenciada pela aglomeração, *ceteris paribus*.

Outro determinante em estudo é a dimensão da unidade hoteleira. Vários estudos apontam para uma relação positiva entre a dimensão e a performance (Claver-Cortés *et al.*, 2007; Dimitrić *et al.*, 2019; Kim *et al.*, 2013; Lado-Sestayo *et al.*, 2016, 2020; Lado-Sestayo & Vivel-Búa, 2018a, 2018b; Menicucci, 2018; Romero *et al.*, 2020; Travanca *et al.*, 2022), apesar de Lado-Sestayo *et al.* (2020) e Romero *et al.*, (2020) referirem

impactos distintos para hotéis próximos/distantes de centros urbanos e próximos/distantes das zonas costeiras. Não obstante:

Hipótese 2: A dimensão influencia positivamente a performance, *ceteris paribus*.

A categoria, medida sob o número de estrelas, infere que quanto maior o número de estrelas, maior a qualidade do serviço e conseqüentemente maior o preço que o mercado estará disposto a pagar pelo serviço. Nesse sentido, Bresciani *et al.*, (2015) concluiu pela existência de uma elevada correlação entre a categoria e a performance, medida neste estudo através do RevPAR.

Em sentido contrário, não foram encontradas relações estatísticas significativas entre a categoria e a performance nos hotéis Espanhóis (Lado-Sestayo & Vivel-Búa, 2018a), e em hotéis de aeroporto (Lado-Sestayo *et al.*, 2020). Também em Portugal, ainda que visando o estudo da eficiência, Oliveira *et al.* (2013) não encontraram evidências de uma relação positiva entre a categoria e eficiência. Tendo por base o exposto anteriormente:

Hipótese 3: A categoria de uma unidade hoteleira não influencia a sua performance, *ceteris paribus*.

As PME possuem de um modo geral e em resultado da sua dimensão, uma fraca capacidade negocial junto de empresas da maior dimensão, como por exemplo os TO (Romero *et al.*, 2020). Adicionalmente, as PME possuem limitações em recursos humanos e tecnológicos (Alvarez *et al.*, 2021) que por outro lado contribuem para uma maior flexibilidade e capacidade de rápida adaptação (Panno, 2019).

Assim formula-se:

H1.4: A classificação como PME não influencia a relação com a performance de uma unidade hoteleira, *ceteris paribus*.

4.2. Operacionalização das variáveis

Conforme foi amplamente referido, a multiplicidade de perspectivas e dimensões sob as quais a performance pode ser analisada, contribui para a existência de inúmeros indicadores utilizados nos estudos e artigos analisados.

Nesse sentido, importa definir a variáveis em estudo, quer relativamente à performance, quer relativamente às determinantes em investigação.

Começando pela performance, o RevPAR e o GOPPAR têm sido utilizados de forma consistente como referenciais nesta métrica (Bresciani *et al.*, 2015; Claver-Cortés *et al.*, 2007; Kim *et al.*, 2013; Lee *et al.*, 2019; Schwartz *et al.*, 2017; Singh & Schmidgall, 2002; Tan *et al.*, 2022).

Apesar das limitações do primeiro, isto é, a não incorporação de outros rendimentos ou a não incorporação de gastos na sua fórmula de cálculo, o estudo de Schwartz *et al.*, (2017) concluiu de um modo geral que o RevPAR e o GOPPAR estariam significativamente relacionados. No entanto, a ausência de um número substancial de estudos comparativos entre estas métricas (Lee *et al.*, 2019; Schwartz *et al.*, 2017), poderão, dada a complexidade das variáveis, limitar as conclusões, pelo que no presente estudo optou-se por utilizar o GOPPAR como indicador de performance e utilizar o RevPAR como indicador secundário de controlo.

Assim, temos:

VARIÁVEIS DEPENDENTES

GOPPAR (*Gross Operating Profit per Available Room*) ou Resultado Operacional por quarto disponível

$$\text{GOPPAR} = \frac{\text{Gross Operating Profit}_n}{\text{Total de quartos disponíveis}_n}$$

RevPAR (*Revenue per Available Room*) ou Rendimento por quarto disponível

$$\text{RevPAR} = \frac{\text{Total de rendimentos de alojamento}_n}{\text{Total de quartos disponíveis}_n}$$

VARIÁVEIS INDEPENDENTES

Aglomeração: Número de empreendimentos turísticos existentes num raio de 5 quilómetros (quantitativa contínua)

Dimensão: Número de unidades de alojamento (quantitativa discreta)

PME (Pequenas e médias empresas): variável binária (*dummy*; 0 – “Não”, 1 – “Sim”)

Categoria: Número de estrelas conforme classificação do Turismo de Portugal, compreendida entre 1 a 5 estrelas (quantitativa discreta)

Localização costeira: variável binária (*dummy*; (0 – “Não”, 1 – “Sim”), representadas pela localização de um NUTSIII Costeiro ou não.

4.3. Definição da população em estudo

A população em estudo teve origem na informação constante dos registos disponíveis na base de dados pública do Registo Nacional de Empreendimentos Turísticos (RNET), consultada a 04 de fevereiro de 2023 e disponível a partir do portal do turismo de Portugal, na área “Business” em:

http://business.turismodeportugal.pt/pt/Planear_Iniciar/Licenciamento_Registo_da_Atividade/Empreendimentos_Turisticos/Paginas/rnet.aspx.

Importa referir que o RNET disponibilizava um total de 5.249 empreendimentos turísticos, sendo que à mesma data a plataforma do TravelBI *Open Data*, acessível em <https://dadosabertos.turismodeportugal.pt/datasets/empreendimentos-turisticos-existentes/data> disponibilizava 4.849 registos para empreendimentos turísticos. A informação dos metadados¹⁷ de acordo com a ISO-19139¹⁸ apontava para uma última atualização a 2021-07-30, o que poderia justificar a divergência.

Tendo em vista a correta utilização dos dados, contactou-se o TP, que informou por escrito que a data indicada corresponderia à data da criação do serviço. Esta informação seria sincronizada diariamente com o registo nacional de turismo. Clarificou ainda que a divergência encontrada corresponde aos complexos turísticos e aos ET da Madeira e Açores, excluídos dos dados disponibilizados no portal business do Turismo de Portugal.

Por esse motivo decidiu-se avançar com os dados recolhidos do RNET, sendo que atendendo ao objeto de estudo desta dissertação, selecionou-se apenas os estabelecimentos hoteleiros. Dessa forma, utilizou-se o filtro disponibilizado pelo RNET, selecionaram-se os estabelecimentos hoteleiros e extraiu-se o *dataset* que compunha Hotéis, Hotéis-Apartamento e Pousadas, num total de 1.886 registos.

¹⁷ Características de um conjunto de dados, tradicionalmente, informáticos

¹⁸ Padrão internacional desenvolvido pela Organização Internacional para Padronização (ISO) que define os requisitos para a implementação de esquemas XML (eXtensible Markup Language) que são derivados de UML (Unified Modeling Language) de acordo com as regras estabelecidas na norma ISO 19118. Este padrão é uma parte da série ISO 19100 de normas geoespaciais e é especificamente focado na representação de metadados geoespaciais

4.4. Instrumento de recolha

O *dataset* mencionado na seção anterior foi analisado e posteriormente transformado através de um *notebook* em *Python*¹⁹ (Apêndice A). O código *Python* em questão foi elaborado com o objetivo de criar uma base para o envio dos inquéritos, que permitisse de forma automatizada e anonimizada, em tratamento posterior, relacionar colunas de informação adicionais com robustez.

Assim, o código importava a necessária biblioteca *Pandas*²⁰, essencial para o tratamento e análise de dados. Após o carregamento do ficheiro, visualização da estrutura e dos primeiros registos, foi criado um subconjunto dos dados originais, focando em colunas específicas e necessárias para o inquérito que foram armazenadas num novo *dataset*. As colunas foram renomeadas para termos mais claros e alinhados com a estrutura predefinida, e a coluna chave foi limpa de caracteres não numéricos e o seu tipo de dados convertido para inteiro. Finalmente, o *dataset* resultante foi exportado para um ficheiro CSV e armazenado para posterior tratamento.

Em suma, o *dataset* continha 42 colunas de informação a partir do qual foi produzido um *dataset* contendo a designação da unidade, a entidade exploradora, o contacto de email e o registo. Conforme referido este *dataset* foi exportado e integrado no *Limesurvey*²¹, ferramenta de questionários adotada para a recolha dos dados.

Nesta, desenhou-se um inquérito subordinado ao tema em investigação sustentados nos seguintes pressupostos:

- Rapidez de resposta – O inquérito deveria ser de resposta fácil e rápida, visando aumentar a probabilidade de resposta;
- Objetividade – O inquérito deveria conter o mínimo de variáveis redundantes focando-se nos dados essenciais e não disponíveis ao público;
- Confidencialidade – Considerando as variáveis em estudo, em particular o GOPPAR, e a resistência setorial em facultar estes dados, era elementar reforçar a garantia de máxima confidencialidade;

¹⁹ Linguagem de programação com ênfase na legibilidade do código, amplamente utilizada para diversas aplicações

²⁰ *Pandas* é uma biblioteca de software criada para a linguagem *Python* para manipulação e análise de dados. Em particular, oferece estruturas e operações para manipular tabelas numéricas e séries temporais. É software livre sob a licença licença BSD. Vide <https://pandas.pydata.org/> para mais informação.

²¹ *LimeSurvey* é um software livre para aplicação de questionários online escrito em *PHP*, podendo utilizar bancos de dados *MySQL*, *PostgreSQL* ou *Microsoft SQL Server* para persistência de dados. Vide www.limesurvey.org para mais informação.

- Funcionalidade e facilidade – O inquérito deveria ser funcional e simples de compreender e preencher;

Admitindo-se a hipótese dos dados financeiros relativos ao RevPAR e GOPPAR suscitarem desconforto junto dos respondentes, assumiu-se a possibilidade de um baixo nível de respostas. Não obstante, terminado a primeira versão do inquérito, desenvolveu-se um inquérito piloto, partilhado com cinco gestores hoteleiros e com os orientadores desta dissertação, e após pequenos ajustamentos, obteve-se o inquérito final (Apêndice B).

O questionário foi concebido para ser preenchido em menos de cinco minutos, sendo composto por sete questões, focadas nas características e na rendibilidade operacional da unidade, sendo subdividido em dois grupos.

O primeiro grupo subordinado ao enquadramento da unidade hoteleira, inclui questões sobre

- A integração do hotel numa Pequena e Média Empresa (PME);
- A existência de outros negócios integrados na empresa além do hotel em análise;
- A conformidade do hotel com o *Uniform System of Accounts Lodging Industry* (USALI);
- O número de estrelas do hotel;
- O número total de unidades de alojamento, com uma resposta limitada entre 0 e 1000;

O segundo grupo incidia sobre os dados financeiros em estudo, para o período de 2017 a 2021, especificamente o *Revenue per Available Room* (RevPAR) e o *Gross Operating Profit per Available Room* (GOPPAR). São fornecidas instruções para inserir 0 (ZERO) nos anos em que não exista registo ou atividade.

O questionário conclui com uma expressão de gratidão pela participação e oferece a possibilidade de partilhar as principais conclusões do estudo com os respondentes, mediante o fornecimento de um contacto.

Após validação do questionário por via dos pré-testes, enviou-se o questionário para os endereços de correio eletrónico constantes do RNET.

Sobre este processo, a primeira nota vai para o considerável número de endereços de email constantes do RNET que se encontravam desatualizados ou inexistentes e que foram analisados, corrigidos e reenviados sempre que tal foi possível. A segunda nota,

para as 108 recusas expressas de participação, através do exercício de opção disponível para o efeito, num valor próximo dos 5,7% da população em estudo.

O inquérito esteve ativo de 07 de fevereiro de 2023 a 06 de junho de 2023. Ao longo destes quatro meses, foram obtidos 97 inquéritos completos. O total de respostas, atingiu aproximadamente 5% da amostra, abaixo da expectativa inicial de 10%.

Reforça-se que este valor só foi possível atendendo às insistências desenvolvidas, através das funcionalidades disponíveis pela ferramenta utilizada²², através de contactos telefónicos e de contactos diretos por telefone ou por via da plataforma *linkedin*.

²² Ao longo do período em que o inquérito esteve ativo, foram utilizados os lembretes, corrigidos emails devolvidos, e respondidas dúvidas ou questões sobre o estudo

4.5. Caracterização da população em estudo

Para melhor compreender a caracterização das unidades hoteleiras, procedeu-se a uma análise ao *dataset* recolhido no RNET. Para o efeito, desenvolveu-se novamente um *notebook* para o efeito (Apêndice C), que permitiu uma análise exploratória de dados, visualização e agregação de informações no contexto da indústria hoteleira.

Foi importada a biblioteca base, *Pandas*, e importado o *dataset*, composto pela população em estudo. Após uma inspeção inicial da estrutura dos dados, procedeu-se ao carregamento das bibliotecas para visualização de dados, nomeadamente *Matplotlib*²³ e *Seaborn*²⁴. Os estilos dos gráficos foram definidos e múltiplas visualizações criadas, incluindo gráficos de barras para representar a distribuição de unidades hoteleiras por tipologia, categoria, NUTSIII (Nomenclatura de unidades territoriais para fins estatísticos 3º nível) e distrito, bem como um histograma para o número de unidades de alojamento.

O código desenvolvido realiza agregações de dados para calcular o número e a percentagem de unidades hoteleiras e unidades de alojamento por distrito. Essas informações são então apresentadas sob a forma de uma tabela.

Na etapa subsequente, a biblioteca *Geopandas*²⁵ é importada para permitir a visualização geoespacial dos dados. Foram carregados e manipulados ficheiros *shapefile* para representar os contornos geográficos de Portugal, obtidos em <https://dados.gov.pt/pt/datasets/>. Os dados agregados foram integrados com esses contornos geográficos para produzir *heat maps* apresentando a concentração do número de UA e número de unidades em Portugal Continental, Açores e Madeira.

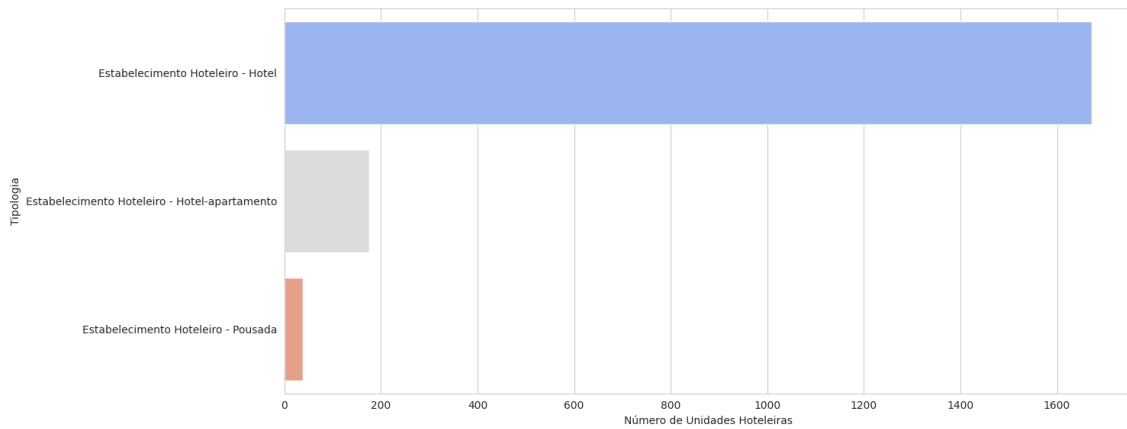
Da análise inicial à população, verifica-se que predominam os hotéis seguidos dos hotéis-apartamentos e por fim as pousadas. Juntos, estes estabelecimentos hoteleiros disponibilizam um total de 136.860 unidades de alojamento.

²³ Biblioteca em linguagem Python para criação de gráficos e visualizações de dados em geral. Para mais informação vide <https://matplotlib.org>.

²⁴ Biblioteca em linguagem Python para visualização de dados baseada na biblioteca Mathplotlib. Para mais informação vide <https://seaborn.pydata.org>.

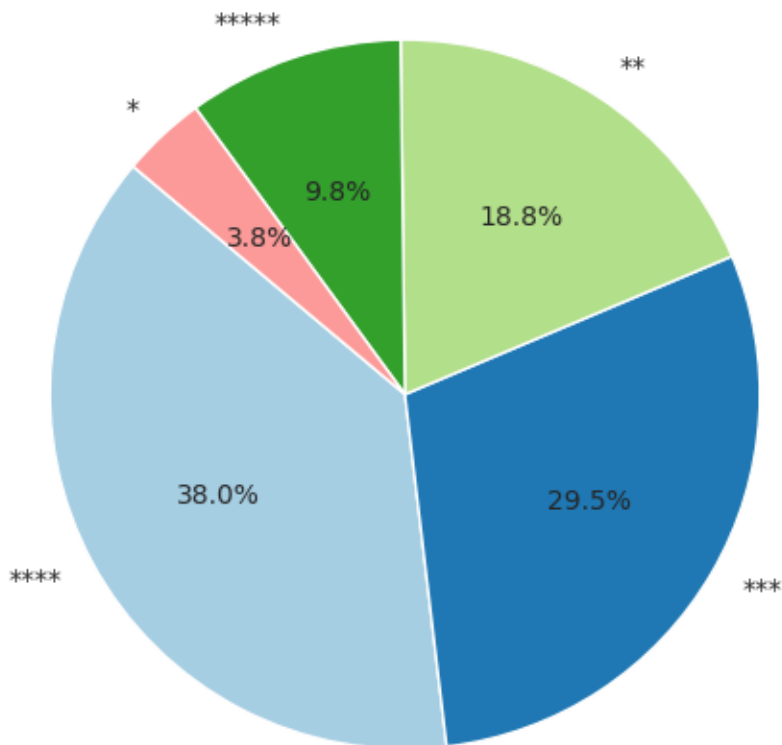
²⁵ Biblioteca em linguagem Python para gestão de dados geoespaciais. Para mais informação vide <https://geopandas.org>

Figura 4.1 - N° de unidades hoteleiras por tipologia



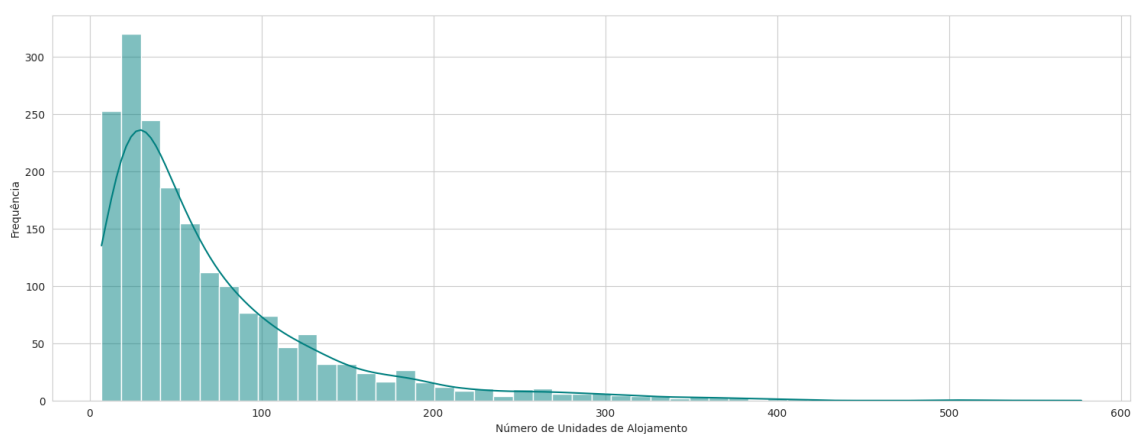
Em termos de categoria, os hotéis de 4 estrelas são os mais prevalentes, com 702 estabelecimentos, seguidos pelos de 3 estrelas e 2 estrelas, com 545 e 348 estabelecimentos, respectivamente.

Figura 4.2- Distribuição da categoria (estrelas)



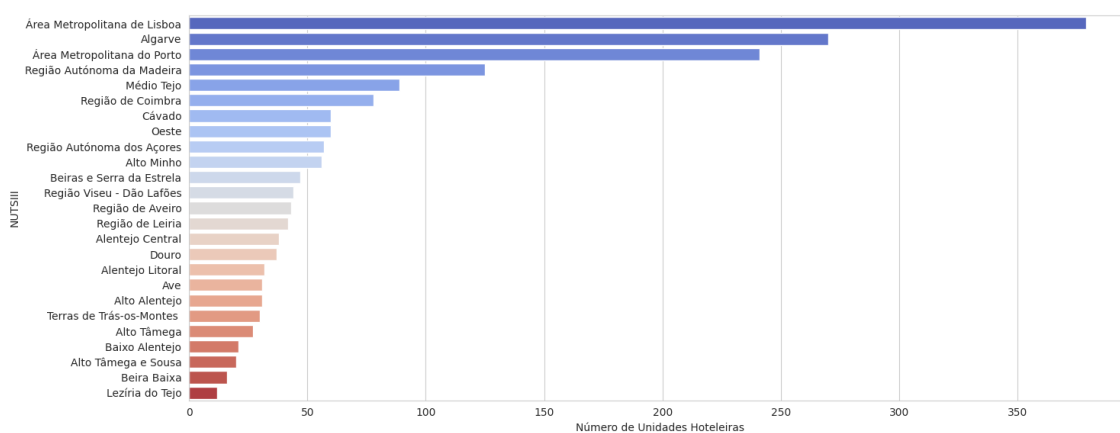
O histograma do número de unidades de alojamento, apresenta uma distribuição com uma assimetria acentuada para a esquerda, com o número médio de UA a situar-se nas 73 UA (figura 4.3).

