



Universidade de Aveiro
2022

**MARTA NUNES
CORREIA**

**ÍNDICE DE VULNERABILIDADE SOCIAL –
APLICAÇÃO AO CASO PORTUGUÊS**



Universidade de Aveiro
2022

**MARTA NUNES
CORREIA**

ÍNDICE DE VULNERABILIDADE SOCIAL – APLICAÇÃO AO CASO PORTUGUÊS

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Ciência de Dados para Ciências Sociais, realizada sob a orientação científica da Doutora Mara Teresa da Silva Madaleno, Professora Auxiliar do Departamento de Economia, Gestão, Engenharia Industrial e Turismo da Universidade de Aveiro.

Este trabalho é financiado por fundos nacionais através da FCT – Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projeto JUST_PLAN (PTDC/GES-OUT/2662/20).

o júri

presidente

Prof. Doutor João José Lourenço Marques
Professor Auxiliar da Universidade de Aveiro

vogais

Prof. Doutor José Manuel de Oliveira Mendes
Professor Catedrático da Faculdade de Economia da Universidade de Coimbra

Prof. Doutora Mara Teresa da Silva Madaleno
Professora auxiliar da Universidade de Aveiro

agradecimentos

Agradeço à minha orientadora, a Professora Mara Madaleno, pela partilha de conhecimento e sugestões, assim como pelo incentivo e motivação ao longo deste processo.

Aos meus amigos, por todos os momentos que passamos juntos, e por terem tornado estes últimos cinco anos nos melhores anos da minha vida.

E, por fim, à minha família, por todo o apoio que me deram ao longo do meu percurso escolar.

palavras-chave

Vulnerabilidade Social, Análise de Componentes Principais, Índice Compósito, Índice de Vulnerabilidade Social (IVS), Portugal.

resumo

A vulnerabilidade social consiste no impacto físico, económico, político ou social que uma população sofre com determinado processo desestruturante. A presente dissertação procura quantificar essa vulnerabilidade através de um índice que possa ser incorporado na ótica das políticas públicas. Este índice é obtido, com recurso à linguagem de programação *Python*, através da aplicação da Análise de Componentes Principais a dados referentes a determinados indicadores da população e infraestruturas dos municípios de Portugal Continental, entre os anos 2009 e 2019. Os valores resultantes deste índice são representados em cartogramas, a partir dos quais é possível identificar os municípios e aglomerados territoriais mais vulneráveis e verificar que, com o passar dos anos, a vulnerabilidade em Portugal tende a apresentar uma distribuição mais homogénea. Os resultados deste estudo poderão constituir uma ferramenta útil de auxílio para os decisores de política, aquando da tentativa de combater os efeitos adversos desta vulnerabilidade, ou seja, no momento de criação e execução de medidas que permitam evitar e reverter tais efeitos.

keywords

Social Vulnerability, Principal Component Analysis, Composite Index, Social Vulnerability Index (SVI), Portugal.

abstract

Social vulnerability consists of a population's physical, economic, political, or social impact caused by a specific destructuring process. The present dissertation aims to quantify this vulnerability through an index that can be incorporated into the standpoint of public policies. This index is obtained through the application of Principal Component Analysis, using the Python programming language, to a dataset referring to specific population and infrastructure indicators of the municipalities of mainland Portugal between the years 2009 and 2019. The values resulting from this index are represented in cartograms, from which it is possible to identify the most vulnerable municipalities and territorial agglomerations and to verify that, over the years, the vulnerability in Portugal tends to present a more homogeneous distribution. The results of this study could constitute a valuable tool to help policymakers combat the adverse effects of this vulnerability, that is, when creating and implementing measures to avoid and reverse such effects.

Índice

Índice de Equações	iii
Índice de Figuras	iv
Índice de Tabelas	v
Lista de Abreviaturas e Siglas	vi
1. Introdução	1
1.1. Definição do Problema.....	1
1.2. Motivação e Relevância do Tema.....	1
1.3. Objetivos	2
1.4. Estrutura.....	2
2. Revisão de Literatura	4
2.1. A Vulnerabilidade	4
2.2. A Vulnerabilidade Social	7
2.3. Medição da Vulnerabilidade Social	8
2.4. Determinantes da Vulnerabilidade Social	12
2.4.1. População	12
2.4.2. Infraestruturas	14
3. Dados e Metodologia	17
3.1. Análise Descritiva dos Dados	17
3.1.1. Seleção e Extração dos Dados.....	18
3.1.2. Descrição e Análise dos Dados	20
3.2. Pré-Processamento dos Dados	21
3.2.1. Limpeza dos Dados	22
3.2.2. Transformação dos Dados.....	22
3.3. Índice Compósito	28
3.3.1. Viabilidade dos Dados	28
3.3.2. Determinação do Número de Componentes.....	29
3.3.3. Análise de Componentes Principais	30

3.3.4.	Criação do Índice	32
3.3.5.	Normalização do Índice.....	34
3.4.	Representação Geográfica	35
3.4.1.	Classificação	35
3.4.2.	Cartogramas	36
3.5.	Análise de Robustez	37
4.	Resultados	39
4.1.	Pré-Processamentos dos Dados	39
4.1.1.	Limpeza dos Dados	39
4.2.	Índice de Vulnerabilidade Social	40
4.2.1.	Viabilidade dos Dados	40
4.2.2.	Determinação do Número de Componentes	40
4.2.3.	Obtenção do Índice	43
4.2.4.	Descrição do Índice.....	44
4.2.5.	Índice de Vulnerabilidade Social	45
4.2.6.	Cartogramas	48
4.3.	Análise de Robustez	52
5.	Conclusão	58
5.1.	Principais Conclusões.....	58
5.2.	Limitações	60
5.3.	Desenvolvimentos Futuros.....	60
	Referências.....	62
	Anexos	69

Índice de Equações

Equação 1 - Fórmula de cálculo do indicador SMED.....	23
Equação 2 - Fórmula de cálculo do indicador TDES.	24
Equação 3 - Fórmula de cálculo do indicador EE.	24
Equação 4 - Fórmula de transformação dos indicadores.....	25
Equação 5 - Fórmula de normalização de indicadores com relação positiva.	26
Equação 6 - Fórmula de normalização de indicadores com relação negativa.....	26
Equação 7 - Matriz dos indicadores a serem usados na ACP.	31
Equação 8 - Fórmulas de cálculo de cada CP.	31
Equação 9 - Fórmula de cálculo de um índice através da ACP.....	33
Equação 10 - Fórmula de atribuição de pesos a cada CP.....	34
Equação 11 - Fórmula de normalização do índice resultante.....	35
Equação 12 - Fórmula de cálculo do IVS para o ano 2009.....	43
Equação 13 - Fórmula de cálculo do novo IVS para a análise de robustez.....	53

Índice de Figuras

Figura 1 - Vulnerabilidade e componentes.....	5
Figura 2 - Exemplo de um Scree Plot.	30
Figura 3 - <i>Scree Plots</i> dos anos de 2009 e 2019.....	41
Figura 4 - Mapa nacional do IVS para 2009.....	48
Figura 5 - Mapa nacional das classes do IVS para 2009.....	48
Figura 6 - Mapa nacional do IVS para 2019.....	49
Figura 7 - Mapa nacional das classes do IVS para 2019.....	49
Figura 8 - Valores do IVS para o distrito do Porto em 2009 e 2019.....	50
Figura 9 - Classes do IVS para o distrito do Porto em 2009 e 2019.....	50
Figura 10 - Valores do IVS para o distrito de Faro em 2009 e 2019.....	51
Figura 11 - Classes do IVS para o distrito de Faro em 2009 e 2019.	51
Figura 12 - Valores do IVS original para o distrito de Setúbal.	55
Figura 13 - Valores do IVS de análise de robustez para o distrito de Setúbal.....	55
Figura 14 - Mapa nacional das classes do IVS original.	56
Figura 15 - Mapa nacional das classes do IVS de análise de robustez.....	56