Figura 4.3 - Distribuição do nº de UA



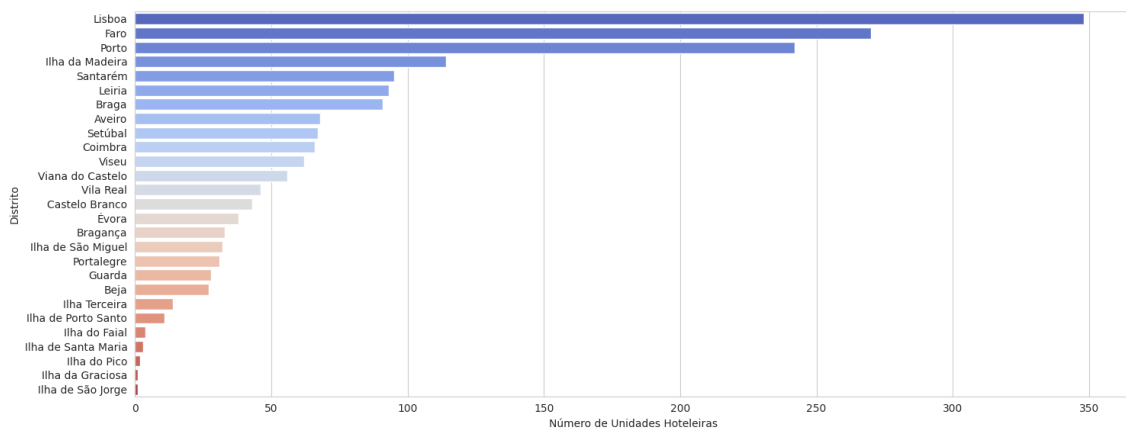
Uma análise à distribuição do número de unidades hoteleiras por NUTSIII evidencia uma concentração de unidades hoteleiras na área metropolitana de Lisboa, Algarve e Área Metropolitana do Porto, cf. figura abaixo.

Figura 4.4 - Unidades hoteleiras por NUTSIII



Em linha com estas indicações, na figura 4.5, verificamos dados similares do número de unidades por Distrito.

Figura 4.5 - Unidades hoteleiras por distrito



Em maior detalhe na tabela 4.1, podemos confirmar a concentração nos distritos de Lisboa, Porto e Faro que representam 32,77% das unidades hoteleiras nacionais, por esta ordem de grandeza.

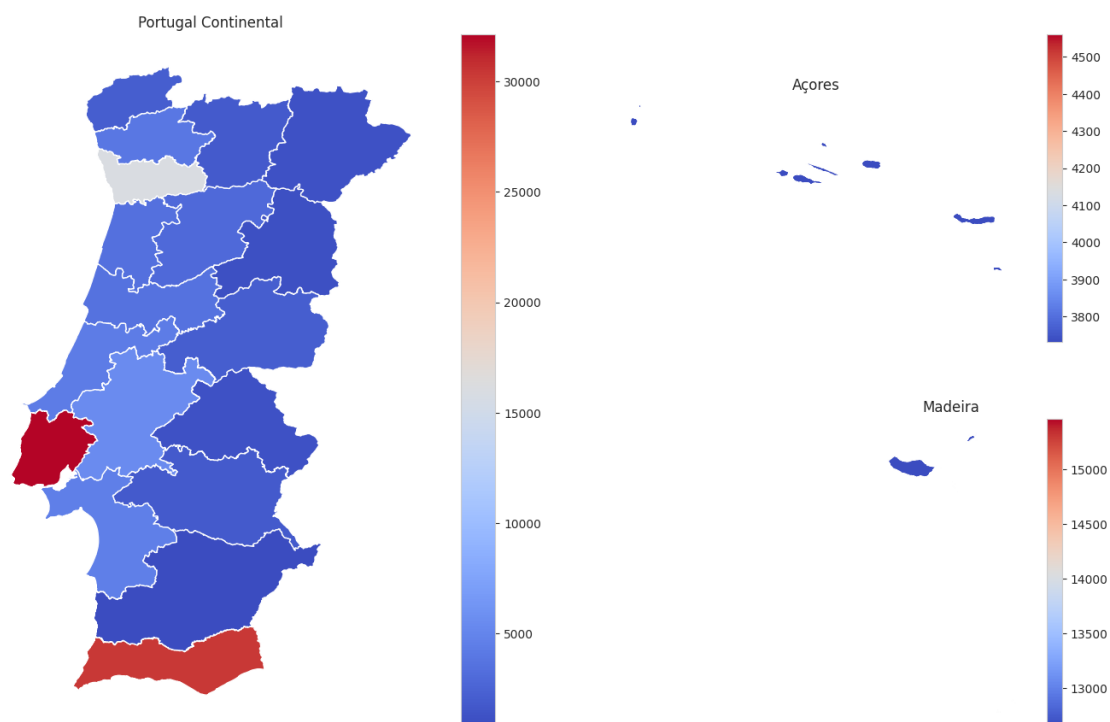
Tabela 4.1 - Estatísticas por distrito (Unidades hoteleiras e unidades de alojamento)

	Nº Unidades Hoteleiras	% do Total UH	Nº UA	% do Total UA	Média de UA	Moda de UA
Aveiro	68.0	3.61	3407.0	2.49	50.0	15.0
Beja	27.0	1.43	791.0	0.58	29.0	24.0
Braga	91.0	4.83	3923.0	2.87	43.0	13.0
Bragança	33.0	1.75	1159.0	0.85	35.0	12.0
Castelo Branco	43.0	2.28	2057.0	1.5	47.0	18.0
Coimbra	66.0	3.5	3496.0	2.55	52.0	15.0
Faro	270.0	14.32	30419.0	22.23	112.0	43.0
Guarda	28.0	1.48	1134.0	0.83	40.0	24.0
Ilha Terceira	14.0	0.74	793.0	0.58	56.0	10.0
Ilha da Graciosa	1.0	0.05	15.0	0.01	15.0	15.0
Ilha da Madeira	114.0	6.04	12896.0	9.42	113.0	52.0
Ilha de Porto Santo	11.0	0.58	1162.0	0.85	105.0	22.0
Ilha de Santa Maria	3.0	0.16	152.0	0.11	50.0	15.0
Ilha de São Jorge	1.0	0.05	58.0	0.04	58.0	58.0
Ilha de São Miguel	32.0	1.7	2610.0	1.91	81.0	54.0
Ilha do Faial	4.0	0.21	341.0	0.25	85.0	27.0
Ilha do Pico	2.0	0.11	177.0	0.13	88.0	40.0
Leiria	93.0	4.93	4280.0	3.13	46.0	30.0
Lisboa	348.0	18.45	32136.0	23.48	92.0	26.0
Portalegre	31.0	1.64	1180.0	0.86	38.0	14.0
Porto	242.0	12.83	16000.0	11.69	66.0	15.0
Santarém	95.0	5.04	5448.0	3.98	57.0	30.0
Setúbal	67.0	3.55	4562.0	3.33	68.0	16.0
Viana do Castelo	56.0	2.97	2025.0	1.48	36.0	20.0
Vila Real	46.0	2.44	1797.0	1.31	39.0	11.0
Viseu	62.0	3.29	2987.0	2.18	48.0	20.0
Evora	38.0	2.01	1855.0	1.36	48.0	25.0

Quando comparada com a percentagem relativa do número de UA, verificou-se que estes três distritos continuam a ocupar o pódio. No entanto, os distritos de Faro e do Porto trocam de posições, com Faro a atingir os 22,23% das UA, valor muito próximo de Lisboa (23,48%). Este dado, é justificado pela média do número de unidades de alojamento no distrito de Faro, que se situa nas 112 UA, só ultrapassada pela Ilha da Madeira com 113 UA em média, destinos associados ao Sol e Praia ou à distribuição via TO.

Considerando uma das hipóteses em estudo, relativa à localização costeira ou não costeira, utilizou-se um *heatmap* ou mapa de calor (técnica de visualização de dados que mostra intensidade de uma frequência ou valor por via de colorização).

Figura 4.7- Heatmap do número de unidades de alojamento



Ademais, também nesta figura podemos observar o peso dos distritos costeiros quando comparados com os do interior.

Acrescenta-se por fim, que em média, os distritos costeiros apresentam 80 UA por estabelecimento, quando os distritos do interior têm uma média de 65 UA. O interior, apresenta não só um menor número de unidades hoteleiras por distrito como unidades hoteleiras mais pequenas.

4.6. Processo de tratamento dos dados

Partindo dos 97 inquéritos completos, as respostas foram exportadas para um ficheiro em formato csv²⁶. Desenvolveu-se um processo de transformação e manipulação de dados, com o intuito de preparar o *dataset* para a análise da regressão. Considerando que a descrição deste processo é fundamental para garantir a replicabilidade e a compreensão da metodologia empregada, descrever-se-á de seguida o procedimento adotado.

Assim, inicialmente, foram importadas as bibliotecas necessárias para a manipulação dos dados, tais como as bibliotecas *numpy* e *pandas*, e realizado o carregamento dos ficheiros, de modo a permitir o acesso aos dados em bruto, inscritos nos ficheiros armazenados.

Nesse âmbito foram carregados dois conjuntos de dados distintos: um contendo as respostas ao inquérito realizado, já referido anteriormente, e outro representando a população. Estes dados foram combinados através de uma chave comum, permitindo a correspondência entre as respostas ao inquérito e os dados integrais. O objetivo foi enriquecer o *dataset* de variáveis adicionais, não incluídas no questionário, mas importantes à análise dos dados.

Na fase subsequente, procedeu-se à eliminação de diversas colunas consideradas redundantes ou fora do espectro de estudo. Esta etapa foi crucial para refinar o conjunto de dados, mantendo apenas as variáveis entendidas como pertinentes à investigação.

Em seguida, efetuou-se a correção dos nomes de algumas variáveis-chave, tornando-os mais compreensíveis e alinhados com o contexto do estudo. Simultaneamente, realizou-se o tratamento de valores em falta nas UA utilizando, nos casos omissos, a informação do RNET. Procedeu-se ainda à substituição de valores inexistentes por "NaN"²⁷, garantindo assim a integridade numérica dos dados relativamente às variáveis dependentes, nomeadamente RevPAR e GOPPAR.

A análise prosseguiu com a eliminação de colunas redundantes e a criação de intervalos para categorizar a proximidade da costa, uma variável de interesse no contexto deste estudo. Algumas colunas foram transformadas em categorias, facilitando a análise subsequente.

²⁶ Formato de ficheiro de dados, separados por vírgulas (*comma-separated values*)

²⁷ Acrónimo de “*Not a Number*”, valor definido em linguagem de programação para um tratamento especial em operações matemáticas

A georreferenciação foi um elemento central nesta fase de tratamento dos dados. Os valores de latitude e longitude foram tratados, separados em colunas distintas e convertidos para o tipo de dados *float*. Registos com valores nulos nessas colunas foram eliminados, assegurando a precisão da análise geográfica. Esta operação reduziu as 97 observações iniciais para 87.

Posteriormente, foram desenvolvidas funções específicas para calcular o número de hotéis e UA circunscritas em determinados raios geográficos. Estas funções foram aplicadas aos dados, permitindo a análise da aglomeração hoteleira em diferentes escalas geográficas (2 km, 5km e 10km).

Depois, procedeu-se à transformação dos dados numa estrutura de painel, processo que multiplicou o número de observações por 5, correspondentes aos cinco anos de registos de RevPAR e GOPPAR que foram inquiridos. Esta operação resultou num *dataset* de 435 observações. Incluiu-se nesta transformação a criação de uma coluna *dummy* 'covid' com base nos valores da coluna 'year', representando o impacto da pandemia na investigação.

Na fase final, procedeu-se ao preenchimento de valores inexistentes em algumas variáveis categóricas e à verificação e tratamento de valores nulos em variáveis como 'PME' e 'USALI', e agrupamento e transformação em variáveis binárias *dummy*. O *dataset* transformado foi então exportado para um ficheiro CSV, preparando-o para análises posteriores, a realizar.

Em suma, através da manipulação de variáveis, tratamento de valores em falta, categorização, e georreferenciação, os dados foram transformados numa forma que potencia a análise e interpretação no contexto da indústria hoteleira.

4.7. Síntese conclusiva

Neste capítulo delineou-se o quadro de investigação para o estudo da performance das unidades hoteleiras. A investigação foi guiada por várias hipóteses, explorando a relação entre a performance e diversos fatores, tais como a proximidade costeira, a aglomeração, a dimensão, a categoria e a classificação como PME.

O capítulo começou com a formulação das hipóteses, fundamentadas numa revisão da literatura existente. Foram identificadas quatro hipóteses principais, relacionando a performance com a proximidade costeira, aglomeração, dimensão e categoria, e uma hipótese adicional relacionada com a classificação como PME.

Em seguida, foram definidas as variáveis dependentes e independentes, incluindo métricas como o GOPPAR e o RevPAR, e variáveis como a aglomeração, dimensão, categoria, entre outras. A população em estudo foi extraída da base de dados pública do RNET, compreendendo um total de 1.886 registos.

O processo de recolha de dados foi detalhado, incluindo a conceção e distribuição de um inquérito, bem como os desafios enfrentados na obtenção de respostas. O inquérito esteve ativo durante quatro meses, resultando em 97 inquéritos completos.

A última parte do capítulo focou-se no tratamento dos dados, descrevendo o processo de transformação e manipulação. Foram realizadas várias etapas, incluindo a eliminação de colunas redundantes, correção de nomes de variáveis, tratamento de valores em falta, categorização, georreferenciação, e transformação dos dados numa estrutura de painel.

Em suma, este capítulo metodológico estabeleceu a base para a análise seguinte, através de uma abordagem que se pretendeu rigorosa e fundamentada. A metodologia adotada reflete uma combinação de revisão teórica, formulação de hipóteses, definição de variáveis, recolha de dados e tratamento de dados, preparando o terreno para uma análise significativa dos determinantes da performance na indústria hoteleira em Portugal.

5. Resultados e discussão

Neste capítulo serão apresentados os resultados produzidos pelos testes estatísticos, relativamente às hipóteses formuladas na seção 3.2. Serão igualmente discutidos os resultados, comparando-os com a investigação já desenvolvida no contexto da realidade da indústria hoteleira em Portugal.

Na segunda secção deste capítulo, far-se-á uma caracterização da amostra, começando por explicar o código utilizado para o efeito, ao que se seguirá a apresentação de um conjunto de representações gráficas que melhor permitam uma adequada perceção das características e estrutura da amostra.

Segue-se, na terceira secção, uma análise aos pressupostos subjacentes aos modelos estatísticos utilizados, incluindo uma breve explicitação dos mesmos.

Na secção seguinte, apresentar-se-á uma descrição ao *notebook* desenvolvido, seguida da apresentação dos resultados da validação dos pressupostos. Depois, apresentar-se-ão os resultados obtidos nos vários modelos, repartidos entre os períodos pré e durante COVID, e entre GOPPAR e RevPAR. Apresentar-se-á uma síntese comparativa dos resultados produzidos e por fim, concluir-se-á este capítulo com uma discussão dos resultados e uma síntese conclusiva.

5.1. Caracterização da amostra

Tal como fez-se anteriormente, para melhor compreensão, inicia-se esta seção com a explicação do *notebook* utilizado para o tratamento da amostra.

Assim, começou-se pelo carregamento das diversas bibliotecas de análise de dados e visualização, já mencionadas nos *notebooks* anteriores. Os dados foram importados para um *dataset* ao que se seguiram verificações quanto à sua estrutura e aos tipos de dados contidos.

Na primeira fase da análise, foi elaborado um gráfico de barras para visualizar a distribuição de unidades hoteleiras por tipologia. Este gráfico foi ajustado para anular a transformação para dados em painel (TDP), efetuada no tratamento dos dados, assegurando desta forma que a qualidade dos resultados. Seguiu-se a produção de um gráfico circular para representar a distribuição das unidades hoteleiras por categorização em estrelas e posteriormente, novo ajustamento à TDP e realização de gráficos de distribuição de unidades hoteleiras por distrito e por NUTSIII.

Este foi complementado com um histograma que apresenta a distribuição do número de unidades de alojamento. Foram também criados dois *heatmaps* que apresentam a distribuição geográfica das unidades hoteleiras. O primeiro mapa exhibe o número total de unidades de alojamento por distrito e o segundo apresenta o número total de unidades hoteleiras por distrito. Adicionalmente, foi desenhada uma função para o cálculo das percentagens para cada variável relativa à proximidade à costa, e visualizados em um gráfico de barras empilhadas.

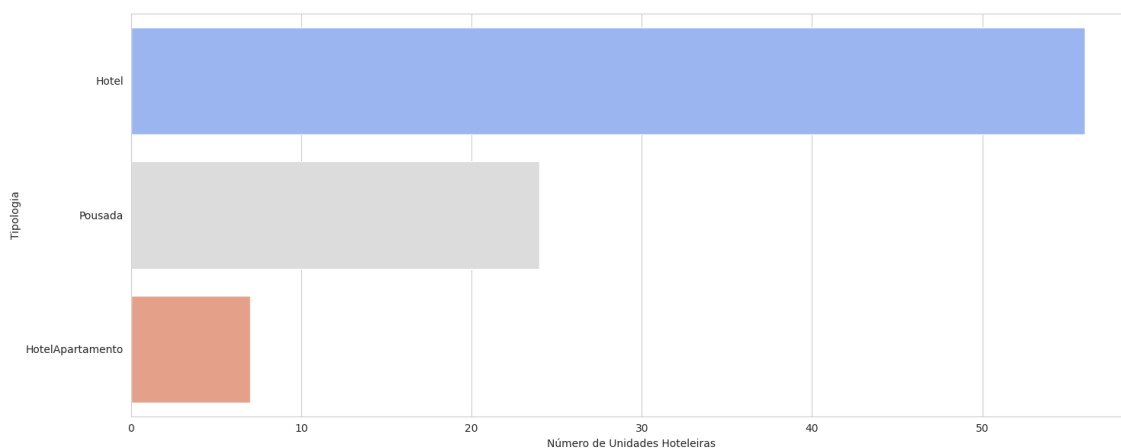
Posteriormente, foram gerados gráficos circulares para as variáveis 'PME' e 'USALI', representando a percentagem de unidades hoteleiras que são PME e que utilizam o sistema USALI, respetivamente, com uma transformação dos valores das legendas de "1" e "0" em "Sim" e "Não" para uma melhor perceção visual.

Foram ainda realizadas análises mais complexas para identificar padrões por regiões NUTSIII e por categoria das unidades hoteleiras, em relação a variáveis independentes como a dimensão da empresa, a aplicação do USALI ou a proximidade costeira.

Estas análises resultaram na criação de gráficos de barras empilhadas, ajustados a 100%, que representam as proporções de cada categoria para as variáveis em questão, para cada NUTSIII e categorização por estrelas. Por fim, o código termina com uma categorização dos hotéis em diferentes níveis de aglomeração com base nos percentis.

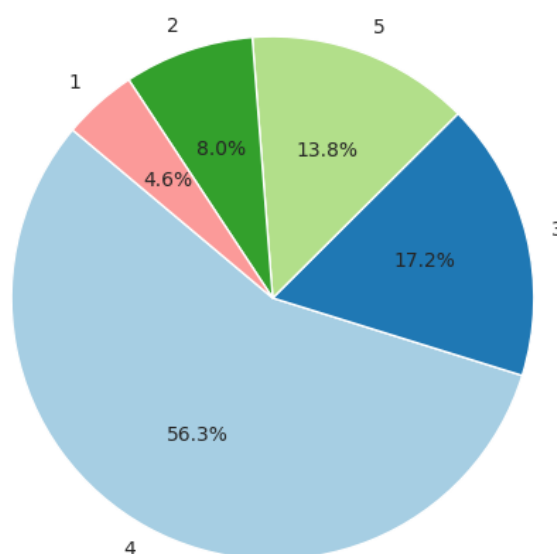
Concluída a descrição do processo metodológico, explicita-se com recurso aos gráficos produzidos, a caracterização da amostra. Composta por 435 registos, correspondentes a 87 unidades hoteleiras, verifica-se uma maior representatividade das pousadas na amostra por comparação à população, conforme figura 5.1. Os hotéis continuam, no entanto, a compor a larga maioria da amostra (56), seguido das pousadas (24) e dos hotéis-apartamentos (7).

Figura 5.1 - Distribuição das unidades hoteleiras por tipologia



Observando as categorias das unidades hoteleiras da amostra, verifica-se que 56,3% das unidades observadas (49) possuem a categoria de quatro estrelas, seguidas de 17,2% das unidades com três estrelas (15) e da categoria de cinco estrelas com 13,8% (12).

Figura 5.2 - Distribuição das unidades hoteleiras por categoria

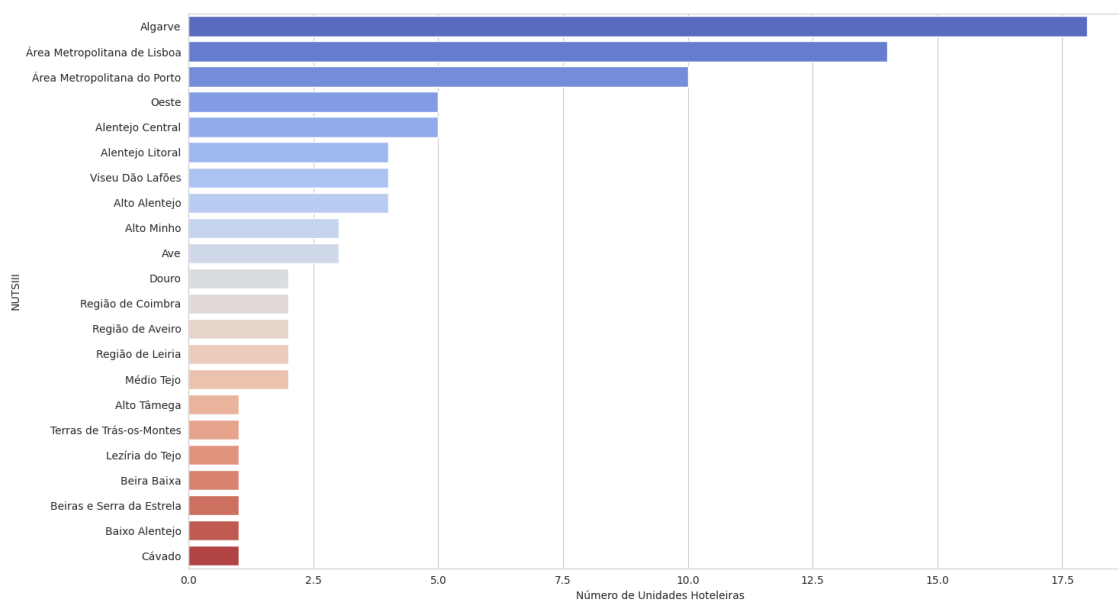


As classificações de cinco e quatro estrelas representam cerca de 70,1% da amostra, que compara com 47,8% representativo do mesmo conjunto de classificações da população.

No que respeita à distribuição geográfica das unidades hoteleiras, em particular quanto ao distrito e regiões NUTSIII, observa-se uma similaridade com a população em estudo.

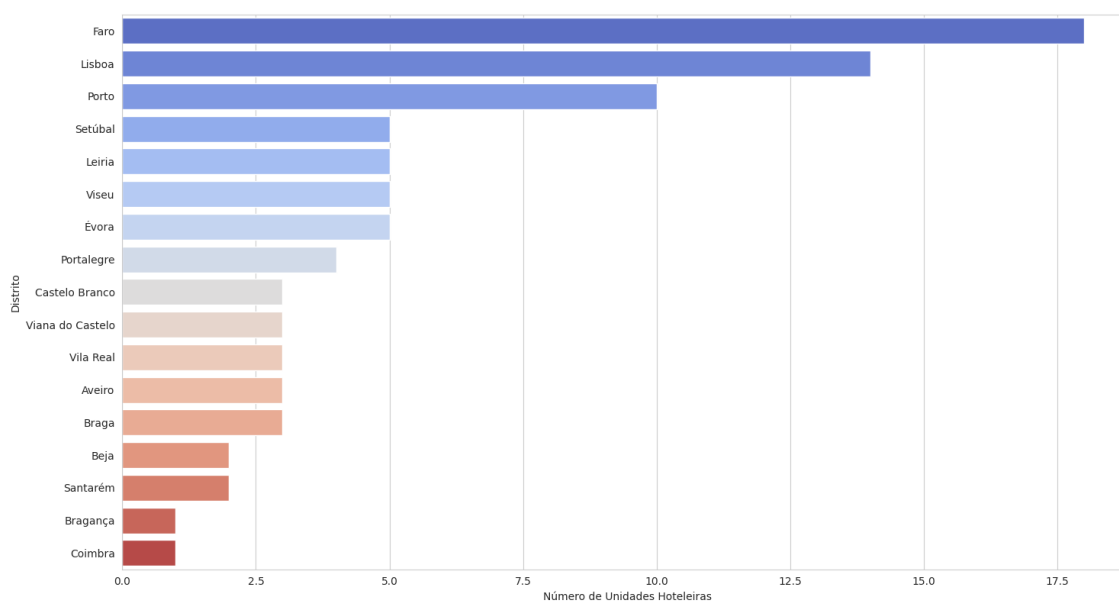
Assim, conforme se pode constatar na figura 5.3, as regiões metropolitanas de Porto e Lisboa e a região do Algarve, destacam-se das demais em linha com o observado anteriormente na população.

Figura 5.3 - Número de unidades hoteleiras por região (NUTSIII)



Nota para uma maior representatividade da região do Algarve, em linha com o observado ao nível dos distritos (figura 5.4), mas diferente do observado na população. Ainda assim, considerou-se a distribuição como semelhante.

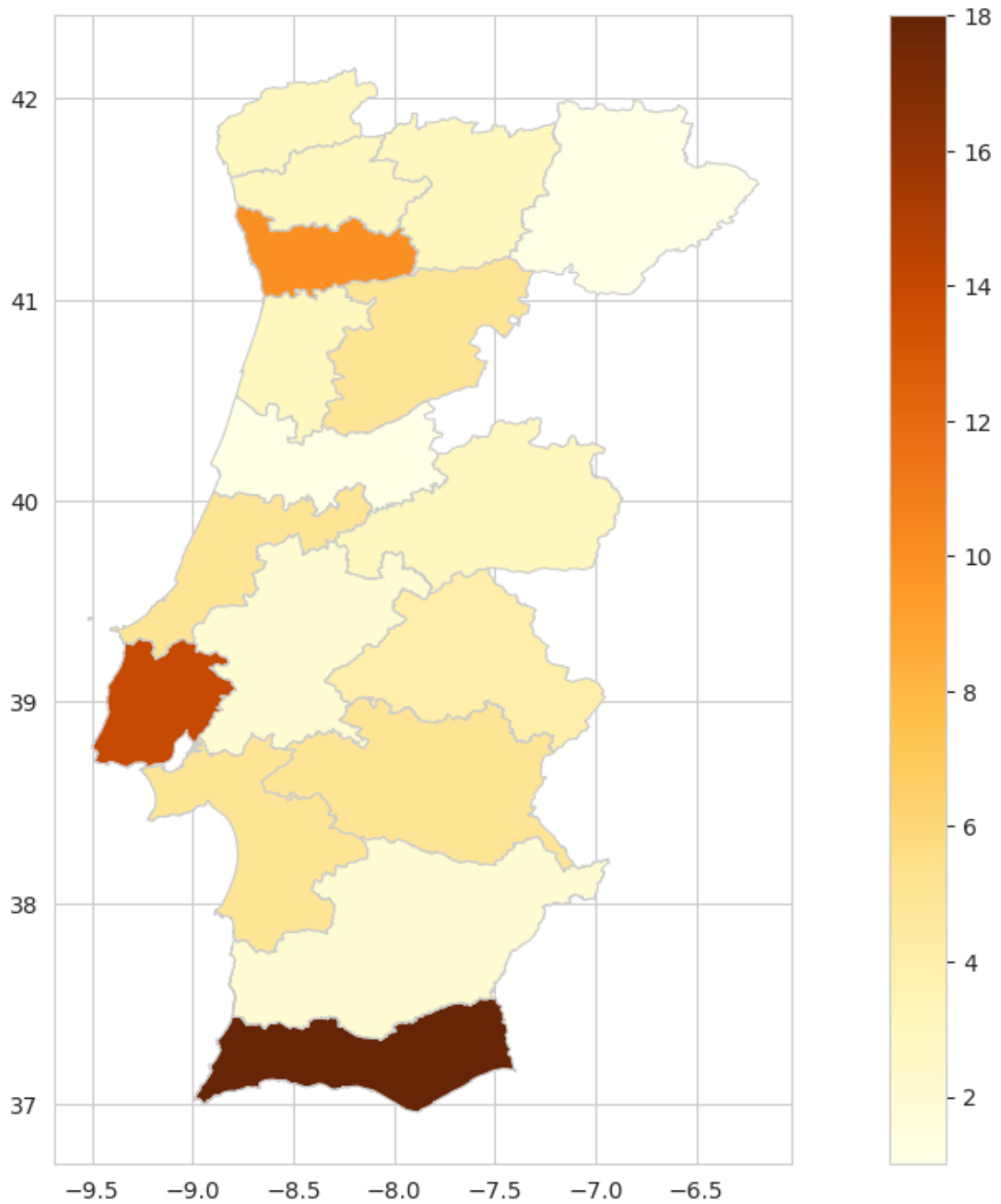
Figura 5.4 - Número de unidades hoteleiras por distrito



De referir, no entanto, a ausência na amostra, de observações de unidades hoteleiras do distrito da Guarda, Ilha da Madeira e Açores.

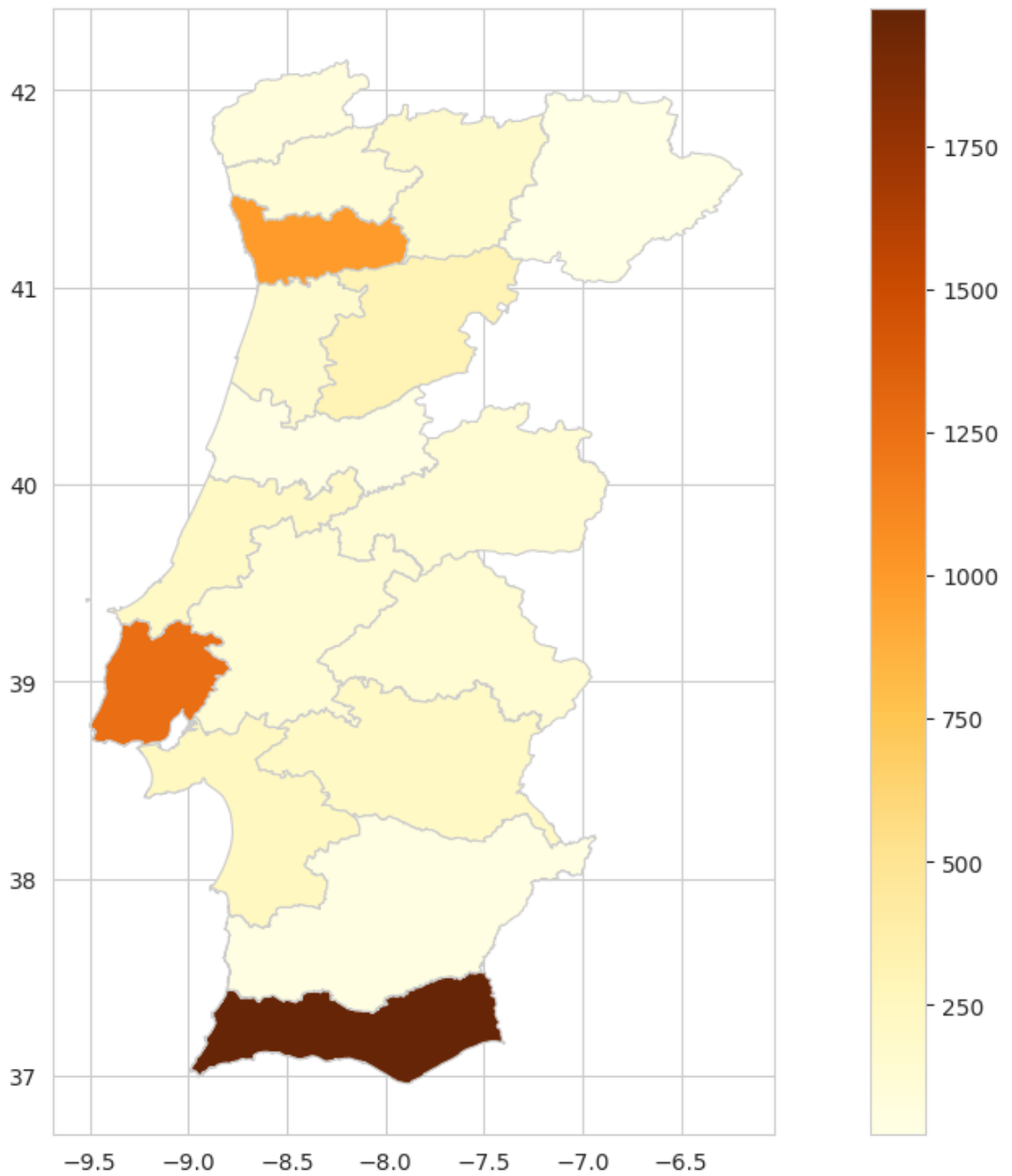
Representadas geograficamente, verifica-se que apesar do destaque dos distritos mencionados anteriormente, a amostra em estudo é mais representativa, em número de unidades hoteleiras, no interior (figura 5.5).

Figura 5.5 - *Heatmap* do número de unidades hoteleiras por distrito



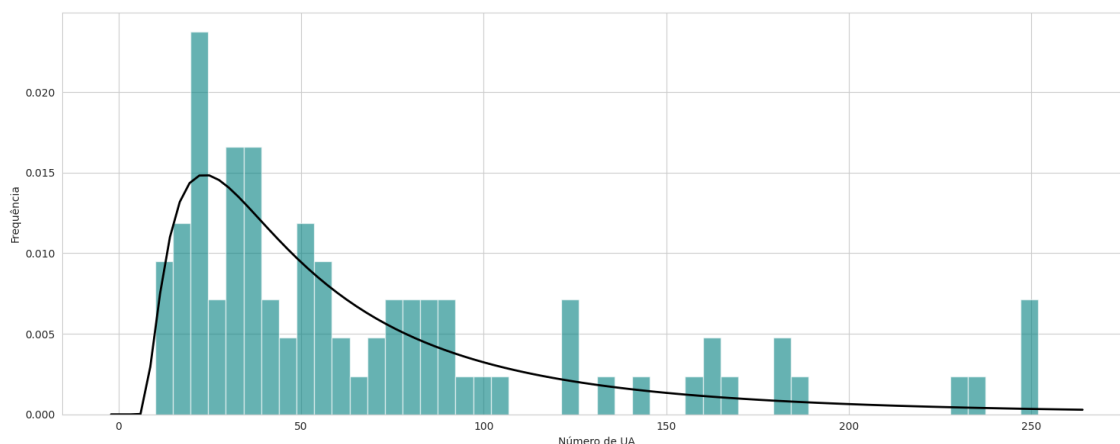
Esta representatividade, perde no entanto força, quando aplicada ao número de unidades de alojamento, em linha com o observado na seção 5.6 para a população em estudo. O facto de o número médios de unidades de alojamento ser inferior no interior assim o proporciona.

Figura 5.6 - Heatmap do número de unidades de alojamento por distrito



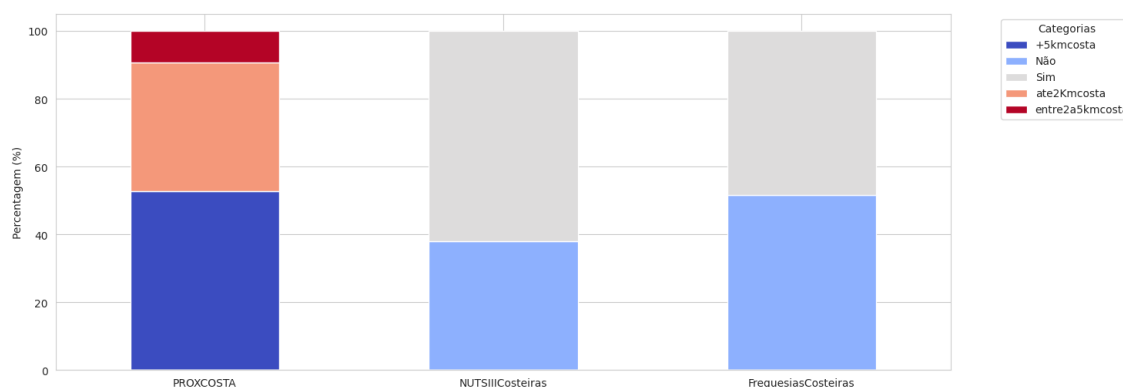
Relativamente ao número de unidades de alojamento, visível na figura 5.7, a amostra apresenta uma maior representatividade no intervalo entre as 25 e as 100 UA, com uma distribuição assimétrica positiva similar à observada na população.

Figura 5.7 - Histograma do número de UA



No que respeita à proximidade costeira, apesar de observada sobre diferentes perspetivas, os resultados observados podem ser considerados pouco divergentes.

Figura 5.8 - Distribuição % das variáveis de proximidade costeira



A variável PROXCOSTA, divide a proximidade à costa em três intervalos: até 2 quilómetros, de 2 a 5 quilómetros e a mais de 5 quilómetros.

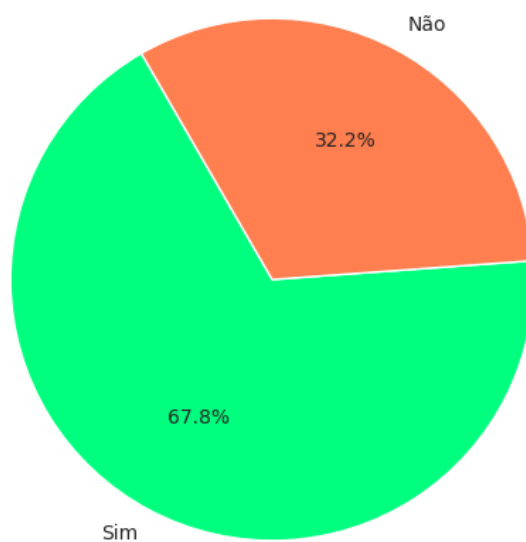
A variável NUTSIIICosteiras e FreguesiasCosteiras, representam unidades que estão (“sim” ou “1”) ou não estão (“não” ou “0”), localizadas em regiões ou freguesias costeiras, isto é que façam fronteira com a costa.

Observada a figura 5.8, e no que respeita a freguesias costeiras, é possível verificar uma distribuição praticamente igual entre observações, com 51,7% das unidades hoteleiras situadas em freguesias costeiras e 48,3% não situadas. Em relação às regiões NUTSIII, estes valores alteram-se ligeiramente com uma maior proporção para regiões costeiras (62,0%), resultado da geografia nacional e dos vastos quilómetros de costa.

No que respeita aos intervalos da variável PROXCOSTA, verifica-se que 52,9% das unidades observadas situam-se a mais de 5 quilómetros da costa, com 37,9% até 2 quilómetros e 9,2% no intervalo entre 2 e 5 quilómetros.

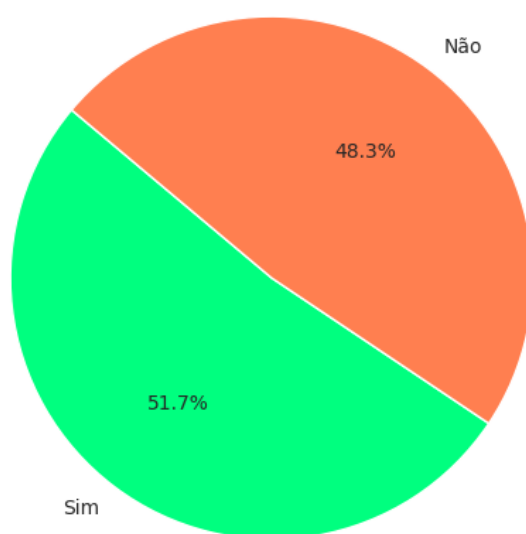
Relativamente à utilização do normativo USALI, 67,8% das unidades hoteleiras da amostra (59), responderam pela utilização, total ou parcialmente, do USALI nas suas contas.