Índice de Tabelas

Tabela 1 - Resiliência e Exposição.....	6
Tabela 2 - Resiliência e Sensibilidade.....	6
Tabela 3 - Resumo dos estudos analisados.....	11
Tabela 4 - Indicadores analisados.....	16
Tabela 5 - Indicadores extraídos da PORDATA.....	19
Tabela 6 - Descrição dos Dados.	20
Tabela 7 - Relação estabelecida entre os indicadores e a vulnerabilidade social.	27
Tabela 8 - Indicadores adicionados para análise de robustez.	38
Tabela 9 - Resultados dos Testes de Viabilidade dos Dados para o ano de 2009.	40
Tabela 10 - Autovalores das 15 Componentes Principais para o ano de 2009.....	41
Tabela 11 - VEA das 15 Componentes Principais para o ano de 2009.	42
Tabela 12 - VEA das 4 Componentes Principais selecionadas.....	43
Tabela 13 - Autovalores e Variâncias das CPs para o ano de 2009.....	43
Tabela 14 - Medidas de Centralidade e Dispersão do IVS.....	44
Tabela 15 - Medidas de centralidade do IVS distrital.....	45
Tabela 16 - IVS para os distritos do Porto e de Faro.	46
Tabela 17 - Distribuição das classes pelos municípios.	47
Tabela 18 - Viabilidade dos dados de análise de robustez.	52
Tabela 19 - VEAs das CPs da análise de robustez.....	53
Tabela 20 - Medidas de centralidade e dispersão para a análise de robustez.	54
Tabela 21 - Medidas de centralidade distritais para a análise de robustez.....	54
Tabela 22 - Distribuição das classes pelos municípios para a análise de robustez.....	56

Lista de Abreviaturas e Siglas

ACP	Análise de Componentes Principais
ANC	Indicador referente ao número de alojamentos não clássicos
ATIVA	Indicador referente à população ativa
CPs	Componentes Principais
DENS	Indicador referente à densidade de alojamentos familiares
DEP	Indicador referente ao índice de dependência
DES	Indicador referente à população desempregada
DIF	Indicador referente à diferença salarial
EB	Indicador referente ao número de escolas de ensino básico
EPE	Indicador referente ao número de escolas de ensino pré-escolar
ES	Indicador referente ao número de escolas de ensino secundário
EUA	Estados Unidos da América
FARM	Indicador referente ao número de farmácias
HOSP	Indicador referente ao número de hospitais
ILIT	Indicador referente ao índice de iliteracia
INCAP	Indicador referente ao número de beneficiários da bonificação por deficiência da segurança social
IVS	Índice de Vulnerabilidade Social
KMO	Kaiser-Meyer-Olkin
MAIS65	Indicador referente ao número de indivíduos com mais de 65 anos
Máx	Valor máximo
Méd	Valor Médio
MENOS5	Indicador referente ao número de indivíduos com menos de 5 anos
Min	Valor mínimo
NAT	Indicador referente à taxa de natalidade
RESID	Indicador referente à população residente
SMED	Indicador referente ao salário médio
SMIN	Indicador referente ao salário mínimo
SoVI	Social Vulnerability Index
SS	Indicador referente ao número de beneficiários do abono de família da segurança social
TDES	Indicador referente à taxa de população desempregada
UNI	Indicador referente ao número de indivíduos com nível de ensino superior
VEA	Variância Explicada Acumulada

1. Introdução

1.1. Definição do Problema

Cada vez mais, a medição e quantificação da vulnerabilidade social é utilizada aquando da criação e aplicação de políticas públicas. No entanto, Portugal, assim como diversos outros países, ainda não inseriu a análise da vulnerabilidade social nas suas resoluções de planeamento e gestão territorial.

Quando se fala de vulnerabilidade social, entende-se o impacto físico, económico, político ou social que determinada população sofre com um dano ou processo desestruturante, quer de origem natural, quer de origem antrópica (Mendes & Tavares, 2011). Desta forma, a quantificação da vulnerabilidade social possibilita o conhecimento e a identificação de territórios mais vulneráveis, tornando, assim, possível a priorização e orientação de medidas de prevenção e políticas públicas para tais territórios (Sigorro, 2018).

Pode-se concluir, então, que, quando integrada na definição de políticas públicas, a quantificação da vulnerabilidade social permite que sejam obtidos conhecimentos territoriais mais aprofundados relativamente à ótica de prevenção e redução dos riscos, de otimização de recursos, e de definição de estratégias de gestão territorial (Mendes & Tavares, 2011).

1.2. Motivação e Relevância do Tema

Esta análise foi motivada pela quantidade reduzida de estudos que procuram comparar a vulnerabilidade social de determinado território com outro. Porém, a quantidade de estudos que estendem esta análise ao longo de um período temporal é ainda mais escassa.

Os estudos de Fraser (2021), Frigerio et al. (2018), e Melidis e Tzagkarakis (2021) realizam uma análise da vulnerabilidade social em determinados territórios, ao longo de diversos períodos temporais. No estudo de Fraser (2021), é efetuada uma quantificação da vulnerabilidade social no Japão, a nível municipal, para os anos compreendidos entre 2000 e 2017, de onde se conclui que, por norma, as regiões mais desenvolvidas apresentam maior vulnerabilidade social que as regiões rurais. Melidis e Tzagkarakis (2021), por sua vez, analisam a evolução da vulnerabilidade social na Grécia, ao longo dos anos de 2008 e 2017, onde se verificou que a crise financeira europeia e o período de austeridade provocaram graves consequências sobre a população grega, algumas ainda bastante presentes nos dias

de hoje. Por fim, Frigerio et al. (2018) avaliam a evolução da vulnerabilidade social, em Itália, durante os anos em que foi realizada a recolha censitária, ou seja 1991, 2001 e 2011, onde se demonstrou que a percentagem de municípios com níveis intermédios de vulnerabilidade social aumentou, e a percentagem de municípios com níveis mais elevados de vulnerabilidade social diminuiu.

Quanto a estudos realizados em território português, Mendes (2009) analisou a vulnerabilidade social da região centro de Portugal Continental. Já Guillard-Gonçalves et al. (2014) analisam essa componente ao nível das freguesias da sub-região da Grande Lisboa, e Gomes (2018) ao nível do arquipélago da Madeira. Existem, ainda, estudos nacionais realizados com detalhe ao nível municipal, como, por exemplo, os estudos de Mendes et al. (2011) e de Marques et al. (2017), no entanto, não foi possível identificar um estudo que realizasse esta análise a nível municipal e a conciliasse com uma evolução ao longo de determinado período temporal.

Desta forma, a relevância do estudo a ser realizado nesta dissertação passa pela análise da vulnerabilidade social, ao longo de 11 anos, de 2009 a 2019, ao nível dos municípios de Portugal Continental.

1.3. Objetivos

Este estudo procura identificar territórios ou pequenos aglomerados de Portugal que apresentem maiores níveis de vulnerabilidade social e, conseqüentemente, auxiliar a criação, orientação e priorização da tomada de medidas de combate e prevenção da vulnerabilidade social, e a aplicação de políticas públicas, com o intuito de mitigar as disparidades sentidas ao nível da qualidade de vida da sociedade portuguesa. Para isto, será necessário quantificar a vulnerabilidade social através de um índice, o Índice de Vulnerabilidade Social, ou IVS, a ser criado durante este estudo.

1.4. Estrutura

Esta dissertação foi organizada em quatro capítulos principais: a Revisão de Literatura, os Dados e Metodologia, os Resultados, e a Conclusão.

Após este capítulo introdutório, é apresentada a Revisão de Literatura onde é descrito todo o enquadramento teórico e conceitos mais relevantes deste estudo, e explorada a

principal literatura existente sobre esses conceitos, assim como o tema da vulnerabilidade social.