Figura 5.9 - Utilização do USALI



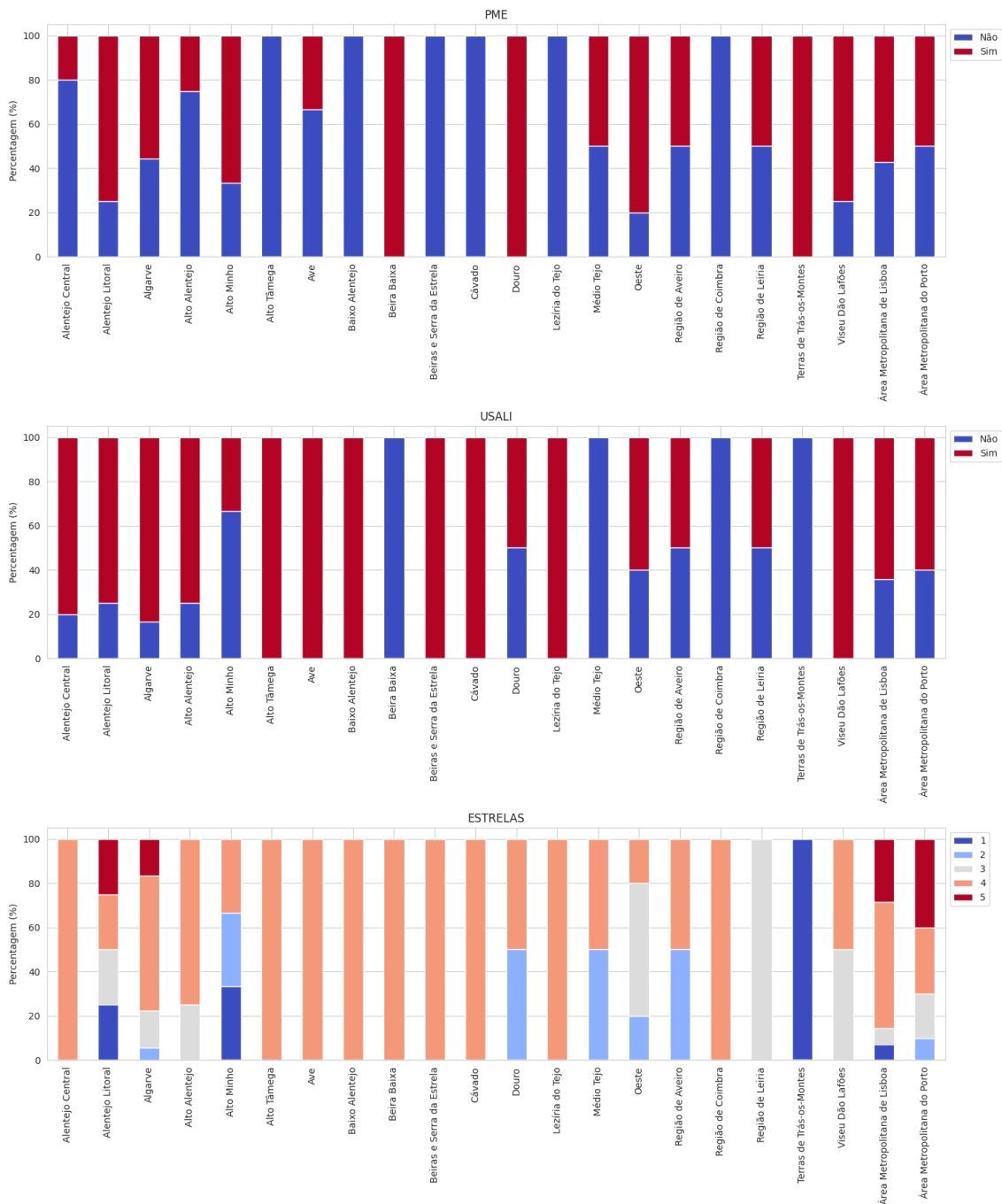
Por fim, considerando a empresa, cerca de 51,7% das unidades hoteleiras respondentes (45), estava integrada em empresas consideradas PME.

Figura 5.10 - Dimensão das empresas - PME



Aprofundando a análise ao distrito, a figura 5.11 apresenta uma análise percentual por distrito de algumas das características anteriormente observadas.

Figura 5.11 - Quadro comparativo por distrito

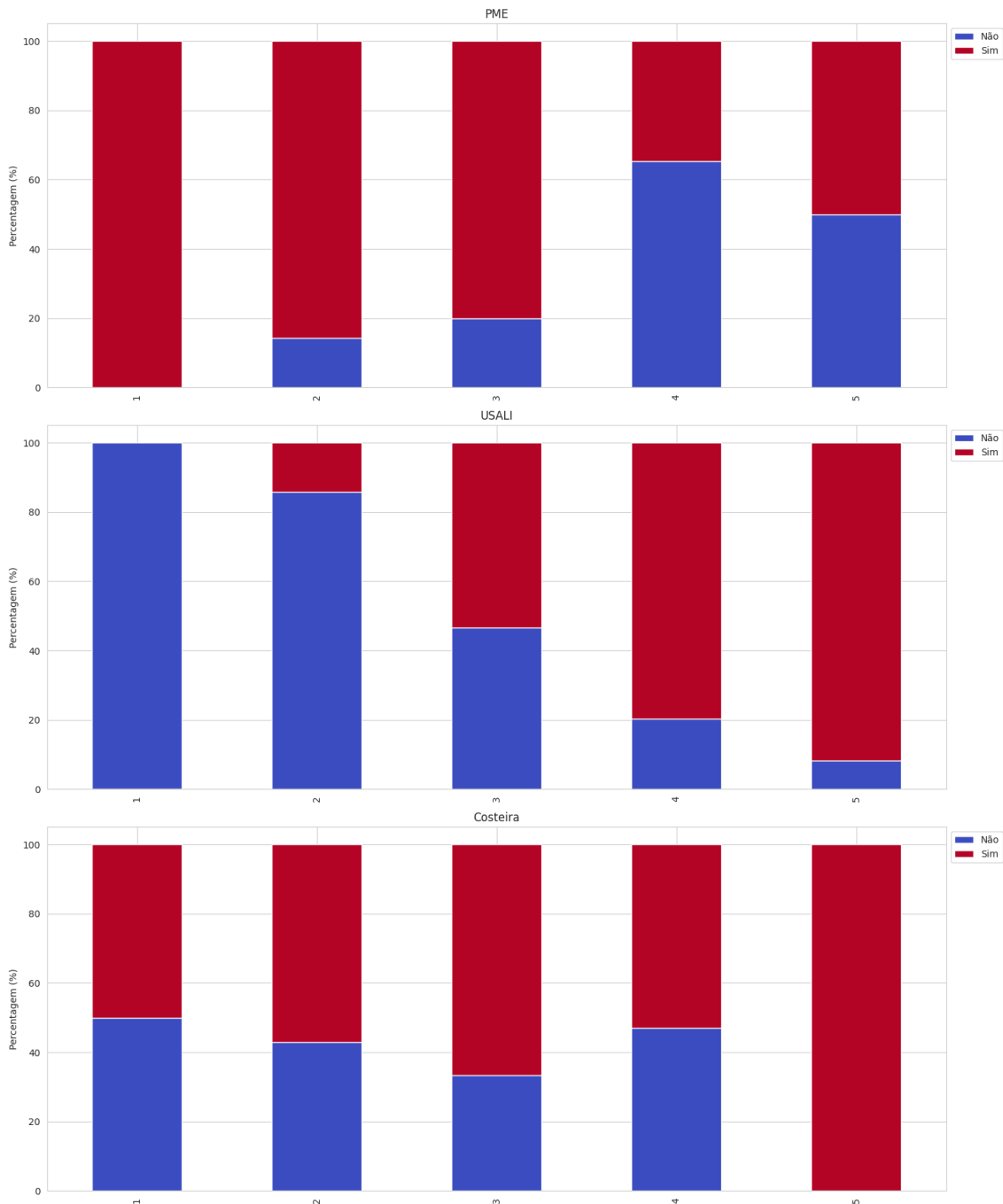


No entanto e ao contrário da figura anterior, não foram observadas tendências particularmente relevantes.

Considerando as características quanto avaliadas as diferentes categorias (estrelas), conforme figura 5.12 apresentada abaixo, é possível observar algumas particularidades da amostra.

As grandes empresas, isto é, empresas não PME, são mais representativas nas categorias mais altas. Do mesmo modo, mas de forma mais evidente, quanto maior a categoria da unidade hoteleira, maior a utilização do USALI.

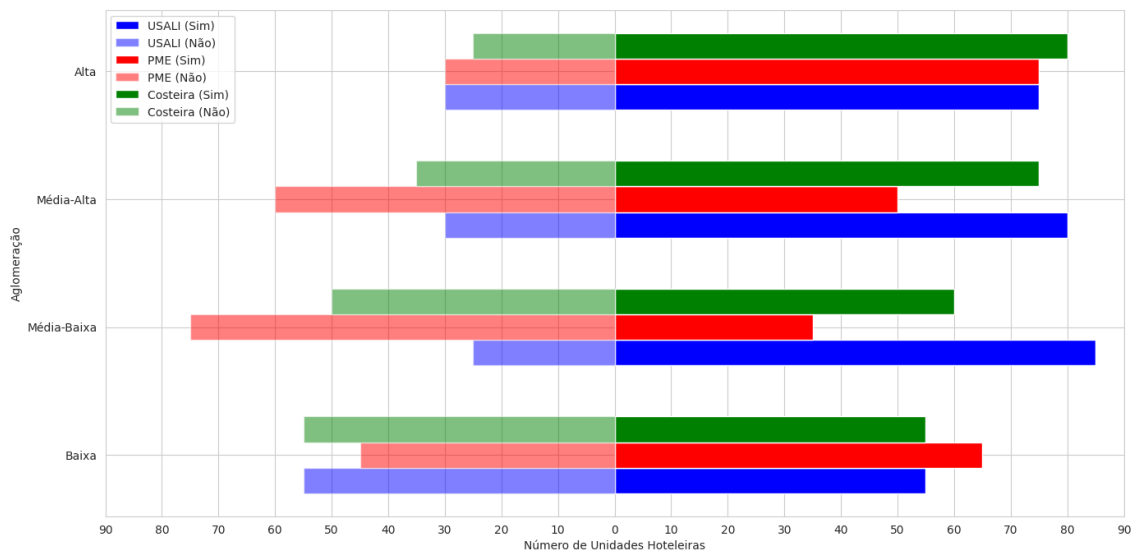
Figura 5.12 - Quadro comparativo por categoria (estrelas)



De referir que todas as unidades de 5 estrelas observadas, tendo por base a variável NUTSIII como indicador de proximidade costeira, situam-se nestas regiões, ao contrário as classificações anteriores, em que a distribuição é similar.

Por fim, e no que à aglomeração de unidades hoteleiras respeita, desagregando esta variável em quatro categorias, nomeadamente, Alta, Média-Alta, Média-Baixa e Baixa, observou-se a distribuição da aplicação do USALI, dimensão da empresa (PME) e proximidade costeira (NUTSIII). No nível alto, existe uma distribuição similar entre todas as variáveis observadas com maior peso do “Sim”. Com exceção da dimensão da empresa, esta distribuição sugere uma convergência para o “Não” com a redução do nível de aglomeração.

Figura 5.13 - Distribuição por nível de aglomeração



5.2. Pressupostos dos modelos de regressão linear múltipla

A implementação de uma análise de dados, com recurso a técnicas estatísticas de regressão linear múltipla, neste caso através do método dos Mínimos Quadrados Ordinários (OLS), requer a consideração dos pressupostos básicos que asseguram a fiabilidade e a validade das inferências obtidas.

Estes pressupostos contribuem para a robustez dos resultados e são imprescindíveis para a correta interpretação das relações estabelecidas entre as variáveis dependentes e independentes. A seguir, enunciam-se e explicam-se os principais pressupostos da regressão linear múltipla.

5.2.1 Linearidade

A relação entre a variável dependente e as variáveis independentes deve ser linear. Este pressuposto pode ser verificado por via de gráficos de dispersão ou testes de não-linearidade.

5.2.2 Autocorrelação

Os resíduos associados a cada observação não devem correlacionado. Este pressuposto é particularmente relevante em séries temporais ou dados em painel, como é o caso da amostra em estudo. Para testar a autocorrelação podemos utilizar testes estatísticos como o teste de Durbin-Watson ou o teste de Ljung-Box.

Teste de Durbin-Watson:

- Hipótese Nula (H_0): Não há autocorrelação nos resíduos;
- Intervalo: O valor da estatística varia entre 0 e 4, onde um valor próximo de 2 sugere ausência de autocorrelação;

Teste de Ljung-Box:

- Hipótese Nula (H_0): Os dados são independentes ao longo do atraso (lag), ou seja, não há autocorrelação nos dados.
- Intervalo: Gera uma estatística de teste e um valor-p para decidir se deve rejeitar a hipótese nula.

5.2.3 Homocedasticidade

A variância dos resíduos deve ser constante em todos os níveis das variáveis independentes. A heterocedasticidade pode ser diagnosticada por meio de gráficos de

resíduos ou por meio do recurso a testes formais como o teste de Breusch-Pagan ou o teste de White. Em ambos os casos temos como hipóteses:

H₀: Homoscedasticidade dos resíduos (ausência de heteroscedasticidade)

H₁: Heteroscedasticidade dos resíduos (presença de heteroscedasticidade)

5.2.4 Normalidade

Para a realização de inferências estatísticas é assumido que os resíduos seguem uma distribuição normal. Para o testar podem ser utilizadas apresentações gráficas ou executados testes como o de Shapiro-Wilk. Este último é um teste estatístico amplamente utilizado para testar a normalidade de uma amostra de dados. Apresenta como hipótese a testar:

H₀: A amostra segue uma distribuição normal

H₁: A amostra não segue uma distribuição normal

5.2.5 Multicolinearidade

As variáveis independentes não devem ser significativamente correlacionadas entre si. A multicolinearidade pode ser detetada através de métodos como o Fator de Inflação da Variância (VIF). Esperam-se valores abaixo de 5 para validar que as variáveis são independentes entre si.

5.3. Resultados dos modelos de regressão linear múltipla

5.3.1 Análise inicial e validação dos pressupostos

Procedeu-se com a validação dos pressupostos do modelo de regressão linear múltipla. Para o efeito foi subdividido o *dataset* em dois subconjuntos, cada um com uma variável dependente (RevPAR e GOPPAR), aos quais foram testados os pressupostos.

Em ambos os casos, os resultados apontaram para distribuições que sugerem relações lineares, conforme figuras abaixo.

Figura 5.14 - Mapas de dispersão das variáveis independentes (amostra RevPAR)

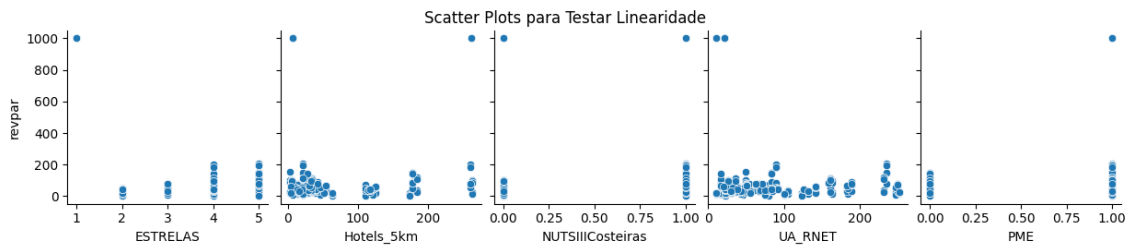
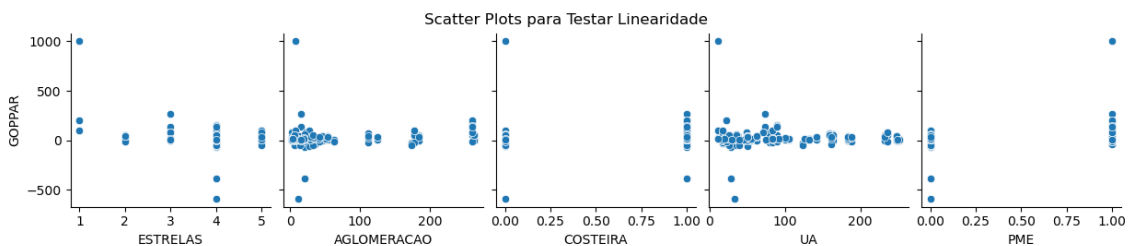


Figura 5.15 - Mapas de dispersão das variáveis independentes (Amostra GOPPAR)



No que respeita à correlação dos resíduos, tanto a apresentação gráfica como os testes desenvolvidos, apontam para a inexistência de correlação entre resíduos. A hipótese nula foi confirmada tanto para a amostra RevPAR como para a amostra GOPPAR ($p=0,6555$ e $p=0,2984$ respetivamente).

Do mesmo modo, o teste de Durbin-Watson, produziu como resultados, os valores de 2.057 e 1,86, valores próximos de 2. Confirma-se que os resíduos são independentes entre si.

Quanto ao teste da homocedasticidade dos resíduos, foram executados os testes de White e de Breusch-Pagan. Em ambos os casos, os resultados obtidos evidenciam que os resíduos são heterocedásticos. Os valores p obtidos, nomeadamente, $6.345e-24$ (White) e $2.043e-06$ (Breusch-Pagan) para o RevPAR e $2.828e-5$ (White) e 0.02689 (Breusch-Pagan) para o GOPPAR, rejeitam as hipóteses nulas.

Relativamente à normalidade dos resíduos, apesar dos gráficos produzidos em ambas as amostras, apresentarem uma distribuição aparentemente regular, o teste de

Shapiro-Wilk rejeitou a hipótese em ambas as amostras, concluindo pela anormalidade da distribuição dos resíduos.

No que respeita à multicolineariedade, os VIF para todas as variáveis independentes, estão abaixo de 5, sugerindo que não existia uma relação entre as variáveis independentes, e como tal não existia um problema de multicolineariedade.

Atendendo a que os pressupostos não se encontravam integralmente cumpridos, em particular quanto à normalidade e homocedasticidade dos resíduos, a validade das inferências realizadas com base num modelo de regressão linear múltipla, poderiam ficar comprometidas.

Nesse sentido optou-se pela adoção de um modelo robusto HC3 (*Heteroscedasticity-Consistent Covariance Matrix Estimator*), uma técnica robusta que ajusta os erros padrão dos coeficientes estimados no modelo OLS.

Assim a execução do modelo OLS Robusto foi aplicada a 4 amostras, representativa dos dados recolhidos relativamente a dois períodos (pré COVID 2017-19 e durante COVID 2020-21) e duas medidas de performance, nomeadamente GOPPAR e RevPAR.

O quadro seguinte, apresenta algumas estatísticas descritivas relativas às quatro amostras.

Tabela 5.1 - Estatísticas descritivas dos modelos

GOPPAR_preCOVID						
	ESTRELAS	AGLOMERACAO	PME	COSTEIRA	UA	GOPPAR
mean	3,93	51,39	0,38	0,65	86,69	33,45
min	1,00	3,00	0,00	0,00	11,00	-52,90
max	5,00	263,00	1,00	1,00	252,00	200,00
std	0,73	73,49	0,49	0,48	71,02	34,05

GOPPAR_duranteCOVID						
	ESTRELAS	AGLOMERACAO	PME	COSTEIRA	UA	GOPPAR
mean	3,91	46,61	0,41	0,67	84,04	12,23
min	1,00	3,00	0,00	0,00	11,00	-593,27
max	5,00	263,00	1,00	1,00	252,00	1000,00
std	0,80	67,56	0,49	0,47	69,19	123,22

REVPAR_preCOVID						
	ESTRELAS	AGLOMERACAO	PME	COSTEIRA	UA	REVPAR
mean	3,87	51,30	0,43	0,65	82,00	73,41
min	1,00	3,00	0,00	0,00	11,00	2,53
max	5,00	263,00	1,00	1,00	252,00	1000,00
std	0,73	70,80	0,50	0,48	67,70	106,72

REVPAR_duranteCOVID						
	ESTRELAS	AGLOMERACAO	PME	COSTEIRA	UA	REVPAR
mean	3,87	46,11	0,46	0,66	79,59	57,43
min	1,00	3,00	0,00	0,00	11,00	1,80
max	5,00	263,00	1,00	1,00	252,00	1000,00
std	0,80	65,21	0,50	0,48	66,08	122,65

Considerando em primeiro lugar as variáveis dependentes, destaca-se a queda significativa, tanto do RevPAR como do GOPPAR no durante o COVID.

No entanto, ao analisarmos a média destas variáveis nas unidades hoteleiras costeiras e não costeiras (tabela 5.2), verificamos que a quebra no GOPPAR foi menos acentuada que nas unidades hoteleiras não costeiras, com as primeiras a recuarem 4,79 euros e as segundas 29,92 euros.

Tabela 5.2 - Médias de performance - costeiras vs não costeiras

	GOPPAR Pre-COVID	GOPPAR Durante COVID	REVPAR Pre-COVID	REVPAR Durante COVID
COSTEIRA				
Não	19,60	14,81	47,04	81,25
Sim	40,86	10,94	87,88	45,24

Curiosamente, os dados recolhidos, apontam no que ao RevPAR respeita, para um comportamento inverso. As unidades hoteleiras costeiras apresentaram reduziram em cerca de 50% o seu RevPAR, enquanto por sua vez as unidades hoteleiras não costeiras cresceram quase 97%.

Ainda relativamente às médias do GOPPAR e RevPAR, após uma agregação baseada nos percentis, nas unidades hoteleiras em localizações geográficas com baixa aglomeração, verifica-se um crescimento do GOPPAR dos 21,03 euros para os 48,23

euros, ao contrário das restantes localizações onde a perda é significativa, com destaque para as localizações de alta aglomeração.