No capítulo dos Dados e Metodologia, é descrito todo o processo metodológico a aplicar neste estudo para que seja criado e calculado um índice de quantificação da vulnerabilidade social dos municípios portugueses, entre os anos 2009 e 2019. Primeiramente, são apresentados os dados e indicadores a serem utilizados nesta quantificação e, de seguida, é efetuada uma breve análise dos mesmos, assim como o seu pré-processamento. Posteriormente, é analisada a viabilidade dos dados e, caso confirmada essa viabilidade, é aplicado o método de Análise de Componentes Principais, ou ACP, através do qual se procede com a aglomeração dos resultados num índice compósito, representativo da vulnerabilidade social. Com os valores da quantificação dessa componente, são criadas categorias, ou níveis de classificação da vulnerabilidade, a serem associados a cada município, e, seguidamente, mapeados os resultados obtidos em cartogramas. O último passo deste capítulo consiste na realização de uma análise de robustez com o intuito de estabelecer uma comparação dos resultados obtidos com o IVS criado com a metodologia descrita e um IVS com leves alterações no conjunto selecionado de indicadores.

O capítulo Resultados reúne todos os resultados obtidos através do processo metodológico descrito anteriormente, assim como algumas discussões dos mesmos, acompanhadas de justificações encontradas na literatura analisada.

Por fim, no último capítulo deste estudo, após uma revisão de todo o processo efetuado, são apresentadas as principais conclusões resultantes deste, assim como algumas limitações encontradas aquando da realização do mesmo, e motivações para investigações futuras.

2. Revisão de Literatura

2.1. A Vulnerabilidade

O conceito de vulnerabilidade é bastante familiar a qualquer indivíduo. No entanto, a sua definição gera algumas indecisões, não existindo, assim, uma definição concisa na literatura (Hufschmidt, 2011). Sabe-se, porém, que este conceito teve origem nas ciências sociais e da natureza, focado na temática da segurança alimentar e dos desastres naturais. Contudo, tem vindo a ser aplicado a outras inúmeras áreas científicas (Vincent, 2004). De uma forma bastante geral, a vulnerabilidade pode ser entendida como a capacidade de um indivíduo, ou população, antecipar, resistir, lidar e responder a um impacto, ou seja, corresponde à tendência dessa entidade ser afetada ou estar suscetível a danos (Paul, 2013; Wisner et al., 1994).

Outra definição do conceito de vulnerabilidade foi desenvolvida por Watts e Bohle (1993) onde a explicam em termos de exposição, capacidade e potencialidade e que, dessa forma, a resposta à vulnerabilidade passa por reduzir a exposição, aumentar a capacidade de reação e melhorar a potencialidade da recuperação. Os autores acrescentam ainda que as consequências da vulnerabilidade deverão ser minimizadas. Complementando esta visão, Wisner et al. (1994) acrescentam que a vulnerabilidade envolve uma combinação de fatores determinantes do grau do risco ao qual a sociedade é exposta. Estes fatores são explícitos na definição de Bohle (2001) que vê a vulnerabilidade como uma medida do bem-estar de uma sociedade, incorporando fatores ambientais, sociais, económicos e políticos. Acrescenta ainda que a vulnerabilidade é um conceito multidimensional, em grande parte, determinado pelas capacidades políticas, económicas e institucionais de determinada sociedade, num determinado período temporal. Por fim, Wolf et al. (2013) resumizam o conceito de vulnerabilidade numa medida de risco futuro.

Após esta breve exploração e análise da literatura existente acerca deste conceito, facilmente se identifica uma divisão da mesma entre duas abordagens teóricas: uma focada no estudo dos perigos e riscos associados à vulnerabilidade, enquanto a outra abordagem se foca na componente social e humana desse mesmo conceito (Wisner et al., 1994). A primeira abordagem, denominada de Escola de Pensamento dos Riscos Naturais, vê o conceito de vulnerabilidade como o risco de exposição de um sistema a um perigo. Por outro lado, a segunda abordagem, a Escola de Pensamento da Ecologia Humana e Economia Política, entende a vulnerabilidade como a exposição de uma determinada unidade social,

tendo em conta as suas estruturas económicas, políticas e sociais que constituem a sua comunidade (Vincent, 2004).

No entanto, Cutter (2003) e Comfort et al. (1999) unificam estas duas abordagens, explicando como a vulnerabilidade incorpora tanto elementos de exposição como elementos de propensão ao risco, isto é, incorpora as circunstâncias que colocam uma população em risco e as circunstâncias que alteram a capacidade da população responder e recuperar desse risco. Assim, facilmente se percebe que a vulnerabilidade está inerente e inevitavelmente associada aos conceitos de risco, exposição e resiliência.

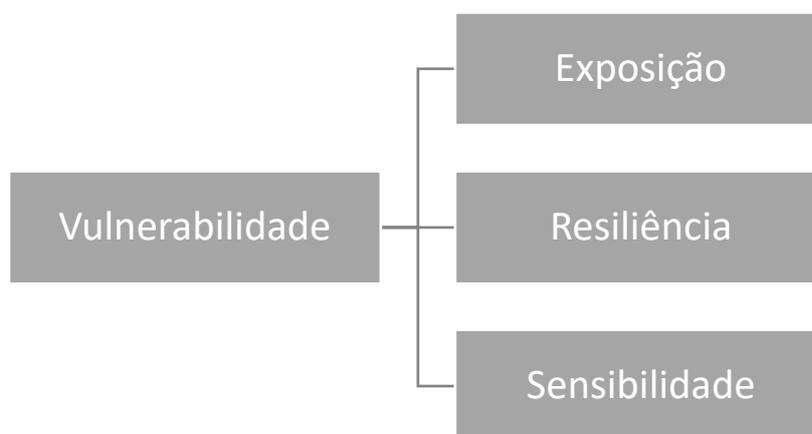


Figura 1 - Vulnerabilidade e componentes.

Fonte: U.S. Department of Transportation (2015). Tradução Própria.

Em primeiro lugar, analisa-se o conceito de exposição. Este conceito é definido, por Balica et al. (2012) e Frischen et al. (2020), como a presença de um sistema, como comunidades, ecossistemas, serviços, recursos ou infraestruturas, em locais e zonas que podem ser afetados de forma adversa por algum perigo. No artigo de Gallopín (2006), a exposição encontra-se definida como o grau, duração e extensão em que o sistema em causa está em contacto ou sujeito a uma perturbação.

O conceito de resiliência, segundo Pelling (2003), consiste na capacidade de adaptação de um sistema potencialmente exposto a uma perturbação, resistindo ou modificando-se com o objetivo de manter ou atingir um nível aceitável de funcionamento estrutural. A esta definição, Afonso (2021) acrescenta que a resiliência passa pela capacidade de adaptação e recuperação do estado inicial de um sistema após um choque, complementando com a visão de Adger (2006) onde a resiliência é vista como a capacidade de evolução e resolução desse mesmo sistema.

Outro conceito associado à vulnerabilidade, é o conceito de sensibilidade, como se pode observar na **Figura 1**. Apesar da sua definição não ser tão clara como as anteriores, a sensibilidade é vista como o grau de modificação ou afetação de um sistema a perturbações ou, mais especificamente, o ponto até onde esse sistema é capaz de absorver as consequências da perturbação, sem sofrer danos de longo prazo ou alterações significativas (Adger, 2006).

No estudo do National Research Council (2001), é referido o facto de serem a exposição e a resiliência que afetam a vulnerabilidade, mais concretamente, que influenciam o nível da vulnerabilidade. Sharma et al. (2000), partilhando essa opinião, classificam a relação entre estes conceitos de acordo com a **Tabela 1**.

Tabela 1 - Resiliência e Exposição.