Do mesmo modo, verifica-se na tabela abaixo, que uma vez mais, o RevPAR cresceu nas localizações com baixa aglomeração e recuou significativamente nas restantes localizações. Do mesmo modo, nas unidades localizações com alta aglomeração, a redução é mais expressiva.

Tabela 5.3 - Médias de performance - aglomeração

	GOPPAR Pre-COVID	GOPPAR Durante COVID	REVPAR Pre-COVID	REVPAR Durante COVID
Aglomeração				
Alta Aglomeração	34,65	2,92	76,57	47,31
Baixa Aglomeração	21,03	48,23	47,86	99,43
Muito Alta Aglomeração	60,92	12,69	117,77	35,72
Média Aglomeração	20,03	-14,90	53,72	39,81

5.3.2 Resultados - Amostra GOPPAR Pré COVID

Nesta amostra, o coeficiente de determinação R^2 é de 0,502, indicando que aproximadamente 50,2% da variabilidade observada na variável dependente é explicada pelo modelo.

Figura 5.16 - Resultados GOPPAR pré COVID

```

=====
                        OLS Regression Results
=====
Dep. Variable:                GOPPAR    R-squared:                0.502
Model:                        OLS      Adj. R-squared:           0.485
Method:                       Least Squares    F-statistic:              9.670
Date:                         Fri, 01 Sep 2023    Prob (F-statistic):       5.41e-08
Time:                         13:32:14      Log-Likelihood:           -698.44
No. Observations:             152        AIC:                     1409.
Df Residuals:                 146        BIC:                     1427.
Df Model:                      5
Covariance Type:              HC3
=====

```

	coef	std err	z	P> z	[0.025	0.975]
Intercept	41.1274	22.882	1.797	0.072	-3.721	85.976
ESTRELAS	-7.1373	6.075	-1.175	0.240	-19.045	4.770
AGLOMERACAO	0.2953	0.051	5.743	0.000	0.195	0.396
PME	-4.5066	5.078	-0.887	0.375	-14.459	5.446
COSTEIRA	17.5924	3.762	4.677	0.000	10.219	24.965
UA	-0.0521	0.034	-1.539	0.124	-0.118	0.014

```

=====
Omnibus:                      25.310    Durbin-Watson:            1.719
Prob(Omnibus):                 0.000    Jarque-Bera (JB):         163.580
Skew:                          0.137    Prob(JB):                  3.01e-36
Kurtosis:                      8.075    Cond. No.                  881.
=====

```

Notes:

[1] Standard Errors are heteroscedasticity robust (HC3)

Adicionalmente, o R^2 ajustado, que penaliza o modelo por incluir variáveis não explicativas, é de 0,485.

Quanto à significância global do modelo, o valor da estatística F é de 9,670, com um valor p de 5,41e-08, o que sugere que o modelo é estatisticamente significativo. O valor Durbin-Watson de 1,719 indica que a autocorrelação dos resíduos não é uma questão significativa neste modelo. Porém, os testes de Omnibus e Jarque-Bera indicam a rejeição da hipótese de normalidade dos resíduos, com valores p muito próximos de zero o que poderá limitar as inferências.

Não obstante, em relação aos coeficientes individuais, o coeficiente para a variável "ESTRELAS" é de -7,1373, mas com um valor p 0,240, o que sugere que esta variável não é estatisticamente significativa. Em contraste, as variáveis "AGLOMERAÇÃO" e "COSTEIRA" têm coeficientes de 0,2953 e 17,5924, respectivamente (com contributo positivo), sendo ambos com um valor p próximo de zero, indicando alta significância estatística. As variáveis "PME" e "UA" apresentam valores p 0,375 e 0,124, respectivamente, indicando que não são estatisticamente significativos ou são apenas marginalmente significativos.

5.3.3 Resultados - Amostra GOPPAR durante COVID

Esta amostra revelou várias limitações do modelo, que questionam a sua aplicabilidade no estudo do impacto das diversas variáveis sobre a variável dependente GOPPAR. O coeficiente de determinação R^2 de 0,154 e o R^2 ajustado de 0,113 apontam para uma explicação insuficiente da variação na variável dependente pelo modelo em questão. Adicionalmente, a estatística F de 1,289 com um valor p de 0,274, sugere que o modelo global não é estatisticamente significativo ao nível convencional de 0,05.

No que concerne aos coeficientes individuais das variáveis, nenhum deles revelou ser estatisticamente significativo. A análise dos diagnósticos residuais, incluindo o valor de Durbin-Watson de 2,156, indica ausência de autocorrelação nos resíduos, uma característica desejável. Contudo, os testes de Omnibus e Jarque-Bera confirmam a violação da suposição de normalidade dos resíduos.

Figura 5.17 - Resultados GOPPAR durante COVID

OLS Regression Results						
Dep. Variable:	GOPPAR		R-squared:	0.154		
Model:	OLS		Adj. R-squared:	0.113		
Method:	Least Squares		F-statistic:	1.289		
Date:	Fri, 01 Sep 2023		Prob (F-statistic):	0.274		
Time:	13:32:14		Log-Likelihood:	-682.09		
No. Observations:	111		AIC:	1376.		
Df Residuals:	105		BIC:	1392.		
Df Model:	5					
Covariance Type:	HC3					
	coef	std err	z	P> z	[0.025	0.975]
Intercept	212.9930	174.035	1.224	0.221	-128.110	554.096
ESTRELAS	-59.0737	42.304	-1.396	0.163	-141.988	23.841
AGLOMERACAO	0.0167	0.076	0.220	0.826	-0.132	0.166
PME	20.0052	14.298	1.399	0.162	-8.019	48.029
COSTEIRA	12.2248	23.734	0.515	0.606	-34.293	58.742
UA	0.1546	0.165	0.936	0.349	-0.169	0.478
Omnibus:	91.827		Durbin-Watson:	2.156		
Prob(Omnibus):	0.000		Jarque-Bera (JB):	4261.752		
Skew:	1.968		Prob(JB):	0.00		
Kurtosis:	33.099		Cond. No.	780.		

Notes:

[1] Standard Errors are heteroscedasticity robust (HC3)

5.3.4 Resultados - Amostra RevPAR pré COVID

O modelo calculado nesta amostra, em que a variável dependente em estudo é o RevPAR, apresenta um coeficiente de determinação de R^2 de 0,255 e um R^2 ajustado de 0,233, indicando que cerca de 25,5% da variação em RevPAR é explicada pelas variáveis independentes do modelo.

Figura 5.18 - Resultados RevPAR pré COVID

OLS Regression Results						
Dep. Variable:	REVPAR	R-squared:	0.255			
Model:	OLS	Adj. R-squared:	0.233			
Method:	Least Squares	F-statistic:	5.828			
Date:	Fri, 01 Sep 2023	Prob (F-statistic):	5.41e-05			
Time:	13:32:14	Log-Likelihood:	-1039.3			
No. Observations:	175	AIC:	2091.			
Df Residuals:	169	BIC:	2110.			
Df Model:	5					
Covariance Type:	HC3					
	coef	std err	z	P> z	[0.025	0.975]
Intercept	175.1308	147.163	1.190	0.234	-113.303	463.565
ESTRELAS	-36.9873	40.183	-0.920	0.357	-115.744	41.769
AGLOMERACAO	0.6320	0.268	2.355	0.019	0.106	1.158
PME	-25.1111	21.340	-1.177	0.239	-66.937	16.715
COSTEIRA	29.6367	12.090	2.451	0.014	5.940	53.333
UA	0.0087	0.100	0.087	0.931	-0.188	0.205
Omnibus:	213.372	Durbin-Watson:	1.595			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	9369.994			
Skew:	4.844	Prob(JB):	0.00			
Kurtosis:	37.513	Cond. No.	870.			

Notes:

[1] Standard Errors are heteroscedasticity robust (HC3)

Neste caso, a estatística F é de 5,828 com um valor p de 5,41e-05, o que sugere que o modelo é estatisticamente significativo ao nível de 0,05. Entre as variáveis independentes, “AGLOMERAÇÃO” e “COSTEIRA” são estatisticamente significativas. O coeficiente de “AGLOMERAÇÃO” é 0,6320, indicando um impacto positivo na RevPAR. Similarmente, “AGLOMERAÇÃO” tem um coeficiente de 29,6367, também sugerindo um impacto positivo.

No que respeita aos diagnósticos dos pressupostos, o valor de Durbin-Watson de 1,595 sugere a possibilidade de autocorrelação residual. Por sua vez, os testes de Omnibus e Jarque-Bera apontam novamente para a rejeição da hipótese de normalidade dos resíduos, uma questão que pode constituir uma limitação em termos de inferência estatística.

5.3.5 Resultados - Amostra RevPAR durante COVID

Tal como sucedeu com o GOPPAR, este modelo apresenta características que limitam a sua utilidade explicativa e inferencial. O coeficiente de determinação R^2 é de

apenas 0,156 e o R^2 ajustado é ainda mais baixo, 0,121, indicando uma explicação muito limitada da variação da variável dependente RevPAR.

Figura 5.19 - Resultados RevPAR durante COVID

OLS Regression Results						
Dep. Variable:	REVPAR	R-squared:	0.156			
Model:	OLS	Adj. R-squared:	0.121			
Method:	Least Squares	F-statistic:	0.4244			
Date:	Fri, 01 Sep 2023	Prob (F-statistic):	0.831			
Time:	13:32:14	Log-Likelihood:	-779.74			
No. Observations:	127	AIC:	1571.			
Df Residuals:	121	BIC:	1589.			
Df Model:	5					
Covariance Type:	HC3					
	coef	std err	z	P> z	[0.025	0.975]
Intercept	306.3619	202.146	1.516	0.130	-89.838	702.562
ESTRELAS	-62.7443	49.195	-1.275	0.202	-159.165	33.676
AGLOMERACAO	-0.0550	0.098	-0.561	0.575	-0.247	0.137
PME	-6.8034	16.255	-0.419	0.676	-38.663	25.056
COSTEIRA	-15.8440	18.843	-0.841	0.400	-52.775	21.087
UA	0.1290	0.194	0.663	0.507	-0.252	0.510
Omnibus:	164.322	Durbin-Watson:	2.381			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	5505.785			
Skew:	4.860	Prob(JB):	0.00			
Kurtosis:	33.757	Cond. No.	748.			

Notes:

[1] Standard Errors are heteroscedasticity robust (HC3)

A estatística F tem um valor de 0,4244 e um valor p associado de 0,831, o que claramente indica que o modelo não é estatisticamente significativo ao nível de 0,05. Em relação aos diagnósticos residuais, o valor de Durbin-Watson é de 2,381, o que sugere a ausência de autocorrelação nos resíduos. No entanto, os testes de Omnibus e Jarque-Bera com valores p muito próximos de zero confirmam que os resíduos não seguem uma distribuição normal, comprometendo a inferência estatística.

5.3.6 Resultados - Comparativo

A utilização de quatro subamostras diferentes, em particular quanto ao período e variável dependente, permitiu obter e comparar os diferentes resultados (Tabela 5.4). O modelo GOPPAR Pré COVID é o que apresenta um coeficiente de determinação mais elevado, com o modelo a explicar 50,19% das variações no GOPPAR. Este valor contrasta com os coeficientes R^2 mais baixos observados nos outros modelos, que variam entre 15,4% e 25,6%, denotando uma capacidade explicativa inferior.

Tabela 5.4 - Quadro comparativo dos modelos

Model	R ²	Coef. ESTRELAS	P-value ESTRELAS	Coef. AGLOMERACAO	P-value AGLOMERACAO
GOPPAR Pre-COVID	50,19%	-7,14	0,240072	0,30	9,28E-09
GOPPAR During COVID	15,37%	-59,07	0,162591	0,02	8,26E-01
REVPAR Pre-COVID	25,53%	-36,99	0,357323	0,63	1,85E-02
REVPAR During COVID	15,58%	-62,74	0,202161	-0,05	5,75E-01

Model	Coef. PME	P-value PME	Coef. COSTEIRA	P-value COSTEIRA	Coef. UA	P-value UA
GOPPAR Pre-	-4,51	0,374822	17,59	0,000003	-0,05	0,123824
GOPPAR During COVID	20,01	0,161774	12,22	0,606497	0,15	0,349137
REVPAR Pre-COVID	-25,11	0,239316	29,64	0,014235	0,01	0,930999
REVPAR During COVID	-6,80	0,675553	-15,84	0,400424	0,13	0,507208

Em relação ao coeficiente associado à variável "ESTRELAS", observa-se que este é negativo em todos os modelos, mas com um nível de significância estatística elevado, superior a 0,05, o que sugere que a categorização por estrelas não é um preditor significativo das métricas em estudo.

Por outro lado, o coeficiente para a variável "AGLOMERAÇÃO" é positivo e altamente significativo no modelo GOPPAR pré COVID, com um valor p inferior a 0,001, indicando uma correlação positiva forte. Contudo, este coeficiente torna-se menos significativo durante o período da COVID-19.

No que concerne à variável "PME", o coeficiente altera-se de positivo para negativo entre os períodos pré COVID e durante a COVID, mas o nível de significância também varia, tornando-se mais significativo no modelo GOPPAR durante a pandemia, com um valor p de 0,162, embora nunca revele significância estatística.

A variável "COSTEIRA" apresenta um coeficiente positivo e altamente significativo no modelo GOPPAR pré COVID, mas este coeficiente perde a sua significância nos outros modelos.

Finalmente, a variável "UA" demonstra um nível de significância estatística geralmente elevado, exceto no modelo GOPPAR pré COVID, onde é negativo e significativo com um valor p de 0,124.

5.4. Discussão dos resultados

A utilização de quatro subamostras diferentes, permitiu nesta dissertação, uma análise simultânea ao objeto de estudo em contextos distintos.

Inicia-se esta discussão por referir que 51,7% das unidades hoteleiras analisadas integram empresas qualificadas como PME. Estas unidades são maioritariamente unidades hoteleiras de 4 estrelas, sendo que os dados demonstram uma forte concentração nas áreas metropolitanas de Lisboa, Porto e Algarve, destinos turísticos consolidados.

A percentagem de unidades hoteleira que indicaram utilizar o USALI (67,8%), sugere um crescimento da adoção deste normativo face ao estudo de Faria *et al.*, (2018). Este número é particularmente expressivo quanto maior a categoria da unidade. Por outras palavras, as unidades de quatro e cinco estrelas mostram uma percentagem de adoção do USALI superior às unidades de três, duas ou uma estrela.

Estes resultados, podem ser justificados pelos custos de contexto que uma implementação do normativo pode implicar em unidades de categorias mais baixas, com a necessidade de um sistema de registo analítico, segmentado por departamentos (centros de custo) e equipas financeiras dotadas de meios e conhecimento para tal. Algo que não será compatível se a contabilidade de gestão de uma unidade hoteleira for desenvolvida pelo contabilista certificado externo, não especializado, solução económica e simplificada.

Verificou-se ainda uma adoção do USALI nas unidades localizadas em zonas de maior aglomeração.

Apesar de não ser objetivo principal desta dissertação, evidencia-se também o impacto do período COVID, compreendido neste estudo pelos anos de 2020 e 2021, que provocou ao longo deste período, um comportamento distinto nas variáveis dependentes em estudo. Assim, os dados recolhidos evidenciam uma quebra na performance entre o período pré COVID e durante COVID, em que o RevPAR médio recua dos 73,41 euros para os 57,43 euros e o GOPPAR médio dos 33,35 euros para os 12,23 euros (tabela 5.1).

No entanto, esta quebra não é igual em todo o território. Nos territórios costeiros, onde a aglomeração é mais alta, verifica-se uma quebra mais acentuada tanto no GOPPAR como no RevPAR, em contraste com as localizações com menor aglomeração, e, portanto, mais isolados, tradicionalmente não costeiros.

Tendo por base a tabela 5.2, pode-se verificar que o RevPAR médio, apresentou um crescimento relativamente ao período pré COVID, de 47,04 euros para os 81,25 euros, quando consideradas as unidades hoteleiras não costeiras.

Em linha com estes resultados, o RevPAR médio das unidades em localizações de baixa aglomeração, apresenta uma evolução dos 47,86 euros para os 99,43 euros. No entanto, esta variação é reforçada pela própria evolução do GOPPAR, com um crescimento dos 21,03 euros de média (pré COVID) para os 48,23 euros de média (durante COVID), dados ilustrados na figura 5.3.

Este comportamento divergente pode ser justificado pela excecionalidade da pandemia, que trouxe períodos com restrições severas à atividade, legislação reativa e um sem número de condicionalismos que limitaram os tradicionais destinos turístico, concentrados em zonas costeiras e com níveis de elevada aglomeração. Em oposição, o interior, com uma baixa aglomeração, permitia à época, o necessário distanciamento social que induzia segurança, e poderá ter provocado uma alteração no perfil do hóspede.

Por tudo o mencionado anteriormente, os resultados obtidos relativamente a este período conduziram a modelos não explicativos e que decidiu-se desconsiderar.

Dito isto, o estudo revela que a rendibilidade operacional dos hotéis em Portugal é melhor explicada num modelo centrado no GOPPAR considerando um período de estabilidade, ou seja, o período pré COVID.

Esta investigação aponta para que a rendibilidade operacional destes hotéis seja explicada em 50% pelos determinantes em estudo.

Relativamente a estes, o modelo apresentado valida a hipótese 1.1, confirmando uma relação positiva da localização costeira com a rendibilidade. O coeficiente positivo e estatisticamente significativo desta variável suporta fortemente a hipótese 1, em linha, com as conclusões de Menicucci (2018). Apesar da autora ter optado por uma adaptação do GOPPAR enquanto variável dependente, o seu estudo focou-se em Itália, um país mediterrâneo com similaridades a Portugal.

Ainda subordinado à localização, a hipótese 1.2 estabelece uma relação positiva entre aglomeração e rendibilidade. O modelo, apesar de apresentar um coeficiente mais baixo quando comparado com a localização costeira, apresenta um valor positivo com uma expressiva significância. Nesse sentido, a hipótese é fortemente aceite, suportando as evidências encontradas em estudos anteriores (Lado-Sestayo *et al.*, 2016), que estabelecem relações entre o destino e a procura e a performance individual,

particularmente o estudo de Travanca *et al.* (2022), cuja investigação incidiu sobre a região do Algarve, caracterizada por uma elevada aglomeração (figura 5.5.).

Em contrapartida, quanto à dimensão da unidade, medida em número de unidades de alojamento, observa-se que o coeficiente estimado é - 0,1546 com um valor *p* de 0,349. Isto sugere que a dimensão da unidade hoteleira não tem influência na sua performance. Portanto, a hipótese 2, que postulava uma relação positiva entre a dimensão e a performance das unidades hoteleiras, não é estatisticamente confirmada, contrariando a maioria dos estudos existentes (Claver-Cortés *et al.*, 2007; Dimitrić *et al.*, 2019; Kim *et al.*, 2013; Lado-Sestayo *et al.*, 2016, 2020; Lado-Sestayo & Vivel-Búa, 2018a, 2018b; Menicucci, 2018; Romero *et al.*, 2020; Travanca *et al.*, 2022).

Em sentido contrário, da literatura analisada, apenas Bresciani *et al.* (2015) encontraram evidências claras da inexistência de uma relação entre a dimensão e a performance, pelo que se reforçam estas conclusões.

Uma justificação para os resultados obtidos, e contrários à maioria dos estudos referenciados, poderá estar na multiplicidade de abordagens encontradas. *Exempli gratia*, uma unidade com um maior número de quartos tem um maior potencial de negócio, mas incorpora um volume superior de custos particularmente fixos, o que invariavelmente condicionará a rentabilidade.

Do mesmo modo, a categoria do hotel, medida pelo número de estrelas, aponta para uma influência negativa na performance, com um coeficiente de -7,1373 e um valor *p* de 0,250, pelo que se considerou este resultado estatisticamente insignificante. Tal confirma a hipótese 3, que sugeria que a categoria do hotel não teria influência na sua performance, confirmando o grupo de estudos que suportam esta evidência (Lado-Sestayo *et al.*, 2020; Lado-Sestayo & Vivel-Búa, 2018a; Oliveira *et al.*, 2013).

Uma unidade de categoria superior terá, teoricamente, um serviço e perceção de mercado, superiores, o que conduzirá a um valor de venda mais elevado. No entanto, a categoria superior obriga a um conjunto de serviços e infraestruturas a disponibilizar ao hóspede, que nem sempre se convertem em vendas e conseqüentemente, incrementam custos. São uma consequência necessária da categoria superior, que na ausência de procura (sazonalidade), traduzem-se em custos sem vendas e conseqüentemente na quebra da rentabilidade.

Por fim, a hipótese 4 apontava a inexistência de uma relação entre a classificação como PME e a rentabilidade de uma unidade hoteleira. Na realidade, observa-se uma influência negativa na rentabilidade das Unidades Hoteleiras que, contudo, não é

estatisticamente significativa a um nível de 5% de significância (valor $p = 0,375$). Assim, a hipótese quatro, que defendia a ausência de influência da classificação como PME na performance é confirmada, suportando o equilíbrio entre as vantagens e desvantagens apontado por Panno (2019) e Alvarez *et al.* (2021).

5.5. Síntese conclusiva

A análise preliminar dos pressupostos dos modelos de regressão linear múltipla, encontrou nas diferentes amostras algumas limitações, ultrapassadas pela utilização de técnicas estatísticas que visavam a sua correção.

Consequentemente, foram testados os quatro modelos de regressão estatística que apontaram diferentes níveis de significância, o que conduziu ao abandono dos modelos que compunham a variável dependente RevPAR e o período durante COVID.

Através do modelo com maior significância, foi possível apurar evidências no sentido de confirmar a existência de relações entre a aglomeração e a proximidade costeira na performance das unidades hoteleiras.

A mesma metodologia rejeitou a existência de relações significativas entre as variáveis: categoria, PME (classificação como PME) e dimensão.

Figura 5.20 - Quadro sumário das hipóteses

Variáveis	Localização Costeira	Aglomeração	Dimensão	Categoria	PME
Hipóteses	1.1	1.2	2	3	4
	+	+	+	inexistente	inexistente
GOPPAR Pré COVID	Suportada	Suportada	Não confirmada	Não confirmada	Não confirmada

6. Conclusão

6.1. Síntese e conclusão

O turismo é um setor relevante para a economia e para o desenvolvimento nacional. Neste âmbito, a investigação nesta área, reveste-se de particular importância, enquanto catalisador do desenvolvimento e performance das empresas turísticas.

O presente trabalho aponta para a adoção do GOPPAR enquanto indicador de rentabilidade mais ajustado de uma unidade hoteleira, contrariando a corrente maioritária de avaliar a performance através do RevPAR. Este é um dos principais contributos deste trabalho, procurando com os seus resultados, reforçar a relevância da discussão sobre qual a métrica mais ajustada para medir a *performance* hoteleira.

Acredita-se que para este resultado contribuí uma crescente adoção do USALI, consequência da atratividade do país a cadeiras internacionais, que com elas, trazem no seu conjunto de regras, procedimentos e instruções de serviço, o normativo contabilístico. Este (USALI) por sua vez difunde e reforça a importância na utilização e monitorização.

Concluiu-se ainda pela evidência estatística de relações positivas entre os determinantes aglomeração e localização costeira e o GOPPAR. Em oposição, conclui-se que a dimensão, a categoria e a qualificação como PME, não influenciam significativamente o nível de rentabilidade.

Este é outro contributo do presente estudo. O contributo suportado em dados primários, que permite aos *stakeholders*, avaliar melhor as suas opções de investimento. Em suma, para uma maior rentabilidade, devem ser preferidas localizações costeiras e com elevado nível de aglomeração, em detrimento da categoria e dimensão.

6.2. Limitações da investigação

Considera-se a reduzida dimensão da amostra, a principal limitação deste estudo, provocada por uma significativa relutância por parte dos quadros superiores e proprietários, na partilha de KPI de atividade.

A indústria hoteleira continua, de acordo com as respostas obtidas, a não partilhar indicadores de rendibilidade operacional. Por outro lado, admite-se que os destinatários dos inquéritos, tendo por suporte a base de dados do RNET, não sejam os mais apropriados. Muitos destes contactos são gerais, indiferenciados, conduzindo a que a leitura dos mesmos seja desconsiderada ou simplesmente eliminada.

Adicionalmente, o reduzido período de tempo do estudo e a inclusão de um evento de rutura estrutural, como foi a COVID, são entendidos como uma limitação adicional.

Apesar de considerada suficiente para as conclusões apresentadas na presente investigação, a reduzida dimensão poderá ter ignorado relações que enriqueceriam o poder explicativo dos modelos, em particular nas localizações com menor representação.

Outra limitação desta investigação prende-se com a não inclusão de um número superior de características ou determinantes adicionais, dada a complexidade do problema, conforme demonstrada em diversos estudos. Fazê-lo, acredita-se, aumentaria a dimensão do inquérito e reduziria a taxa de resposta.

6.3. Sugestão para futuras investigações

O horizonte temporal desta investigação, que compreendeu dois anos de pandemia COVID, nomeadamente 2020 e 2021, permitiu verificar diferentes comportamentos no período COVID para as unidades hoteleiras em distintas localizações geográfica e com níveis diferentes de aglomeração. Esta evidência, aplicava-se em diferentes escalas, mas em sentidos iguais para ambos os indicadores, ou seja, RevPAR e GOPPAR.

A primeira indicação para futuras investigações, compreende o estudo das características das unidades hoteleiras que determinam este comportamento distinto perante uma crise, alargando para o efeito o horizonte temporal em investigação.

O presente estudo apresentou variáveis com níveis de significância similares entre modelos, ainda que limitados pelo seu grau explicativo. Apesar do número crescente de estudos relativos à performance hoteleira, a utilização do GOPPAR enquanto indicador de performance não tem colhido a preferência dos investigadores. A necessidade de se estudar a relação entre o GOPPAR e o RevPAR, e de como a dimensão, a localização e as características podem influenciar esta relação, é a segunda sugestão de investigação.

Por fim, o presente estudo apontou evidências da relação entre a performance e a localização costeira e em localizações com maiores níveis de aglomeração. Estudar o que caracteriza estas localizações e o que as faz potenciar o resultado, procurando também responder se a localização atrai investimento ou se o investimento potencia o desenvolvimento estrutural do destino, é a terceira sugestão de uma futura investigação.

REFERÊNCIAS BIBLIOGRÁFICAS

- Alvarez, P., Sensini, L., Vazquez, M., & Bello, C. (2021). Management Accounting Practices and Performance of SMEs in the Hotel Industry: Evidence from an emerging economy. *Article in International Journal of Social Science and Business*, 12(2). <https://doi.org/10.30845/ijbss.v12n2p3>
- Assaf, A. G., & Tsionas, M. (2018). Measuring hotel performance: Toward more rigorous evidence in both scope and methods. *Tourism Management*, 69, 69–87. <https://doi.org/10.1016/J.TOURMAN.2018.05.008>
- Barros, C. P. (2005). Measuring efficiency in the hotel sector. *Annals of Tourism Research*, 32(2), 456–477. <https://doi.org/10.1016/j.annals.2004.07.011>
- Barros, C. P., Botti, L., Peypoch, N., & Solonandrasana, B. (2011). Managerial efficiency and hospitality industry: the Portuguese case. *Applied Economics*, 43(22), 2895–2905. <https://doi.org/10.1080/00036840802600145>
- Bresciani, S., Thrassou, A., & Vrontis, D. (2015). Determinants of performance in the hotel industry - An empirical analysis of Italy. *Global Business and Economics Review*, 17(1), 19–34. <https://doi.org/10.1504/GBER.2015.066531>
- Centro de Competências em Turismo Moneris. (2022). *Anuário Turismo edição 2022*.
- Claver-Cortés, E., Molina-Azorín, J. F., & Pereira-Moliner, J. (2007). The impact of strategic behaviours on hotel performance. *International Journal of Contemporary Hospitality Management*, 19(1), 6–20. <https://doi.org/10.1108/09596110710724125>
- Correia, H., Santos, L. L., Gomes, C., & Ferreira, R. R. (2016). USALI adapted to the small independent hotels. *EATSA - Euro-Asian Tourism Studies Association Annual Conference, 2nd, Lisbon, Peniche, Coimbra, Portugal, 2016*, 621–638. <https://iconline.ipleiria.pt/handle/10400.8/5805>
- Costa, V., & Almeida, C. (2015). Low-Cost Carriers, Local Economy and Tourism Development at Four Portuguese Airports. A Model of Cost–Benefit Analysis. *Journal of Spatial and Organizational Dynamics*, 3(4), 245–261. <https://www.jsod-cieo.net/journal/index.php/jsod/article/view/55>
- Daniel, A. C. M. (2010). Caracterização do Sector Turístico em Portugal. *Revista de Estudos Politécnicos Polytechnical Studies Review*, 14, 255–276.

- Diário imobiliário. (2023, Junho 30). *Sector hoteleiro português em alta neste Verão*. <https://www.diarioimobiliario.pt/Sector-hoteleiro-portugues-em-alta-neste-Verao>
- Dimitrić, M., Tomas Žiković, I., & Arbula Blečić, A. (2019). Profitability determinants of hotel companies in selected Mediterranean countries. *Economic Research-Ekonomska Istraživanja*, 32(1), 1977–1993. <https://doi.org/10.1080/1331677X.2019.1642785>
- Duro, J. A., & Turrión-Prats, J. (2019). Tourism seasonality worldwide. *Tourism Management Perspectives*, 31, 38–53. <https://doi.org/10.1016/j.tmp.2019.03.010>
- Eurostat. (2022, Julho 26). *2021: 43% of tourism nights spent in July and August*. <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20220726-1>
- Faria, A. R., Ferreira, L., & Trigueiros, D. (2018). Analyzing customer profitability in hotels using activity based costing. *Tourism & Management Studies*, 14(3), 65–74. <https://doi.org/10.18089/tms.2018.14306>
- Harris, P. J., & Mongiello, M. (2001). Key performance indicators in European hotel properties: general managers' choices and company profiles. *International Journal of Contemporary Hospitality Management*, 13(3), 120–128. <https://doi.org/10.1108/09596110110388909>
- HFTP. (2023). *USALI 12th Revised Edition | Hospitality Accounting Text | HFTP*. https://www.hftp.org/hospitality_resources/usali_guide/usali_12/
- Hilton Worldwide Holdings Inc. (2023). *Hilton Reports Fourth Quarter and Full Year Results*. <https://ir.hilton.com/>
- Horwath HTL. (2022). *Portugal Hotels & Chains Report 2022*.
- Hotel Travel News. (2009, Outubro 27). *A New “Location, Location, Location” for the Hotel Industry*. <https://www.breakingtravelnews.com/focus/article/a-new-location-location-location-for-the-hotel-industry/>
- Hotstats. (2019). *Why you should - and can! - benchmark profitability*. www.hotstats.com
- Hotstats. (2021a). *For Hotels, Revenue is the Sprint, Profit is the Marathon*. <https://www.hotstats.com/blog/for-hotels-revenue-is-the-sprint-profit-is-the-marathon>

- Hotstats. (2021b). *HotStats' COO Talks Profitability—Today and Tomorrow*.
<https://www.hotstats.com/blog/hotstats-coo-talks-profitability-today-and-tomorrow>
- Hotstats. (2022). *PROFIT MATTERS Global Hotel Performance Review*.
- INE. (2022). *Estatística do Turismo 2021*.
- INE. (2023a). *Atividade Turística - Dezembro de 2022*.
- INE. (2023b). *Atividade turística - Estatísticas Rápidas - Dez 2022*.
- INE. (2023c). *Remuneração bruta mensal média por trabalhador*.
- ITB Berlin, & IPK International. (2022, Março 17). *Upturn in international tourism*.
https://www.itb.com/en/press/press-releases/news_4803.html
- Kala, D., & Bagri, S. C. (2014). Key performance indicators for hospitality industry: a study from the tourist state of Uttarakhand, India. *Tourismos: An international multidisciplinary journal of tourism*, 9, 187–206.
- Kim, W. G., Cho, M., & Brymer, R. A. (2013). Determinants affecting comprehensive property-level hotel performance: The moderating role of hotel type. *International Journal of Hospitality Management*, 34(1), 404–412.
<https://doi.org/10.1016/j.ijhm.2012.12.002>
- Kimes, S. E. (2003). Revenue Management: A Retrospective. *Cornell Hotel and Restaurant Administration Quarterly*, 44(5–6), 131–138.
<https://doi.org/10.1177/001088040304400518>
- Lado-Sestayo, R., Otero-González, L., Vivel-Búa, M., & Martorell-Cunill, O. (2016). Impact of location on profitability in the Spanish hotel sector. *Tourism Management*, 52, 405–415. <https://doi.org/10.1016/j.tourman.2015.07.011>
- Lado-Sestayo, R., & Vivel-Búa, M. (2018a). Las corporaciones hoteleras españolas: determinantes internos y externos de su desempeño. *Contaduría y Administración*, 64(1), 84. <https://doi.org/10.22201/fca.24488410e.2018.1427>
- Lado-Sestayo, R., & Vivel-Búa, M. (2018b). Profitability in the hotel sector: a PLS approach. *Journal of Hospitality and Tourism Technology*, 9(3), 455–470.
<https://doi.org/10.1108/JHTT-10-2017-0118>
- Lado-Sestayo, R., & Vivel-Búa, M. (2019). Hotel profitability: a multilayer neural network approach. *Journal of Hospitality and Tourism Technology*, 11(1), 35–48. <https://doi.org/10.1108/JHTT-08-2017-0072>
- Lado-Sestayo, R., Vivel-Búa, M., & Otero-González, L. (2017). Determinants of TRevPAR: hotel, management and tourist destination. *International Journal of*

- Contemporary Hospitality Management*, 29(12), 3138–3156.
<https://doi.org/10.1108/IJCHM-03-2016-0151>
- Lado-Sestayo, R., Vivel-Búa, M., & Otero-González, L. (2020). Connection between hotel location and profitability drivers: an analysis of location-specific effects. *Current Issues in Tourism*, 23(4), 452–469.
<https://doi.org/10.1080/13683500.2018.1538203>
- Lee, J. (2009). Does Size Matter in Firm Performance? Evidence from US Public Firms. *International Journal of the Economics of Business*, 16(2), 189–203.
<https://doi.org/10.1080/13571510902917400>
- Lee, S., Pan, B., & Park, S. (2019). RevPAR vs. GOPPAR: Property- and firm-level analysis. *Annals of Tourism Research*, 76, 180–190.
<https://doi.org/10.1016/j.annals.2019.04.006>
- Marques, A. C. (2010). Caracterização do Sector Turístico em Portugal. *Revista de Estudos Politécnicos Polytechnical Studies Review*, 14, 255–276.
- Menicucci, E. (2018). The influence of firm characteristics on profitability. *International Journal of Contemporary Hospitality Management*, 30(8), 2845–2868. <https://doi.org/10.1108/IJCHM-04-2017-0219>
- Mihailescu, R. (2010). An assessment of Charter airline benefits for Port Elizabeth and the Eastern Cape. *Chinese Business Review*, 9(2).
<http://cybercapetown.com/Maps/images/sa.gif>.
- Decreto-Lei N°39/2008 de 7 de Março do Ministério da Economia, (2008).
- Moreira, C. O. (2018). Portugal as a tourism destination. *Méditerranée*, 130.
<https://doi.org/10.4000/mediterranee.10402>
- Oliveira, R., Pedro, M. I., & Marques, R. C. (2013). Efficiency and its determinants in Portuguese hotels in the Algarve. *Tourism Management*, 36, 641–649.
<https://doi.org/10.1016/j.tourman.2012.06.009>
- Panno, A. (2019). Performance measurement and management in small companies of the service sector; evidence from a sample of Italian hotels. *Measuring Business Excellence*, 24(2), 133–160. <https://doi.org/10.1108/MBE-01-2018-0004>
- Pnevmatikoudi, K., & Stavrinoudis, T. (2016). Classification of hotel performance measurement indicators presented in international scientific research. *European Journal of Tourism Research*, 12, 82–98.
<https://doi.org/10.54055/ejtr.v12i.214>

- Resolução do Conselho de Ministros N° 134/2017, (2017).
- Publituris. (2023, Maio 2). *Março trouxe 2,1 milhões de hóspedes para o alojamento turístico português*. <https://www.publituris.pt/2023/05/02/marco-trouxe-21-milhoes-de-hospedes-para-o-alojamento-turistico-portugues>
- Romero, I., Fernández-Serrano, J., & Cáceres-Carrasco, F. R. (2020). Tour operators and performance of SME hotels: Differences between hotels in coastal and inland areas. *International Journal of Hospitality Management*, *85*, 102348. <https://doi.org/10.1016/j.ijhm.2019.102348>
- Sainaghi, R. (2010). Hotel performance: state of the art. *International Journal of Contemporary Hospitality Management*, *22*(7), 920–952. <https://doi.org/10.1108/09596111011066617>
- Sainaghi, R., Baggio, R., Phillips, P., & Mauri, A. G. (2020). Hotel performance and research streams: a network cluster analysis. *International Journal of Contemporary Hospitality Management*, *32*(2), 425–462. <https://doi.org/10.1108/IJCHM-05-2017-0260>
- Sainaghi, R., Phillips, P., Baggio, R., & Mauri, A. (2018). Cross-citation and authorship analysis of hotel performance studies. *International Journal of Hospitality Management*, *73*, 75–84. <https://doi.org/10.1016/J.IJHM.2018.02.004>
- Santos, L. L., Gomes, C., Faria, A. R., Lunkes, R. J., Malheiros, C., Silva da Rosa, F., & Nunes, C. (2016). *Contabilidade de gestão hoteleira* (ATF - Edições Técnicas, Ed.; 1.^a ed.). ATF - Edições Técnicas.
- Schwartz, Z., Altin, M., & Singal, M. (2017). Performance measures for strategic revenue management: RevPAR versus GOPPAR. *Journal of Revenue and Pricing Management*, *16*(4), 357–375. <https://doi.org/10.1057/rpm.2016.23>
- Serrasqueiro, Z. S., & Nunes, P. M. (2008). Performance and size: empirical evidence from Portuguese SMEs. *Small Business Economics*, *31*(2), 195–217. <https://doi.org/10.1007/s11187-007-9092-8>
- Singh, A. J., & Schmidgall, R. S. (2002). Analysis of financial ratios commonly used by US lodging financial executives. *Journal of Retail & Leisure Property*, *2*(3), 201–213. <https://doi.org/10.1057/palgrave.rlp.5090210>
- Siteminder. (sem data). *Your comprehensive guide to hotel classifications*. Obtido 26 de Setembro de 2023, de <https://www.siteminder.com/r/hotel-classifications/#hotel-size-and-number-of-rooms>

- Song, B. D., & Ko, Y. D. (2017). Quantitative Approaches for Location Decision Strategies of a Hotel Chain Network. *International Journal of Hospitality Management*, 67, 75–86. <https://doi.org/10.1016/j.ijhm.2017.08.004>
- Sousa, V. (2021). Perceção internacional de Portugal como destino turístico International perception of Portugal as a tourism destination. *Journal of Tourism & Development / n.º*, 36. <https://doi.org/10.34624/rtd.v36i2.8797>
- Tan, K. P.-S., Li, X. (Robert), & Lee, S. (2022). Exploring the Determinants of Hotel Operating Performance Stabilization in Emerging Markets: Deciphering the Myth. *Journal of Hospitality & Tourism Research*, 46(4), 742–770. <https://doi.org/10.1177/1096348021996440>
- Travanca, C. M. R., Vieira, C. M. R., & Félix, E. G. S. (2022). Determinants of Profitability in the Tourism Sector in Portugal. Em *Tourism, Hospitality and Event Management* (pp. 193–215). Springer Nature. https://doi.org/10.1007/978-3-030-89232-6_10
- Turismo de Portugal. (2022). *Turismo em Portugal 2021*.
- Turismo de Portugal. (2023). *Turismo em Números 2022*.
- Yang, Y., Luo, H., & Law, R. (2014). Theoretical, empirical, and operational models in hotel location research. *International Journal of Hospitality Management*, 36, 209–220. <https://doi.org/10.1016/j.ijhm.2013.09.004>

APÊNDICE A: NOTEBOOK_PREP_QUESTIONARIO_FINAL.IPYNB

```
# Importação da biblioteca pandas
import pandas as pd

# Importação do ficheiro base
df=pd.read_excel('ListaRegistos_EstHoteleiros_RNET_ (2023-02-04).xlsx')
df.info()

#Visualização primária dos dados
df.head(5)

# Criação de um novo dataframe para o questionário
df_questionario = df[['Designacao', 'Entidade Exploradora', 'Contacto Email', 'Numero do Registo']].copy()

# Alteração dos nomes das colunas
df_questionario.columns=['firstname', 'lastname', 'email', 'token']

# Eliminação de caracteres não numéricos da coluna chave denominada token
df_questionario['token']=df_questionario['token'].str.replace(r'\D', '')

# Alteração do tipo para número inteiro
df_questionario.token = df_questionario.token.astype(int)

#Exportação do ficheiro para integração no limesurvey
df_questionario.to_csv('Convites.csv', index=False)
```

APÊNDICE B: QUESTIONÁRIO

O impacto da localização e da dimensão na rendibilidade operacional dos hotéis em Portugal



Obrigado pela sua participação neste estudo.

Este questionário foi desenvolvido para que o seu preenchimento não exceda 5 minutos do seu tempo, fazendo parte de uma investigação académica no âmbito de uma dissertação de mestrado.

É composto por **7 questões** sobre as características e a rendibilidade operacional no período de 2017 a 2021, do hotel que representa.

As respostas são confidenciais e garantimos o anonimato dos dados recolhidos.

Existem 6 perguntas neste questionário.