		Resiliência	
		Alta	Baixa
Exposição	Alta	Vulnerabilidade baixa	Vulnerabilidade alta
	Baixa	Vulnerabilidade muito baixa	Vulnerabilidade baixa

Fonte: Sharma et al. (2000). Tradução própria.

Na **Tabela 1**, é exposta a forma como determinada comunidade e as suas componentes físicas e naturais são perceptíveis como vulneráveis caso exista exposição a um risco. Todavia, esta vulnerabilidade irá variar de acordo com o nível de resiliência dessa comunidade.

Neste sentido, é também analisada, no estudo de Davies (1996), a influência da sensibilidade na vulnerabilidade, acompanhada da resiliência, através de uma tabela semelhante, representada pela **Tabela 2**.

Tabela 2 - Resiliência e Sensibilidade.

		Resiliência	
		Alta	Baixa
Sensibilidade	Alta	Vulnerável	Muito vulnerável
	Baixa	Não vulnerável	Vulnerável

Fonte: Davies (1996). Tradução própria.

Quanto ao conceito de risco, este é interpretado por Cutter (2001) como a perceção de um perigo em termos dos danos causados por este. Alwang et al. (2001) exploram mais este

conceito, afirmando que é caracterizado por uma distribuição de probabilidades, conhecida, ou não, de eventos que, por sua vez, são caracterizados pela sua magnitude, frequência e duração.

A análise da vulnerabilidade é, então, aplicável a estudos de inúmeras áreas, como, por exemplo, a Economia, a Sociologia, a Saúde, o Ambiente, a Geografia, entre outras. Como cada uma destas áreas se foca em diferentes riscos e fatores, entende-se que o conceito de vulnerabilidade seja manipulado e modificado de forma que se adeque ao estudo em causa. Desta forma, a criação de uma definição mais concreta e globalmente aceite do conceito de vulnerabilidade consiste numa difícil tarefa com escassa probabilidade de ser tornada realidade (Alwang et al., 2001; Paul, 2013).

Como solução a este problema, são então formuladas inúmeras definições de vulnerabilidade, cada uma focada nos riscos, fatores e objetivos da área em causa, como, por exemplo, a vulnerabilidade económica, a vulnerabilidade social, a vulnerabilidade ambiental, entre outras.

Uma vez que este estudo terá incidência na vulnerabilidade social, importa estudar mais detalhadamente este conceito.

2.2. A Vulnerabilidade Social

A vulnerabilidade social, como notado anteriormente, constitui uma das diversas vertentes da vulnerabilidade. Esta vertente é comumente descrita através de características individuais de uma sociedade, por exemplo, a idade, raça, saúde, rendimento, emprego, entre outras. Desta forma, a vulnerabilidade social resulta da conjugação das desigualdades sociais com as desigualdades de lugar. Enquanto as primeiras dizem respeito aos fatores sociais que influenciam ou manipulam a suscetibilidade da sociedade ao dano em causa e controlam a sua capacidade de resposta ao mesmo, as últimas dizem respeito às características da comunidade e do seu meio envolvente, como o nível de urbanização, a sua taxa de crescimento e a sua saúde económica, por exemplo (Cutter et al., 2003).

Outra definição desta vertente da vulnerabilidade é descrita por Mavhura et al. (2017), que referem como os fatores ou processos físicos, sociais, económicos e ambientais contribuem para o incremento da suscetibilidade de um sistema às consequências de uma perturbação. Mendes (2009) define a vulnerabilidade social como um estado de fragilidade coletiva que reflete as condições de desigualdade social que antecedem a ocorrência de

uma perturbação. Estas condições estão, em grande parte, relacionadas com a qualidade do tecido social, nomeadamente com fatores como a pobreza, a idade, o género, a classe social, a literacia, e a experiência acumulada para lidar com essas perturbações, como referido anteriormente. No estudo do National Research Council (2001), a vulnerabilidade social é vista como a capacidade da sociedade lidar com eventos de risco e perturbações, traduzidas pelos aumentos de potenciais perdas e reduções da capacidade de recuperação, em grande parte provocados pelos processos sociais, económicos, históricos e políticos dessa sociedade tomados pelas ações e decisões da sociedade. A vulnerabilidade social é vista como um estado potencial das sociedades que pode afetar a forma como estas experienciam perigos e perturbações, sendo que este estado depende das interações dinâmicas dos processos já referidos (Comfort et al., 1999; Wisner et al., 1994).

Um sistema social vulnerável é aquele que se encontra mais exposto a perturbações, com uma limitada capacidade tanto de reação, como de recuperação (Bohle et al., 1994). Diversos estudos e investigações mostram que os grupos sociais marginalizados sofrem maiores consequências com desastres e perturbações, podendo, desta forma, ser estabelecida uma relação entre um elevado nível de vulnerabilidade com as populações mais pobres, politicamente marginalizadas ou excluídas da sociedade (Chakraborty et al., 2005).

Cutter et al. (2003) frisam a importância da compreensão e investigação dos principais motivos, processos e fatores que causam de facto a vulnerabilidade social. Com isto, considera-se relevante identificar e estudar, nos capítulos seguintes, como calcular a vulnerabilidade social de determinada população e quais os fatores e condições que estabelecem algum tipo de relação com a vulnerabilidade.

2.3. Medição da Vulnerabilidade Social

Antes de ser possível tomar qualquer medida de combate ou prevenção da vulnerabilidade social, é necessário conseguir mensurá-la. Para isto, foram desenvolvidos inúmeros estudos que serão explorados e analisados ao longo deste subcapítulo.

A vulnerabilidade social é classificada como uma variável latente, isto é, representa uma variável que não é observada diretamente e que depende de uma entidade ou lugar físico. Desta forma, para que a vulnerabilidade social seja quantificada, é necessário recorrer a procedimentos estatísticos incidentes sobre a informação de variáveis que podem ser

medidas e observadas de forma direta como, por exemplo, atributos demográficos de determinado território (Spielman et al., 2020).

O método de quantificação da vulnerabilidade social de determinada população mais comum consiste na criação de um índice. Por definição, um índice representa um sistema numérico utilizado para comparar valores de determinados indicadores que variam de acordo com algum critério (Cambridge Dictionary, 2022). Entre as diversas tipologias de índices, existem os índices sociais que, por norma, combinam um elevado número de variáveis relacionadas com temas específicos de determinada sociedade num valor único. Este tipo de índices permitem agrupar a informação anteriormente dispersa em inúmeras fontes de informação num só indicador, possibilitando, não só uma visão mais abrangente da situação em causa, como também comparações dessa mesma situação ao longo do tempo. Este facto acaba por tornar os índices sociais numa ferramenta bastante atrativa ao interesse público, no entanto, esta ferramenta pode também deter algumas inconveniências. Ao acumular informação numa só componente, o índice poderá omitir problemas imergentes através das suas interpretações simplistas, assim como encobrir fortes desigualdades presentes na informação acumulada (Krause, 2019). Com isto, não se pretende desencorajar a utilização de índices sociais nos mais diversos estudos, mas realçar a importância da observação dos elementos individuais que compõem o índice em causa, assim como a observação do índice em si.

O primeiro instrumento de medição da vulnerabilidade social foi desenvolvido num estudo elaborado por Susan Cutter, Bryan Buroff e Lynn Shirley em 2003. Neste estudo, os autores criaram, através de dados demográficos e socioeconómicos dos condados dos Estados Unidos da América (EUA), um índice de vulnerabilidade social e perigos ambientais, o SoVI. Este índice foi criado com recurso a 42 variáveis que, através de método de Análise de Componentes Principais, ou ACP, foram reduzidas a apenas 11 componentes a partir das quais foi calculado o índice pretendido. Neste estudo, os autores conseguiram concluir que este indicador é bastante funcional e próximo da realidade quando utilizado como uma medida de comparação espacial reforçando, porém, o facto de que são ainda necessárias algumas melhorias e ajustes como, por exemplo, a adaptação do índice a uma comparação temporal (Cutter et al., 2003).