Enquadramento do hotel

Neste grupo identificaremos as características do hotel

O Hotel está integrado numa empresa certificada como PME (Pequena e Média Empresa)? *

Por favor, seleccione **apenas uma** das seguintes opções:

Sim

Não

nos termos da Recomendação da Comissão 2003-361-CE

A empresa tem integrados outros negócios para além do Hotel em análise? *

Por favor, seleccione **apenas uma** das seguintes opções:

- Sim
- Não

O Hotel reporta de acordo com o Uniform System of Accounts Lodging Industry (USALI)? *

🗳 Escolha uma das seguintes respostas

Por favor, seleccione **apenas uma** das seguintes opções:

- Sim
- Não
- Parcialmente/Adaptado
- Não sei

Número de estrelas *

Por favor, seleccione **apenas uma** das seguintes opções:

- 1
- 2
- 3
- 4
- 5

Número total de unidades de alojamento *

❶ A resposta deve estar entre 0 e 1000

❷ Neste campo apenas pode ser introduzido um valor inteiro.

Por favor, escreva aqui a sua resposta:

Dados financeiros do hotel

Nesta secção abordaremos os dados financeiros em estudo (RevPAR e GOPPAR)

Total revenue per available room e gross operating profit per available room *

	RevPAR	GOPPar
2021	<input type="text"/>	<input type="text"/>
2020	<input type="text"/>	<input type="text"/>
2019	<input type="text"/>	<input type="text"/>
2018	<input type="text"/>	<input type="text"/>
2017	<input type="text"/>	<input type="text"/>

Nos anos em que não exista registo ou atividade, favor inserir 0 (ZERO).

Obrigado pela sua valiosa participação.

Se desejar, podemos partilhar as principais conclusões deste estudo. Para tal queira por favor enviar-nos o seu contacto via: {ADMINEMAIL}

APÊNDICE C: NOTEBOOK_ANALISEUNIVERSO_FINAL.IPYNB

```
# importação das bibliotecas iniciais, montagem da drive e
carregamento do ficheiro base
import pandas as pd
df=pd.read_excel(ListaRegistos_EstHoteleiros_RNET_ (2023-02-
04).xlsx')

# Mostrar a característica do dataframe carregado
df.info()

# Importação de bibliotecas necessárias para a visualização e
tratamento de dados
import matplotlib.pyplot as plt
import seaborn as sns

# Definição de estilo
sns.set_style("whitegrid")

# Criação de gráfico de barras para a distribuição por Tipologia
plt.figure(figsize=(15, 6))
sns.countplot(y='Tipologia', data=df,
order=df['Tipologia'].value_counts().index, palette="coolwarm")
plt.title(' ')
plt.xlabel('Número de Unidades Hoteleiras')
plt.ylabel('Tipologia')
plt.tight_layout()
plt.show()

# Produzir um "pie chart" para a variável "Classificação"
plt.figure(figsize=(15, 6))
df['Classificacao'].value_counts().plot(kind='pie',
autopct='%1.1f%%', startangle=140, colors=plt.cm.Paired.colors)
plt.title(' ')
plt.ylabel('')
plt.show()

# Produção de um gráfico de barras para a distribuição por NUTS III
plt.figure(figsize=(15, 6))
sns.countplot(y='NUT III', data=df, order=df['NUT
III'].value_counts().index, palette="coolwarm")
plt.title(' ')
plt.xlabel('Número de Unidades Hoteleiras')
plt.ylabel('NUTSIII')
plt.tight_layout()
plt.show()

# Produção de um gráfico de barras para a distribuição por Distrito
plt.figure(figsize=(15, 6))
```

```

sns.countplot(y='Localizacao Distrito', data=df,
order=df['Localizacao Distrito'].value_counts().index,
palette="coolwarm")
plt.title(' ')
plt.xlabel('Número de Unidades Hoteleiras')
plt.ylabel('Distrito')
plt.tight_layout()
plt.show()

# Produção de histograma para a distribuição do número de unidades
de alojamento
plt.figure(figsize=(15, 6))
sns.histplot(df['Numero unidades alojamento'], bins=50,
color="teal", kde=True)
plt.title(' ')
plt.xlabel('Número de Unidades de Alojamento')
plt.ylabel('Frequência')
plt.tight_layout()
plt.show()

# Criação de um quadro com contagens, percentagens relativas,
médias e modas por distrito

# Número de unidades hoteleiras por distrito
num_unidades = df['Localizacao Distrito'].value_counts()
# Percentagem do total do número de unidades hoteleiras por
distrito
perc_num_unidades = (num_unidades / num_unidades.sum()) * 100
# Número total de unidades de alojamento por distrito
total_unidades_alojamento = df.groupby('Localizacao
Distrito')['Numero unidades alojamento'].sum()
# Média de unidades de alojamento por distrito
media_unidades_alojamento = df.groupby('Localizacao
Distrito')['Numero unidades alojamento'].mean()
# Moda de unidades de alojamento por distrito
moda_unidades_alojamento = df.groupby('Localizacao
Distrito')['Numero unidades alojamento'].apply(lambda x:
x.mode().iloc[0])
# Percentagem do total do número de unidades de alojamento por
distrito
perc_total_unidades_alojamento = (total_unidades_alojamento /
total_unidades_alojamento.sum()) * 100

# Junção dos valores calculados num único dataframe
quadro = pd.DataFrame({
    'N° Unidades Hoteleiras': num_unidades,
    '% do Total UH': perc_num_unidades,
    'N° UA': total_unidades_alojamento,
    '% do Total UA': perc_total_unidades_alojamento,

```

```

    'Média de UA': media_unidades_alojamento,
    'Moda de UA': moda_unidades_alojamento
))

# Arredondamento das percentagens para duas casas decimais
quadro['% do Total UH'] = quadro['% do Total UH'].round(2)
quadro['% do Total UA'] = quadro['% do Total UA'].round(2)
quadro['Média de UA'] = quadro['Média de UA'].round(0)
quadro['Nº UA'] = quadro['Nº UA'].round(0)
quadro['Nº Unidades Hoteleiras'] = quadro['Nº Unidades
Hoteleiras'].round(0)
# Ordenação do quadro pelo nome do distrito
quadro.sort_index(inplace=True)
# Função para exibir a tabela de forma mais visual
def display_table(data):
    fig, ax = plt.subplots(figsize=(15, 10))
    ax.axis('off')
    ax.axis('tight')
    table_data = []
    columns = [""]
    columns.extend(data.columns)
    table_data.append(columns)
    for row in data.iterrows():
        table_data.append([row[0]] + list(row[1]))
    ax.table(cellText=table_data, cellLoc='center', loc='center')

# Exibição da tabela
display_table(quadro)

# Criação de um "heatmap" com o número de unidades por distrito

# Importação da biblioteca necessária para a visualização e
tratamento dos dados
import geopandas as gpd
# Carregamento do arquivo shapefile (shp) de Portugal
gdf_portugal = gpd.read_file("georef-portugal-distrito-
millesime.shp")
# Mapeamento manual para corrigir os nomes com erros de codificação
mapeamento_nomes_correcao = {
    'Ã\x89vora': 'Évora',
    'Setã°bal': 'Setúbal',
    'Aãores': 'Açores',
    'Santarã©m': 'Santarém',
    'Braganãa': 'Bragança'
}
# Aplicação do mapeamento para corrigir os nomes no geodataframe
gdf_portugal['dis_name'] =
gdf_portugal['dis_name'].replace(mapeamento_nomes_correcao)

```

```

# Mapeamento manual dos nomes dos distritos para agregação das
ilhas
mapeamento_nomes_agregacao = {
    'Ilha da Madeira': 'Madeira',
    'Ilha de Porto Santo': 'Madeira',
    'Ilha Terceira': 'Açores',
    'Ilha da Graciosa': 'Açores',
    'Ilha de Santa Maria': 'Açores',
    'Ilha de São Jorge': 'Açores',
    'Ilha de São Miguel': 'Açores',
    'Ilha do Faial': 'Açores',
    'Ilha do Pico': 'Açores'
}
# Aplicando o mapeamento no dataframe original para agregação
df['Localizacao Distrito Agregado'] = df['Localizacao
Distrito'].replace(mapeamento_nomes_agregacao)
# Agregação dos dados do dataframe original com base na agregação
quadro_agregacao = df.groupby('Localizacao Distrito
Agregado').agg({'Localizacao Distrito': 'count'})
quadro_agregacao.rename(columns={'Localizacao Distrito': 'Número de
Unidades'}, inplace=True)
# Integração dos dados agregados com o geodataframe corrigido
gdf_portugal_integrado =
gdf_portugal.merge(quadro_agregacao['Número de Unidades'],
left_on='dis_name', right_index=True, how='left')
# Verificação dos dados após integração
gdf_portugal_integrado[['dis_name', 'Número de Unidades']].head()

# Separação dos geodataframes para Portugal Continental, Açores e
Madeira
gdf_continental =
gdf_portugal_integrado[gdf_portugal_integrado['dis_name'] !=
'Açores']
gdf_continental = gdf_continental[gdf_continental['dis_name'] !=
'Madeira']
gdf_acores =
gdf_portugal_integrado[gdf_portugal_integrado['dis_name'] ==
'Açores']
gdf_madeira =
gdf_portugal_integrado[gdf_portugal_integrado['dis_name'] ==
'Madeira']
# Configurar a figura e os eixos
fig = plt.figure(figsize=(15, 9))
# Definir as posições dos mapas na figura
ax1 = fig.add_axes([0.05, 0.05, 0.45, 0.9]) # Portugal Continental
ax2 = fig.add_axes([0.55, 0.55, 0.4, 0.4]) # Açores
ax3 = fig.add_axes([0.55, 0.05, 0.4, 0.4]) # Madeira

# Plotar os mapas

```

```

# Portugal Continental
gdf_continental.plot(column='Número de Unidades', cmap='coolwarm',
legend=True, ax=ax1)
ax1.set_title('Portugal Continental')
ax1.set_axis_off()
# Açores
gdf_acores.plot(column='Número de Unidades', cmap='coolwarm',
legend=True, ax=ax2)
ax2.set_title('Açores')
ax2.set_axis_off()

# Madeira
gdf_madeira.plot(column='Número de Unidades', cmap='coolwarm',
legend=True, ax=ax3)
ax3.set_title('Madeira')
ax3.set_axis_off()
# Apresentação dos mapas
plt.show()

# Criação de um "heatmap" com o número de unidades de alojamento
por distrito

# Agregação dos dados do dataframe original com base na agregação
quadro_agregacao_alojamento = df.groupby('Localizacao Distrito
Agregado').agg({'Numero unidades alojamento': 'sum'})
# Integração dos dados agregados de alojamento com o geodataframe
corrigido
gdf_portugal_integrado_alojamento =
gdf_portugal.merge(quadro_agregacao_alojamento['Numero unidades
alojamento'], left_on='dis_name', right_index=True, how='left')
# Configuração da figura e os eixos para os mapas
fig = plt.figure(figsize=(15, 9))
# Definição das posições dos mapas na figura
ax1 = fig.add_axes([0.05, 0.05, 0.45, 0.9]) # Portugal Continental
ax2 = fig.add_axes([0.55, 0.55, 0.4, 0.4]) # Açores
ax3 = fig.add_axes([0.55, 0.05, 0.4, 0.4]) # Madeira

# Plotar os mapas com número de unidades de alojamento
# Portugal Continental
gdf_portugal_integrado_alojamento[gdf_portugal_integrado_alojamento
['dis_name'] !=
'Açores'][gdf_portugal_integrado_alojamento['dis_name'] !=
'Madeira'].plot(column='Numero unidades alojamento',
cmap='coolwarm', legend=True, ax=ax1)
ax1.set_title('Portugal Continental')
ax1.set_axis_off()
# Açores

```

```
gdf_portugal_integrado_alojamento[gdf_portugal_integrado_alojamento
['dis_name'] == 'Açores'].plot(column='Numero unidades alojamento',
cmap='coolwarm', legend=True, ax=ax2)
ax2.set_title('Açores')
ax2.set_axis_off()
# Madeira
gdf_portugal_integrado_alojamento[gdf_portugal_integrado_alojamento
['dis_name'] == 'Madeira'].plot(column='Numero unidades
alojamento', cmap='coolwarm', legend=True, ax=ax3)
ax3.set_title('Madeira')
ax3.set_axis_off()
# Apresentação dos mapas
plt.show()
```

APÊNDICE D: NOTEBOOK_TRANSFORMACAO DADOS_FINAL.IPYNB

```
import os
import csv
import numpy as np
import pandas as pd
from geopy.distance import geodesic

# Carregamento do ficheiro com o as respostas ao inquérito
dfrespostas=pd.read_csv('results-survey767592_Final.csv')

# Carregamento do ficheiro com o universo em estudo
dfdados=pd.read_csv('Empreendimentos_Turisticos_Existentes_4943.csv')
dfdados.info()
dfrespostas.info()
dfrespostas.head(5)

# Combinação das duas dataframes através do 'Código de acesso' da
df resposta e 'NrRNET' dos dados integrais
dfa = dfrespostas.merge(dfdados, left_on='Código de acesso',
right_on='NrRNET', how='left')

# Eliminação de colunas redundantes ou fora do espectro de estudo
dfa.drop(columns=['ID da resposta', 'Data de submissão', 'Última
página', 'Idioma inicial', 'Semente', 'Código de acesso', 'Data de
início', 'Endereço IP', 'URL de origem', 'UNESCO', 'Geoparques',
'Biosfera', 'RedeAldeias', 'Email', 'Website', 'ImovelClassificado', 'De
sigOutrasCertif', 'OutrasCertifPremios', 'DesigCertifQualServ', 'Certi
fQualServ', 'DesigCertifAmbiental', 'CertifAmbiental', 'RecTurNatureza
', 'AtribuicaoUT', 'DataPostClassif', 'ParecerPostClassif', 'Tempototal
', 'NrRNET', 'Denominacao', 'EntProprietaria', 'EntExploradora', 'NrSuit
es', 'NrApart', 'NrMoradias', 'NrCampistas', 'IntegraCT', 'DesignacaoCT'
, 'CampoGolfe', 'LimitePROT', 'LimiteAlbufeiras', 'DataAtualiza', 'SeloC
leanSafe', 'RNET', 'RNAP', 'RedeNatura2000Global', 'RedeNatura2000Sitio
s', 'RedeNatura2000ZPE', 'ZonaTerrestreProtecao', 'LimitePOC', 'ERT', 'P
NCT', 'Data da última .. ', 'Tempo do grupo:... ', 'Tempo da pergun..
.1', 'Tempo da pergun.. .2', 'Tempo da pergun.. .4', 'Tempo da
pergun.. .5', 'Tempo do grupo:... .1', 'Tempo da pergun.. ', 'Tempo da
pergun...3', 'OutrosEquip', 'AnoAberturaEmpreendimento', 'DataUltClass
if', 'ValidadeClassif', 'Endereco', 'OBJECTID', 'SHAPESTArea', 'SHAPEST
Length', 'ClassificacaoTdP'], inplace=True)

# Correção dos nomes de algumas variáveis chave
new_names = {'Total revenue p.. [2021][RevPAR]': 'REVPAR21',
              'Total revenue p.. [2020][RevPAR]': 'REVPAR20',
              'Total revenue p.. [2019][RevPAR]': 'REVPAR19',
              'Total revenue p.. [2018][RevPAR]': 'REVPAR18',
              'Total revenue p.. [2017][RevPAR]': 'REVPAR17',
```

```

        'Total revenue p.. [2021][GOPPar]': 'GOPPAR21',
        'Total revenue p.. [2020][GOPPar]': 'GOPPAR20',
        'Total revenue p.. [2019][GOPPar]': 'GOPPAR19',
        'Total revenue p.. [2018][GOPPar]': 'GOPPAR18',
        'Total revenue p.. [2017][GOPPar]': 'GOPPAR17',
        'O Hotel está in.. ': 'PME',
        'A empresa tem i.. ': 'OUTNEGOCIOS',
        'O Hotel reporta.. ': 'USALI',
        'Número de estrelas': 'ESTRELAS',
        'Número total de.. ': 'UA',
        'NrUnidAloj': 'UA_RNET'}
dfa = dfa.rename(columns=new_names)

# Correção aos números de UA em falta
dfa['UA_RNET'].fillna(value=dfa['UA'], inplace=True)
dfa['UA_RNET'] = dfa['UA_RNET'].round().astype(int)

# Substituição de valores inexistentes por "nan" e validação
dfa[['REVPAR21', 'REVPAR20',
     'REVPAR19', 'REVPAR18', 'REVPAR17', 'GOPPAR21', 'GOPPAR20', 'GOPPAR19',
     'GOPPAR18', 'GOPPAR17']] = dfa[['REVPAR21', 'REVPAR20',
     'REVPAR19', 'REVPAR18', 'REVPAR17', 'GOPPAR21', 'GOPPAR20', 'GOPPAR19',
     'GOPPAR18', 'GOPPAR17']].replace(0, np.nan)
selected_columns = dfa.loc[:, ['REVPAR21', 'REVPAR20',
     'REVPAR19', 'REVPAR18', 'REVPAR17', 'GOPPAR21', 'GOPPAR20', 'GOPPAR19',
     'GOPPAR18', 'GOPPAR17']]
print(selected_columns.to_string(index=False))

# Considerando a duplicação, eliminei colunas redundantes
dfa = dfa.drop(['Categoria', 'FiabilidadeGeo', 'UA'], axis=1)

# criação de intervalos de proximidade da costa
conditions = [
    dfa['ZonaCosteira'] == 'Zona Costeira (2km)',
    (dfa['FaixaCosteira'] == 'Faixa Costeira (5km)') &
    (dfa['ZonaCosteira'] == 'Não Abrangido'),
    (dfa['ZonaCosteira'].isnull()) &
    (dfa['FaixaCosteira'].isnull())
]

choices = ['ate2Kmcosta', 'entre2a5kmcosta', '+5kmcosta']
dfa['PROXCOSTA'] = np.select(conditions, choices,
default='+5kmcosta')
# Transforma colunas em categorias

cols = ['PME', 'OUTNEGOCIOS',
        'USALI', 'TipologiaET', 'ESTRELAS', 'PROXCOSTA',
        'FreguesiasCosteiras', 'NUTSIIICosteiras', 'SPA', 'SalasReunioes']
dfa[cols] = dfa[cols].apply(lambda x:x.astype('category'))

```

```

dfa.info()
# Tratamento dos dados para geo referenciação

# Tratamento dos valores ausentes ou nulos na coluna "LatLong", com
separação 'Latitude' e 'Longitude' nos datasets

dfdados['LatLong'] = dfdados['LatLong'].astype(str)
dfdados['Latitude'], dfdados['Longitude'] =
zip(*dfdados['LatLong'].apply(lambda x: x.replace(',', '.').split('
; ') if ' ; ' in x else (None, None)))
dfa['LatLong'] = dfa['LatLong'].astype(str)
dfa['Latitude'], dfa['Longitude'] =
zip(*dfa['LatLong'].apply(lambda x: x.replace(',', '.').split('
; ') if ' ; ' in x else (None, None)))

# Conversão das novas colunas para o tipo float

dfdados['Latitude'] = pd.to_numeric(dfdados['Latitude'],
errors='coerce')
dfdados['Longitude'] = pd.to_numeric(dfdados['Longitude'],
errors='coerce')
dfa['Latitude'] = pd.to_numeric(dfa['Latitude'], errors='coerce')
dfa['Longitude'] = pd.to_numeric(dfa['Longitude'], errors='coerce')

# Drops das coluna "LatLong" originais nos datasets
dfa.drop(columns=['LatLong'], inplace=True)
dfdados.drop(columns=['LatLong'], inplace=True)
dfa.info()

# Localização e apresentação dos valores nulos ou inexistentes nas
colunas 'Latitude' e 'Longitude' no conjunto de dados 'dfa'

null_latlong_dfa = dfa.loc[dfa['Latitude'].isnull() |
dfa['Longitude'].isnull()]
null_latlong_dfa[['Latitude', 'Longitude']]

# Eliminação dos registos associados aos valores nulos nas colunas
"Latitude" e "Longitude", tendo por base a impossibilidade de os
preencher. Verificação
dfa = dfa.dropna(subset=['Latitude'])
null_latlong_dfa = dfa[dfa['Latitude'].isnull() |
dfa['Longitude'].isnull()]
null_latlong_dfa[['Latitude', 'Longitude']]

# Criação de função para o cálculo do número de unidades hoteleiras
num determinado raio, excluindo qualquer valor inválido

def count_hotels_within_radius(lat, long, radius_km):
    count = 0