No estudo de Flanagan et al. (2011) é também desenvolvido um índice de vulnerabilidade social, desta vez, através de 15 variáveis extraídas para todos os setores dos censos do ano 2000 dos EUA. Estas 15 variáveis foram ordenadas e, posteriormente, foi criado um *ranking*, através da classificação dos valores das variáveis em percentis, cuja soma permitiu obter o

índice de vulnerabilidade social. Com este índice, os autores pretendem auxiliar os agentes responsáveis pela garantia da segurança e bem-estar da sociedade uma vez que este permite identificar a localização de comunidades socialmente vulneráveis e, desta forma, transmitir essa informação de forma mais eficiente, para que fossem tomadas as medidas necessárias o mais rapidamente possível. Os autores concluíram ainda que o índice calculado é bastante flexível e poderá ser aplicado a diversos outros contextos. Não obstante, identificaram também algumas limitações como as céleres alterações demográficas de populações de menor escala e as novas subdivisões geográficas que podem alterar o esquema populacional verificado anteriormente, assim como o uso de dados dos censos nacionais que limita o horizonte temporal da análise efetuada.

O seguinte estudo analisado é o estudo de Solangaarachchi et al. (2012) onde é calculado um índice de vulnerabilidade social para os distritos incluídos nos censos australianos, no ano 2006, dos quais foram utilizadas 51 variáveis incluídas nas temáticas socioeconómicas e demográficas da região definida. Estas variáveis foram reduzidas a cerca de 25 após as restantes serem eliminadas com um teste de multicolinearidade e, posteriormente, foi ainda aplicado o método ACP de onde resultaram 6 componentes. O índice de vulnerabilidade social foi, então, criado através destas componentes e permitiu estabelecer comparações da vulnerabilidade social entre os distritos da Austrália. Neste estudo, os autores defendem a utilidade do instrumento calculado e que esta metodologia pode ser incorporada a diferentes estudos de forma a possibilitar uma melhor perceção da vulnerabilidade social de determinadas sociedades e desenvolver programas mais eficazes de combate à mesma.

Num estudo mais recente, Cumberbatch et al. (2020) calculam um índice de vulnerabilidade social através de 8 variáveis retiradas dos censos de 2012 de Barbados. À semelhança do efetuado no estudo de Flanagan et al. (2011), estas 8 variáveis são ordenadas e classificadas em percentis que, somados, constituem o índice em causa. Os mapeamentos dos resultados obtidos permitem identificar as unidades geográficas mais e menos socialmente vulneráveis e, desta forma, dirigir políticas e medidas de combate a este problema.

No estudo de Avila-Vera et al. (2012) são utilizados e manipulados 12 indicadores, extraídos dos censos de 2010, para calcular o índice de vulnerabilidade social do México. Este índice é obtido através do método ACP e permite identificar quais os municípios cuja vulnerabilidade social é mais elevada e, conseqüentemente, os que mais necessitam de algum tipo de intervenção política e comunitária. Os autores defendem que apenas através

de políticas realmente afetivas e eficientes podem melhorar as condições de vida da sociedade e que este índice é uma ferramenta bastante útil para dirigir essas políticas.

Por fim, no estudo de Davino et al. (2021), a proposta de medição da vulnerabilidade social na zona urbana de Roma passa pela criação de um índice através da sumarização de diversos indicadores observáveis num índice único ou compósito. Esta sumarização é realizada pelo método ACP onde os 21 indicadores selecionados, a partir dos censos nacionais, foram reduzidos a 7 componentes correspondentes a 7 dimensões da vulnerabilidade social, que, por sua vez foram agrupadas num único valor, criando assim um índice compósito como uma medida quantitativa da vulnerabilidade social. Neste estudo, os autores relembram uma das limitações dos índices, onde estes podem levar à omissão de informações relevantes, e mencionam que a falta de dados ou a indisponibilidade dos mesmos limitou a análise efetuada. Porém, os autores concluem que, de um ponto de vista prático, o índice calculado pode ser tomado em conta aquando da formulação de políticas de combate à vulnerabilidade social, dirigi-las para as zonas mais vulneráveis e, desta forma, melhorar as condições de vida dessas zonas.

Para dar este subcapítulo como concluído, apresenta-se, de seguida, uma tabela resumo, a **Tabela 3**, onde se sumarizam as informações transmitidas pelos estudos analisados ao longo desta revisão de literatura.

Tabela 3 - Resumo dos estudos analisados.

Estudo	Perigo	Contexto espacial	Divisões geográficas	Nº de indicadores	Método utilizado
Cutter et al. (2003)	Perigos Ambientais	EUA	Condados	42	ACP
Flanagan et al. (2011)	Geral	EUA	Setores censitários	15	<i>Ranking</i>
Cumberbatch et al. (2020)	Geral	Barbados	Distrito de enumeração	8	<i>Ranking</i>
Solangaarachchi et al. (2012)	Incêndios	Sydney, Austrália	Setores censitários	25	ACP
Avila-Vera et al. (2012)	Geral	México	Municípios	12	ACP
Davino et al. (2021)	Geral	Roma, Itália	Municípios	21	ACP

Fonte: Elaboração própria.

Nota: ACP - Análise de Componentes Principais.

2.4. Determinantes da Vulnerabilidade Social

Ao longo desta secção, têm vindo a ser referidos alguns fatores e indicadores que alteram a vulnerabilidade de determinado sistema ou entidade e que são utilizados aquando da construção de um índice de vulnerabilidade social. Neste capítulo procura-se explorar, de forma mais detalhada, esses fatores e identificar o tipo de relação que estabelecem com a vulnerabilidade social desse mesmo sistema.

Os indicadores a serem apresentados são divididos em duas grandes categorias que caracterizam determinada sociedade: indicadores relativos à população, e indicadores relativos às infraestruturas.

2.4.1. População

Nesta primeira secção são descritos indicadores que refletem as características da população e que são considerados relevantes para a definição da vulnerabilidade da mesma, como, por exemplo, a idade, saúde, a educação, o rendimento e a situação de emprego dos indivíduos. Desta forma, estes indicadores, para além de refletirem as características demográficas da sociedade, refletem também o seu estado socioeconómico.

Começando pelos indicadores mais comuns, e presentes na grande maioria da literatura analisada, é atribuída elevada relevância à idade da população. Nos estudos de Cutter et al. (2003), Guillard-Gonçalves et al. (2014), Rabby et al. (2019), Schmidlein et al. (2008), e Tasnuva et al. (2020) são utilizados dois indicadores referentes à idade da população. O primeiro referente ao número de indivíduos com menos de 5 anos, e o segundo referente ao número de indivíduos com mais de 65 anos, ou seja, referentes aos extremos do espetro etário de uma população, sendo definidas como as faixas etárias mais frágeis e, por essa mesma razão, mais vulneráveis. A idade das crianças com menos de 5 anos, apesar de moldar as suas necessidades, expõe-nas a potenciais novos riscos uma vez que, para além de serem completamente dependentes e requererem cuidados constantes, são particularmente sensíveis ao meio e pessoas que as rodeiam (Organization for Economic Cooperation and Development, 2019). Por outro lado, os idosos, com o passar do tempo, tornam-se mais propícios a desenvolver restrições ou problemas de mobilidade, aumentando a sua falta de resiliência e celeridade de recuperação de um potencial perigo (Ngo, 2001). Neste sentido, um maior número quer de crianças com menos de 5 anos, quer

de adultos com mais de 65 anos, reflete-se numa maior vulnerabilidade da população, sendo estabelecida uma relação positiva com a mesma.