```

```

    for index, row in dfdados.iterrows():
        hotel_lat = row['Latitude']
        hotel_long = row['Longitude']
        if pd.notnull(hotel_lat) and pd.notnull(hotel_long):
            distance = geodesic((lat, long), (hotel_lat,
hotel_long)).km
            if distance <= radius_km:
                count += 1
    return count

# Aplicação da função para calcular o número de hotéis em raios de
1 km, 5 km e 10 km

dfa['Hotels_1km'] = dfa.apply(lambda row:
count_hotels_within_radius(row['Latitude'], row['Longitude'], 1),
axis=1)
dfa['Hotels_5km'] = dfa.apply(lambda row:
count_hotels_within_radius(row['Latitude'], row['Longitude'], 5),
axis=1)
dfa['Hotels_10km'] = dfa.apply(lambda row:
count_hotels_within_radius(row['Latitude'], row['Longitude'], 10),
axis=1)

# Validação das primeiras linhas para verificar as novas colunas
dfa[['Latitude', 'Longitude', 'Hotels_1km', 'Hotels_5km',
'Hotels_10km']].head()
# Criação de função para o cálculo do número de UA num determinado
raio, excluindo qualquer valor inválido

def sum_UA_within_radius(lat, long, radius_km):
    total_UA = 0
    if np.isfinite(lat) and np.isfinite(long):
        for index, row in dfdados.iterrows():
            hotel_lat = row['Latitude']
            hotel_long = row['Longitude']
            if pd.notnull(hotel_lat) and pd.notnull(hotel_long) and
np.isfinite(hotel_lat) and np.isfinite(hotel_long):
                distance = geodesic((lat, long), (hotel_lat,
hotel_long)).km
                if distance <= radius_km:
                    total_UA += row['NrUnidAloj']
    return total_UA

# Aplicação da função para somar o número de UA em raios de 1 km, 5
km e 10 km

dfa['UA_1km'] = dfa.apply(lambda row:
sum_UA_within_radius(row['Latitude'], row['Longitude'], 1), axis=1)

```

```

dfa['UA_5km'] = dfa.apply(lambda row:
sum_UA_within_radius(row['Latitude'], row['Longitude'], 5), axis=1)
dfa['UA_10km'] = dfa.apply(lambda row:
sum_UA_within_radius(row['Latitude'], row['Longitude'], 10),
axis=1)

# Validação das primeiras linhas para verificar as novas colunas

dfa[['Latitude', 'Longitude', 'UA_1km', 'UA_5km',
'UA_10km']].head()

dfa.info()
# Código para transformação dos dados em painel
# Definindo as colunas que serão mantidas como identificadores
(todas as colunas exceto as de REVPAR e GOPAR)

id_vars = list(dfa.columns.difference(['REVPAR21', 'REVPAR20',
'REVPAR19', 'REVPAR18', 'REVPAR17', 'GOPPAR21', 'GOPPAR20',
'GOPPAR19', 'GOPPAR18', 'GOPPAR17']))

# Reshape das colunas REVPAR
revpar_melted = dfa.melt(id_vars=id_vars, value_vars=['REVPAR21',
'REVPAR20', 'REVPAR19', 'REVPAR18', 'REVPAR17'], var_name='Year',
value_name='revpar')
revpar_melted['Year'] =
revpar_melted['Year'].str.extract('(\\d+)').astype(int)

# Reshape das colunas GOPAR
gopar_melted = dfa.melt(id_vars=id_vars, value_vars=['GOPPAR21',
'GOPPAR20', 'GOPPAR19', 'GOPPAR18', 'GOPPAR17'], var_name='Year',
value_name='goppar')
gopar_melted['Year'] =
gopar_melted['Year'].str.extract('(\\d+)').astype(int)

# Junção dos resultados
final_df = pd.merge(revpar_melted, gopar_melted, on=id_vars +
['Year'])

# Criação da coluna 'covid' com base nos valores da coluna 'Year'
final_df['covid'] = final_df['Year'].apply(lambda x: 0 if x in [17,
18, 19] else 1)
dfa = final_df

# Preenchimento dos valores inexistentes nas variáveis de
proximidade costeira
dfa['FreguesiasCosteiras'] =
dfa['FreguesiasCosteiras'].fillna('Não')
dfa['NUTSIIICosteiras'] = dfa['NUTSIIICosteiras'].fillna('Não')

```

```

# Verificação da existência de valores nulos nas variáveis
categóricas 'PME' e 'USALI'
null_values_pme = dfa['PME'].isnull().sum() if 'PME' in dfa.columns
else "Variável 'PME' não encontrada."
null_values_usali = dfa['USALI'].isnull().sum() if 'USALI' in
dfa.columns else "Variável 'USALI' não encontrada."

# Agrupamento dos valores da variável 'USALI' em 1 e 0
if 'USALI' in dfa.columns:
    dfa['USALI'] = dfa['USALI'].apply(lambda x: '1' if x in
['Parcialmente/Adaptado', 'Sim'] else '0')

# Descrição atualizada da variável categórica 'USALI'
usali_description_updated =
dfa['USALI'].value_counts().reset_index() if 'USALI' in dfa.columns
else "Variável 'USALI' não encontrada."

# visualização das primeiras linhas do dataframe atualizado
first_rows_updated = dfa.head()
null_values_pme, null_values_usali, usali_description_updated,
first_rows_updated
dfa.info()

#Exportação ficheiro de trabalho
dfa.to_csv('Dados_transf_base.xlsx', index=False)
output_path = 'Dados_transf_base.csv'
dfa.to_csv(output_path, index=False)

```

APÊNDICE E: NOTEBOOK_ANALISEAMOSTRA_FINAL.IPYNB

```
import os
import csv
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import geopandas as gpd
%matplotlib inline

# Importação dos dados transformados para análise
df=pd.read_csv('Dados_transf_base.csv')
df.info()

# Criação de gráfico de barras para a distribuição por Tipologia
# Cálculo das contagens por Tipologia com ajustamento à
transformação para dados em painel (TDP)
tipologia_counts = df['TipologiaET'].value_counts() / 5
tipologia_counts = tipologia_counts.reset_index()
tipologia_counts.columns = ['Tipologia', 'count']
# Definição do estilo
sns.set_style("whitegrid")
# Execução do gráfico de barras para a distribuição por Tipologia
plt.figure(figsize=(15, 6))
sns.barplot(y='Tipologia', x='count', data=tipologia_counts,
order=tipologia_counts['Tipologia'], palette="coolwarm")
plt.title('')
plt.xlabel('Número de Unidades Hoteleiras')
plt.ylabel('Tipologia')
plt.tight_layout()
plt.show()

# Produção de # Produção de um gráfico de barras para a
distribuição por Distrito com ajustamento à TDP
Distrito_counts = df['Distrito'].value_counts() / 5
Distrito_counts = Distrito_counts.reset_index()
Distrito_counts.columns = ['Distrito', 'count']
plt.figure(figsize=(15, 8))
sns.barplot(y='Distrito', x='count', data=Distrito_counts,
order=Distrito_counts['Distrito'], palette="coolwarm")
plt.title('')
plt.xlabel('Número de Unidades Hoteleiras')
plt.ylabel('Distrito')
plt.tight_layout()
plt.show()
um "pie chart" para a variável "Classificação"
plt.figure(figsize=(15, 6))
```

```

df['ESTRELAS'].value_counts().plot(kind='pie', autopct='%1.1f%%',
startangle=140, colors=plt.cm.Paired.colors)
plt.title('')
plt.ylabel('')
plt.show()

# Produção de um gráfico de barras para a distribuição Por NUTSIII
com ajustamento à TDP
NUTSIII_counts = df['NUTSIII'].value_counts() / 5
NUTSIII_counts = NUTSIII_counts.reset_index()
NUTSIII_counts.columns = ['NUTSIII', 'count']
plt.figure(figsize=(15, 8))
sns.barplot(y='NUTSIII', x='count', data=NUTSIII_counts,
order=NUTSIII_counts['NUTSIII'], palette="coolwarm")
plt.title('')
plt.xlabel('Número de Unidades Hoteleiras')
plt.ylabel('NUTSIII')
plt.tight_layout()
plt.show()

# Produção de histograma para a distribuição do número de unidades
de alojamento ajustada à TDP

# Definição de uma matriz de pesos, onde cada peso é 1/5 para
ajustamento à TDP
weights = [1/5] * len(df['UA_RNET'])
plt.figure(figsize=(15, 6))
plt.hist(df['UA_RNET'], bins=50, color="teal", weights=weights,
density=False)
plt.xlabel('Número de UA')
plt.ylabel('Frequência')
plt.tight_layout()
plt.show()

# Criação de um "heatmap" com o número de unidades por distrito

# Agrupamento pelo campo "Distrito" e somar o número de unidades em
'UA_RNET' ajustado à TDP
distrito_soma =
df.groupby('Distrito')['UA_RNET'].sum().reset_index()
distrito_soma['UA_RNET'] = distrito_soma['UA_RNET'] / 5
# Contagem do número total de registros em 'UA_RNET'
total_registros = df['UA_RNET'].sum() / 5
# Carregamento do arquivo shapefile (shp) de Portugal
distritos_geo = gpd.read_file("georef-portugal-distrito-
millesime.shp")
# Mapeamento manual para corrigir os nomes com erros de codificação
mapeamento_nomes_correcao = {
    'Ã\x89vora': 'Évora',

```

```

    'Setúbal': 'Setúbal',
    'Açores': 'Açores',
    'Santarém': 'Santarém',
    'Bragança': 'Bragança'
}
# Aplicação do mapeamento para corrigir os nomes no geodataframe
distritos_geo['dis_name'] =
distritos_geo['dis_name'].replace(mapeamento_nomes_correcao)
# Realização de uma junção entre os dados geográficos e o dataframe
distrito_soma
merged =
distritos_geo.set_index('dis_name').join(distrito_soma.set_index('D
istrito'))
# Plotar o heatmap
fig, ax = plt.subplots(1, 1, figsize=(15, 9))
merged.plot(column='UA_RNET', cmap='YlOrBr', linewidth=0.8, ax=ax,
edgecolor='0.8', legend=True)
plt.show()

# Criação de um "heatmap" com o número de unidades de alojamento
por distrito

# Contagem do número de registros por distrito ajustado à TDP
distrito_count =
df.groupby('Distrito').size().reset_index(name='count')
distrito_count['count'] /= 5
# Realização de uma junção entre os dados geográficos e o dataframe
distrito_count
merged_count =
distritos_geo.set_index('dis_name').join(distrito_count.set_index('
Distrito'))
# Plotar o heatmap
fig, ax = plt.subplots(1, 1, figsize=(15, 9))
merged_count.plot(column='count', cmap='YlOrBr', linewidth=0.8,
ax=ax, edgecolor='0.8', legend=True)
plt.show()
# Comparação da distribuição percentual
# Função para calcular a contagem e percentagem para cada variável
def get_distribution(column_name):
    counts = df[column_name].value_counts()
    percentagens = df[column_name].value_counts(normalize=True) *
100
    return pd.DataFrame({'Contagem': counts, 'Percentagem (%)':
percentagens})
# Obtenção da distribuição das três variáveis
proxcosta_distribution = get_distribution('PROXCOSTA')
nutsiiicosteiras_distribution =
get_distribution('NUTSIIICosteiras')

```

```

freguesiascosteiras_distribution =
get_distribution('FreguesiasCosteiras')
# Preparação dos dados para o gráfico de barras empilhadas
stacked_data = pd.DataFrame({
    'PROXCOSTA': proxcosta_distribution['Percentagem (%)'],
    'NUTSIIICosteiras': nutsiiicosteiras_distribution['Percentagem (%)'],
    'FreguesiasCosteiras':
freguesiascosteiras_distribution['Percentagem (%)']
}).fillna(0)
# Transposição dos dados
transposed_data = stacked_data.T
# Criação de um gráfico de barras empilhadas
transposed_data.plot(kind='bar', stacked=True, figsize=(15, 6),
colormap='coolwarm')
plt.title('')
plt.ylabel('Percentagem (%)')
plt.xticks(rotation=0, ha='center')
plt.legend(title='Categorias', bbox_to_anchor=(1.05, 1), loc='upper
left')
plt.show()

# Realização de um pie chart para a variável 'PME'
# Contagem das frequências de cada valor na coluna 'PME'
pme_counts = df['PME'].value_counts()
# Definição das labels e os valores para o gráfico
labels = ['Sim' if value == 1 else 'Não' for value in
pme_counts.index]
values = pme_counts.values
plt.figure(figsize=(15, 6))
plt.pie(values, labels=labels, autopct='%1.1f%%', startangle=140,
colors=['springgreen', 'coral'])
plt.title('')
plt.axis('equal')
plt.show()

# Realização de um pie chart para a variável 'USALI'

# Contar a frequência de cada valor na coluna 'USALI'
usali_counts = df['USALI'].value_counts()
# Definir as labels e os valores para o gráfico
labels = ['Sim' if value == 1 else 'Não' for value in
usali_counts.index]
values = usali_counts.values
plt.figure(figsize=(15, 6))
plt.pie(values, labels=labels, autopct='%1.1f%%', startangle=120,
colors=['springgreen', 'coral'])
plt.title('')
plt.axis('equal')

```

```

plt.show()

# Realização de uma conjunto de gráficos de barras empilhados para
# analisar padrões por NUTSIII

# Cálculo das contagens para cada combinação de 'NUTSIII' e as
# variáveis 'PME', 'USALI', e 'ESTRELAS'
pme_NUTSIII_counts = df.groupby(['NUTSIII',
    'PME']).size().unstack().fillna(0)
usali_NUTSIII_counts = df.groupby(['NUTSIII',
    'USALI']).size().unstack().fillna(0)
estrelas_NUTSIII_counts = df.groupby(['NUTSIII',
    'ESTRELAS']).size().unstack().fillna(0)
# Cálculo das proporções para cada NUTSIII, de modo que a soma das
# proporções para cada distrito seja 100%
pme_proportions =
pme_NUTSIII_counts.divide(pme_NUTSIII_counts.sum(axis=1), axis=0) *
100
usali_proportions =
usali_NUTSIII_counts.divide(usali_NUTSIII_counts.sum(axis=1),
axis=0) * 100
estrelas_proportions =
estrelas_NUTSIII_counts.divide(estrelas_NUTSIII_counts.sum(axis=1),
axis=0) * 100
# Execução de um gráfico de barras empilhadas a 100%, representando
# essas proporções
fig, axes = plt.subplots(3, 1, figsize=(15, 18), sharex=False)
pme_proportions.plot(kind='bar', stacked=True, colormap='coolwarm',
ax=axes[0], legend=True)
axes[0].set_title('PME')
axes[0].set_xlabel('')
axes[0].set_ylabel('Porcentagem (%)')
axes[0].legend(loc='upper left', bbox_to_anchor=(1, 1))
usali_proportions.plot(kind='bar', stacked=True,
colormap='coolwarm', ax=axes[1], legend=True)
axes[1].set_title('USALI')
axes[1].set_xlabel('')
axes[1].set_ylabel('Porcentagem (%)')
axes[1].legend(labels=['Não', 'Sim'], loc='upper left',
bbox_to_anchor=(1, 1))
estrelas_proportions.plot(kind='bar', stacked=True,
colormap='coolwarm', ax=axes[2], legend=True)
axes[2].set_title('ESTRELAS')
axes[2].set_xlabel('')
axes[2].set_ylabel('Porcentagem (%)')
axes[2].legend(loc='upper left', bbox_to_anchor=(1, 1))
plt.tight_layout()
plt.show()

```

```

# Realização de uma conjunto de gráficos de barras empilhados para
# analisar padrões por ESTRELAS

# Cálculo das contagens para cada combinação de 'ESTRELAS' e as
# variáveis 'PME', 'USALI'
pme_ESTRELAS_counts = df.groupby(['ESTRELAS',
                                  'PME']).size().unstack().fillna(0)
usali_ESTRELAS_counts = df.groupby(['ESTRELAS',
                                     'USALI']).size().unstack().fillna(0)
proxcoasteira_ESTRELAS_counts = df.groupby(['ESTRELAS',
                                              'NUTSIIICosteiras']).size().unstack().fillna(0)
# Cálculo das proporções para cada NUTSIIII, de modo que a soma das
# proporções para cada distrito seja 100%
pme1_proportions =
pme_ESTRELAS_counts.divide(pme_ESTRELAS_counts.sum(axis=1), axis=0)
* 100
usali1_proportions =
usali_ESTRELAS_counts.divide(usali_ESTRELAS_counts.sum(axis=1),
axis=0) * 100
proxcoasteira1_proportions =
proxcoasteira_ESTRELAS_counts.divide(proxcoasteira_ESTRELAS_counts.su
m(axis=1), axis=0) * 100
# Execução de um gráfico de barras empilhadas a 100%, representando
# essas proporções
fig, axes = plt.subplots(3, 1, figsize=(15, 18), sharex=False)
pme1_proportions.plot(kind='bar', stacked=True,
colormap='coolwarm', ax=axes[0], legend=True)
axes[0].set_title('PME')
axes[0].set_xlabel('')
axes[0].set_ylabel('Porcentagem (%)')
axes[0].legend(loc='upper left', bbox_to_anchor=(1, 1))
usali1_proportions.plot(kind='bar', stacked=True,
colormap='coolwarm', ax=axes[1], legend=True)
axes[1].set_title('USALI')
axes[1].set_xlabel('')
axes[1].set_ylabel('Porcentagem (%)')
axes[1].legend(labels=['Não', 'Sim'], loc='upper left',
bbox_to_anchor=(1, 1))
proxcoasteira1_proportions.plot(kind='bar', stacked=True,
colormap='coolwarm', ax=axes[2], legend=True)
axes[2].set_title('Costeira')
axes[2].set_xlabel('')
axes[2].set_ylabel('Porcentagem (%)')
axes[2].legend(loc='upper left', bbox_to_anchor=(1, 1))
plt.tight_layout()
plt.show()

# Criação de um mapa

```

```

# Mapeamento dos valores 'Sim' e 'Não' para 1 e 0, respectivamente,
nas colunas 'PME' e 'NUTSIIICosteiras'
mapping = {'Sim': 1, 'Não': 0}
df['PME'] = df['PME'].map(mapping)
df['NUTSIIICosteiras'] = df['NUTSIIICosteiras'].map(mapping)
# Cálculo dos percentis para a coluna 'Hotels_5km'
percentiles = np.percentile(df['Hotels_5km'], [25, 50, 75])
# Categorização dos hotéis em diferentes níveis de aglomeração com
base nos percentis
df['Aglomeracao'] = pd.cut(df['Hotels_5km'], bins=[-np.inf,
percentiles[0], percentiles[1], percentiles[2], np.inf],
labels=['Baixa', 'Média-Baixa',
'Média-Alta', 'Alta'])
# Filtro das colunas de interesse: 'USALI', 'PME',
'NUTSIIICosteiras', e 'Aglomeracao'
data_filtered = df[['USALI', 'PME', 'NUTSIIICosteiras',
'Aglomeracao']]
# Contagem do número de hotéis para cada nível de aglomeração e
para cada variável de interesse
count_data = data_filtered.groupby('Aglomeracao').agg({'USALI':
['sum', 'count'],
'PME':
['sum', 'count'],
'NUTSIIICost
eiras': ['sum', 'count']})
# Cálculo do número de hotéis com "Não" ou 0 para as variáveis de
interesse
count_data[('USALI', 'No')] = count_data[('USALI', 'count')] -
count_data[('USALI', 'sum')]
count_data[('PME', 'No')] = count_data[('PME', 'count')] -
count_data[('PME', 'sum')]
count_data[('NUTSIIICosteiras', 'No')] =
count_data[('NUTSIIICosteiras', 'count')] -
count_data[('NUTSIIICosteiras', 'sum')]
# Preparação dos dados para o gráfico
labels = count_data.index
usal_yes = count_data[('USALI', 'sum')]
usal_no = count_data[('USALI', 'No')]
pme_yes = count_data[('PME', 'sum')]
pme_no = count_data[('PME', 'No')]
nuts_yes = count_data[('NUTSIIICosteiras', 'sum')]
nuts_no = count_data[('NUTSIIICosteiras', 'No')]
# Criação do gráfico de barras horizontais com dois sentidos e
agrupado
fig, ax = plt.subplots(figsize=(16, 8))
# Definição da posição para os grupos
y_pos = np.arange(len(labels))
# Largura das barras
width = 0.2

```

```

# Ajustamento da posição para cada barra
pos_yes = y_pos - width
pos_no = y_pos + width
# Criação das barras para 'USALI' (Sim à direita e Não à esquerda)
ax.barh(pos_yes, usal_yes, width, label='USALI (Sim)',
color='blue')
ax.barh(pos_yes, -usal_no, width, label='USALI (Não)',
color='blue', alpha=0.5)
# Criação das barras para 'PME' (Sim à direita e Não à esquerda)
ax.barh(y_pos, pme_yes, width, label='PME (Sim)', color='red')
ax.barh(y_pos, -pme_no, width, label='PME (Não)', color='red',
alpha=0.5)
# Criação das barras para 'NUTSIIICosteiras' (Sim à direita e Não à
esquerda)
ax.barh(pos_no, nuts_yes, width, label='Costeira (Sim)',
color='green')
ax.barh(pos_no, -nuts_no, width, label='Costeira (Não)',
color='green', alpha=0.5)
# Configuração das legendas e rótulos
ax.set_yticks(y_pos)
ax.set_yticklabels(labels)
ax.set_xlim([-90, 90])
ax.set_xticks(list(range(-90, 91, 10)))
ax.set_xticklabels([str(abs(x)) for x in range(-90, 91, 10)])
ax.set_xlabel('Número de Unidades Hoteleiras')
ax.set_ylabel('Agglomeração')
ax.legend()

# Apresentação do gráfico
plt.show()

```

Apêndice F: Notebook_AnaliseDados_Final.ipynb