Ainda relacionado com a idade da população, o indicador referente à porção de população dependente é também bastante utilizado na literatura, nomeadamente nos estudos de Chen et al. (2013), Cutter et al. (2003), Guillard-Gonçalves et al. (2014), Rabby et al. (2019), e Tasnuva et al. (2020). Este indicador diz respeito às faixas etárias e indivíduos social e economicamente dependentes de outros indivíduos que acentuam a necessidade de apoio no período pós-desastre do indivíduo responsável, contribuindo assim para o aumento da vulnerabilidade geral da população (Cutter et al., 2003).

Outro indicador comumente utilizado, também relativo às características demográficas da população, é a taxa de natalidade da mesma. Esta taxa reflete o número de nascimentos de bebés numa determinada população e é considerado como um fator que afeta a resiliência e a recuperação dos perigos, aumentando a vulnerabilidade da população, como verificado estudos de Cutter et al. (2003), e de Mendes (2009). Isto deve-se ao facto das famílias com recém-nascidos terem, na maioria das vezes, recursos financeiros limitados para o cuidado destes, consistindo este fator uma nova despesa. Assim, estas famílias necessitam de conciliar este aspeto com responsabilidades e necessidades anteriores, tornando-os menos resilientes e preparados para qualquer perigo (Gomes, 2018).

Passando para o estado socioeconómico da população, os indicadores mais utilizados dizem respeito à taxa de desemprego e ao salário médio da população. A taxa de desemprego reflete o número de indivíduos, em idade ativa, desempregados, de determinada população, e é utilizado nos estudos de Cumberbatch et al. (2020), Cutter et al. (2003), Dintwa et al. (2019), Mendes (2009), Rabby et al. (2019), e de Schmidtlein et al. (2008). Segundo Gomes (2018), níveis mais elevados de riqueza pessoal estão relacionados com uma maior resiliência e capacidade de recuperação de desastres e, conseqüentemente, os indivíduos tendem a possuir menos recursos, o que se traduz numa maior vulnerabilidade social. Quanto ao salário médio, o pensamento é o mesmo, a riqueza permite que a população absorva e recupere de perdas mais rapidamente, tornando-a menos vulnerável (Cutter et al. 2003). Assim, conclui-se que uma maior taxa de desemprego contribui para o aumento da vulnerabilidade, e um maior salário médio para a diminuição da mesma.

Também relacionado com o estado socioeconómico da população, é utilizado, nos estudos de Guillard-Gonçalves et al. (2014), Rabby et al. (2019), e Tasnuva et al. (2020), um indicador relativo à educação da população: a percentagem de população iliterata ou analfabeta. Por norma, os indivíduos incluídos nesta percentagem não têm possibilidades

de alcançar uma colocação de emprego com um rendimento favorável, ou até suficiente para garantir uma posição económica e socialmente estável, e, juntamente com as dificuldades de compreender informações de alerta e de recuperação, tornam-se mais vulneráveis a riscos e perigos (Cutter et al. 2003). Para além deste indicador, Chen et al. (2013), Gomes (2018), Guillard-Gonçalves et al. (2014) e Tasnuva et al. (2020) utilizam também, nos seus estudos, indicadores referentes à população com grau de ensino universitário o que, de acordo com o que foi dito anteriormente neste parágrafo, contribui para uma diminuição da vulnerabilidade social da população.

Os últimos indicadores desta secção são referentes aos apoios sociais disponibilizados às populações. Os estudos de Guillard-Gonçalves et al. (2014), Mendes (2009), e de Rabby et al. (2019) utilizam o indicador referente à taxa de famílias que recebem pensões de incapacidade ou deficiência uma vez que indivíduos com estas condições apresentam uma menor probabilidade e mais demorada recuperação após ocorrência de inconveniências, tornando-os, assim, mais vulneráveis (Gomes, 2018). Por outro lado, Chen et al. (2013), Cutter et al. (2003), Dintwa et al. (2019), e Schmidlein et al. (2008) utilizam um indicador mais genérico, referente à taxa de famílias que recebem apoios da segurança social. Estas famílias, estando dependentes dos serviços sociais, já se encontram económica e socialmente marginalizadas, necessitando assim de um apoio adicional no período pós-desastre, aumentando a vulnerabilidade da população em que estão inseridas (Cutter et al. 2003).

2.4.2. Infraestruturas

Nesta secção são descritos indicadores que refletem algumas características da sociedade em termos das suas infraestruturas. Estas infraestruturas são relevantes para a definição da vulnerabilidade de uma sociedade visto que podem afetar a forma como esta se prepara e recupera de riscos e perigos. Os indicadores aqui referidos dizem respeito ao tipo, à quantidade e à qualidade das infraestruturas existentes numa determinada área geográfica.

O primeiro indicador analisado nesta secção diz respeito à densidade de alojamentos familiares. Este indicador, utilizado nos estudos de Cutter et al. (2003) e de Rabby et al. (2019), representa o número de alojamentos familiares clássicos por definida unidade de medida, como o quilómetro quadrado ou a milha quadrada. Esta densidade estabelece uma relação negativa com a vulnerabilidade social da população, uma vez que áreas com uma

maior densidade, quer populacional, quer de infraestruturas, tendem a oferecer melhores redes de apoio (Gomes, 2018). Gomes (2018), Guillard-Gonçalves et al. (2014) e Rabby et al. (2019) aprofundam as suas abordagens a este indicador explorando as características e condições dos alojamentos em causa: saneamento, água potável, eletricidade, etc. A existência destes últimos indicadores, que acabam por representar a qualidade do alojamento em si, está também negativamente relacionada com a vulnerabilidade social, pelo que um alojamento que apresente estas condições e características, contribuirá para uma menor vulnerabilidade da população, em geral (Gomes, 2018).

Outros indicadores também utilizados são referentes às infraestruturas da área da saúde. Cutter et al. (2003), Guillard-Gonçalves et al. (2014), e Mendes (2009) usam, nos seus estudos, indicadores relativos ao número de hospitais e farmácias existentes em determinada área geográfica, indicadores que afetam positivamente a vulnerabilidade da sociedade. Isto deve-se ao facto de aumentarem a disponibilidade de cuidados e de prestadores de serviços de saúde (Gomes, 2018).

De forma a dar como concluída esta secção, apresenta-se, uma tabela-resumo, a **Tabela 4**, onde são referidos os indicadores mencionados anteriormente e os respetivos estudos onde foram utilizados como determinantes da vulnerabilidade social. Nesta tabela, são ainda apresentadas os tipos de relações estabelecidas entre o indicador em causa e a vulnerabilidade social, conforme foi descrito nos parágrafos anteriores.

Tabela 4 - Indicadores analisados.

Tema	Indicador	Relação	Estudos
População	Indivíduos com menos de 5 anos	+	Cutter et al. (2003), Guillard-Gonçalves et al. (2014), Rabby et al. (2019), Schmidtlein et al. (2008), e Tasnuva et al. (2020)
	Indivíduos com mais de 65 anos	+	Cutter et al. (2003), Guillard-Gonçalves et al. (2014), Rabby et al. (2019), Schmidtlein et al. (2008), e Tasnuva et al. (2020)
	População dependente	+	Chen et al. (2013), Cutter et al. (2003), Guillard-Gonçalves et al. (2014), Rabby et al. (2019), e Tasnuva et al. (2020)
	Taxa de natalidade	+	Cutter et al. (2003) e Mendes (2009)
	Taxa de desemprego	+	Cumberbatch et al. (2020), Cutter et al. (2003), Dintwa et al. (2019), Mendes (2009), Rabby et al. (2019), e Schmidtlein et al. (2008)
	Salário médio	-	Cumberbatch et al. (2020), Cutter et al. (2003), Dintwa et al. (2019), Mendes (2009), Rabby et al. (2019), e Schmidtlein et al. (2008)
	Taxa de analfabetismo	+	Guillard-Gonçalves et al. (2014), Rabby et al. (2019), e Tasnuva et al. (2020)
	População com grau de ensino universitário	-	Chen et al. (2013), Gomes (2018), Guillard-Gonçalves et al. (2014) e Tasnuva et al. (2020)
	Famílias que recebem pensões de incapacidade	+	Guillard-Gonçalves et al. (2014), Mendes (2009), e Rabby et al. (2019)
	Famílias que recebem apoios da Segurança Social	+	Chen et al. (2013), Cutter et al. (2003), Dintwa et al. (2019), e Schmidtlein et al. (2008)
Infraestruturas	Densidade de alojamentos familiares	-	Cutter et al. (2003), e Rabby et al. (2019)
	Qualidade dos alojamentos familiares	-	Gomes (2018), Guillard-Gonçalves et al. (2014), e Rabby et al. (2019)
	Nº de hospitais	-	Cutter et al. (2003), Guillard-Gonçalves et al. (2014), e Mendes (2009)
	Nº de farmácias	-	Cutter et al. (2003), Guillard-Gonçalves et al. (2014), e Mendes (2009)

Fonte: Elaboração própria.

3. Dados e Metodologia

Durante esta secção é descrito todo o procedimento metodológico a aplicar neste estudo. Como já foi referido anteriormente, a análise a ser efetuada consiste na tentativa de criação de um índice que quantifique e classifique a vulnerabilidade social a nível municipal de Portugal continental, num período de 11 anos, desde 2009 até 2019.

Este capítulo é, então, iniciado com uma análise descritiva dos dados a serem manipulados. Essa análise consiste na exposição dos passos de seleção, extração, e descrição desses mesmos dados. De seguida, é efetuado o pré-processamento dos dados, onde se efetua o tratamento e a transformação dos mesmos. O passo seguinte consiste na criação de um índice composto, o IVS, através do método de Análise de Componentes Principais, caso seja testada a viabilidade dos dados. Posteriormente, são criadas algumas categorias de classificação dos valores do IVS, com o intuito de serem realizados mapeamentos e representações geográficas dos resultados obtidos. Por fim, é ainda efetuada uma breve análise de robustez onde são comparados dois índices de vulnerabilidade social.

Todos estes passos estão descritos de forma mais detalhada nos subcapítulos seguintes.

3.1. Análise Descritiva dos Dados

Antes de efetuar qualquer processamento de dados, é importante realizar uma análise inicial dos dados em estudo. Esta análise deverá atuar sobre dados passados de variáveis ou indicadores, de forma a identificar como estes são classificados e, a partir dessa classificação, perceber que operações e processos podem ser aplicados. Para além disto, é esta etapa que permite formar as primeiras impressões sobre cada indicador, identificar as suas unidades de medida e o seu significado, verificando se estes, no seu estado bruto, são adequados, ou não, para o seu processamento (Rajagopalan, 2020).

Nesta etapa, são então apresentadas as características mais relevantes de cada indicador, assim como algumas das suas medidas de dispersão e de centralidade.

3.1.1. Seleção e Extração dos Dados

Para a criação de um índice de vulnerabilidade social, procurou-se, inicialmente, identificar indicadores que representassem, de alguma forma, o estado de vulnerabilidade da unidade geográfica em causa, tendo por base os estudos analisados durante a revisão de literatura efetuada anteriormente.

Para o efeito, foram exploradas bases de dados que fornecessem informações a este nível, das quais se destacou a PORDATA. Esta base de dados recolhe, organiza, sistematiza e divulga informação oficial e certificada sobre diversas áreas da sociedade, quer de Portugal e dos seus municípios, quer de países europeus. Assim, tendo sido selecionada uma fonte de dados de onde é possível retirar informação para a criação de um índice de vulnerabilidade social, pode-se proceder à etapa de recolha dessa informação.

Explorando a informação fornecida pela PORDATA, ao nível dos municípios de Portugal, nota-se que esta se encontra dividida em múltiplas categorias. Como se pretende criar um índice que meça a vulnerabilidade social em diversos aspetos da sociedade, optou-se por selecionar indicadores de algumas dessas categorias disponíveis, nomeadamente indicadores que se incluam nas duas grandes categorias mencionadas no capítulo Revisão de Literatura: População e Infraestruturas.

A seleção dos dados procurou ir de encontro aos indicadores selecionados por outros autores, referidos durante a revisão de literatura, quando realizaram um estudo semelhante ao desta dissertação, tendo sido também influenciada pela disponibilidade e divulgação dos dados para os anos e divisão geográfica selecionada, que nem sempre cumpriu as expectativas criadas, tendo então sido encontrada a primeira limitação deste estudo. De forma a serem utilizados indicadores mais próximos ou semelhantes aos analisados durante a revisão de literatura, terão de ser recolhidos dados que não são disponibilizados com a mesma periodicidade temporal que a selecionada, mais concretamente, alguns indicadores apenas estão disponíveis decenalmente, nos anos em que foram realizados os censos, em vez de anualmente.

De forma a dar continuidade a este estudo, optou-se por utilizar alguns desses indicadores decenais, usando o valor do ano pertencente ao período a analisar, 2011, para todos os anos desse período, sendo, assim possível ter em conta as variáveis representadas por esses indicadores.

Foi, então, selecionado um conjunto de indicadores, para os anos compreendidos entre 2009 e 2019, para os 278 municípios de Portugal Continental, apresentado na **Tabela 5**,

juntamente com a nomenclatura que será utilizada para identificar cada indicador no decorrer deste estudo. Este conjunto de indicadores compreende, não só, indicadores a serem utilizados na análise a ser efetuada, como também indicadores auxiliares a algumas transformações que serão necessárias e efetuadas num subcapítulo posterior.

Tabela 5 - Indicadores extraídos da PORDATA.

Nome	Indicador	Frequência
MENOS5	Número de indivíduos com menos de 5 anos	Anual
MAIS65	Número de indivíduos com mais de 65 anos	Anual
DEP	Percentagem de população dependente	Anual
NAT	Taxa de natalidade	Anual
ILIT	Taxa de analfabetismo	Decenal
UNI	Número de indivíduos com nível de ensino universitário	Decenal
DES	Número de desempregados inscritos nos centros de desemprego	Anual
ATIVA	População residente em idade ativa	Anual
DIF	Diferença entre o salário mínimo nacional e a remuneração média	Anual
INCAP	Número de famílias que recebem pensões de incapacidade	Anual
SS	Número de famílias que recebem apoios da Segurança Social	Anual
DENS	Densidade de alojamentos familiares por km ²	Anual
ANC	Número de alojamentos familiares não clássicos	Decenal
HOSP	Número de hospitais	Anual
FARM	Número de farmácias	Anual
EPE	Número de estabelecimentos de Ensino Pré-Escolar	Anual
EB	Número de estabelecimentos de Ensino Básico	Anual
ES	Número de estabelecimentos de Ensino Secundário	Anual
RESID	População média anual	Anual

Fonte: Elaboração própria.

Nota: Os indicadores foram retirados da base de dados PORDATA (2022).

São excluídos, desde já, os dados referentes às Regiões Autónomas dos Açores e da Madeira uma vez que estas possuem características específicas a elas mesmas e que se refletem num estatuto político-administrativo e em órgãos de governo próprios, tornando a sua incorporação nesta análise bastante mais complexa (Sigorro, 2018).

Após selecionados os indicadores a usar neste estudo, assim como os indicadores auxiliares, procedeu-se à extração dos dados necessários da PORDATA e, posteriormente,

à sua importação para um ficheiro *Notebook* do *Python*, através do software *Visual Studio Code*.

Já no *Notebook*, os dados foram organizados em *dataframes*, para os 11 anos a serem analisados, onde, neste caso, cada linha dessa *dataframe* corresponde a um município, e cada coluna corresponde a um indicador. Desta forma, é obtida uma tabela com 278 linhas e 20 colunas, sendo que a primeira coluna corresponde ao nome do município.

Tendo sido os dados já devidamente preparados para ser em analisados e explorados, procede-se à etapa seguinte, a descrição dos dados.

3.1.2. Descrição e Análise dos Dados

Aqui, são descritas as principais informações transmitidas pelos dados extraídos, como o tipo de variável e a sua unidade de medida, registadas na **Tabela 6**.

Tabela 6 - Descrição dos Dados.

Nome	Tipo de Variável	Unidade de Medida	Tipo de Valor
MENOS5	Quantitativa Discreta	Indivíduos	Valor absoluto
MAIS65	Quantitativa Discreta	Indivíduos	Valor absoluto
DEP	Quantitativa Contínua	Percentagem	Rácio
NAT	Quantitativa Contínua	Percentagem	Rácio
ILIT	Quantitativa Contínua	Percentagem	Rácio
UNI	Quantitativa Discreta	Indivíduos	Valor absoluto
DES	Quantitativa Discreta	Indivíduos	Valor absoluto
ATIVA	Quantitativa Discreta	Indivíduos	Valor absoluto
DIF	Quantitativa Contínua	Euros	Média
INCAP	Quantitativa Discreta	Indivíduos	Valor absoluto
SS	Quantitativa Discreta	Indivíduos	Valor absoluto
DENS	Quantitativa Contínua	Alojamentos	Média
ANC	Quantitativa Discreta	Alojamentos	Valor absoluto
HOSP	Quantitativa Discreta	Hospitais	Valor absoluto
FARM	Quantitativa Discreta	Farmácias	Valor absoluto
EPE	Quantitativa Discreta	Estabelecimentos	Valor absoluto
EB	Quantitativa Discreta	Estabelecimentos	Valor absoluto
ES	Quantitativa Discreta	Estabelecimentos	Valor absoluto
RESID	Quantitativa Discreta	Indivíduos	Valor absoluto

Fonte: Elaboração Própria.

Como se pode perceber a partir da descrição dos indicadores, a sua maioria não reflete uma relação direta com a população ou dimensão do município em causa, não sendo possível efetuar comparações corretas entre os municípios sem antes proceder à transformação desses indicadores. Assim, e dando este processo como concluído, procede-se à etapa seguinte, o pré-processamentos dos dados.

3.2. Pré-Processamento dos Dados

O pré-processamentos de dados corresponde a uma etapa realizada em qualquer processo de análise e manipulação de dados onde são executadas diversas atividades de preparação, organização e estruturação dos mesmos. É, por norma, das primeiras etapas a ser executada, precedendo a realização de qualquer modelação ou previsão de dados (Zhao, 2021). De forma geral, esta etapa consiste então num conjunto de diversas técnicas e práticas utilizadas para transformar dados brutos em dados cujos formatos são, para além de mais eficientes, mais úteis à análise a ser realizada, e deverá, assim, ser ajustada aos objetivos finais da mesma, escolhendo-se efetuar apenas os passos necessários para atingir esses objetivos (García et al., 2016).

Deste modo, o pré-processamento de dados consiste num passo fundamental e de elevada importância, visto que determina a qualidade final dos dados a serem analisados e evita que ocorra o fenómeno de *“garbage in, garbage out”* no processo de análise de dados. Isto é, apesar de poderem entrar dados sem sentido e com pouca qualidade, no estudo a ser efetuado, o pré-processamento evita que resultem, deste estudo, dados não tratados, sem utilidade ou significado relevante. Por conseguinte, este processo é realizado tendo em vista o melhor desempenho e a obtenção dos melhores resultados possíveis (Silva, 2021).

As técnicas de pré-processamento de dados mais comuns passam pela transformação, integração, limpeza e redução dos dados (Zhao, 2021). No entanto, realizando um ajuste ao estudo atual, este processo foca-se apenas em dois passos principais, a limpeza de dados e a transformação de dados que serão descritos mais detalhadamente nas secções seguintes.

3.2.1. Limpeza dos Dados

A limpeza de dados foca-se no preenchimento de informação em falta, na redução do seu ruído, na identificação de *outliers*, e na correção de outras inconsistências presentes no conjunto de dados a serem processados (Sivakumar & Gunasundari, 2017). No entanto, neste estudo, apenas será abordado o preenchimento dos dados em falta uma vez que é o único procedimento que realmente se torna necessário perante o conjunto de dados selecionados. Os dados em falta, ou seja, as observações cujo valor que lhes foi atribuído é não existente, podem ser eliminados ou podem ser preenchidos com um valor, conforme o que for mais correto para a análise efetuada. Normalmente, este valor de preenchimento pode corresponder à média, mediana ou moda dos valores dessa variável, ou ainda a um valor mais provável obtido por uma técnica preditiva, como uma regressão (García et al., 2016).

Uma vez que se considera mais relevante manter todos os distritos para análise, procede-se com o preenchimento dos valores em falta por um outro valor. Observando a restante literatura, este preenchimento é normalmente efetuado pela substituição dos dados em falta pela média geral do indicador em causa, como nos artigos de Rabby et al. (2019) e de Tarling (2017). No entanto, de forma a manter a distribuição dos dados o mais semelhante com o original quanto possível, optou-se pela substituição dos dados em falta pela média distrital, à semelhança do que é feito nos artigos de Chen et al (2013) e de Karaye e Horney (2020).

Após todos os dados em falta terem sido devidamente preenchidos, procede-se à Transformação dos Dados.

3.2.2. Transformação dos Dados

Durante a etapa de transformação de dados são executados processos que modificam os dados originais em dados com formatos mais adequados e apropriados para a análise a ser realizada. Um dos processos mais comuns desta etapa é a normalização, um processo que passa por redimensionar os valores de dados numéricos entre um intervalo específico, de forma que todas as variáveis se encontrem na mesma escala e, desta forma, seja possível estabelecer comparações e relações entre as mesmas (Sivakumar & Gunasundari, 2017).

São também aplicados outros processos durante a transformação de dados, como, por exemplo, a agregação, a generalização, a criação e a manipulação de dados. A criação de novos dados consiste na adição de dados ao estudo, quer estes sejam provenientes de uma

fonte de dados, diferente da original, ou criados através de dados já existentes (Han et al., 2012). Por outro lado, a manipulação dos dados consiste em modificar ou alterar os dados existentes com o objetivo de os tornar mais corretos e organizados (Pratt & Bernstein, 2022).

Neste estudo, a etapa de transformação de dados inicia-se com a criação de novos indicadores, e, posteriormente, é realizada a manipulação de alguns dados medidos em valores absolutos em dados per capita. Por fim, procede-se à normalização dos dados relativos aos indicadores transformados.

Realiza-se, então, em primeiro lugar, a criação de novos indicadores através de relações entre os indicadores já existentes, tornando assim o novo conjunto de dados mais consistente e mais direcionado para o objetivo final deste estudo, indo de encontro aos indicadores referidos durante o capítulo Revisão de Literatura. Os indicadores a serem criados são referentes ao salário médio, à taxa de população desempregada de cada município, e ao número de estabelecimentos de ensino.

O indicador referente ao salário médio, SMED, será obtido através da subtração do indicador DIF ao Salário Mínimo Nacional verificado em cada ano, como se pode observar na **Equação 1**. O indicador TDES, referentes à taxa de população desempregada, corresponde à proporção de indivíduos desempregados, em relação ao número de indivíduos em idade ativa, como observado na **Equação 2**. Quanto ao indicador EE, referente ao número total de estabelecimentos de ensino existentes num município, este é obtido através da soma do número de estabelecimentos de ensino pré-escolar, básico e secundário num único indicador, como descrito na **Equação 3**.

Salário Médio

$$SMED = SMIN - DIF \quad (1)$$

Onde:

SMED = Salário Médio

SMIN = Salário Mínimo

DIF = Diferença entre o Salário Mínimo e a Remuneração Média Mensal

Taxa de População Desempregada

$$TDES = \frac{DES}{ATIVA} * 100 \quad (2)$$

Onde:

TDES = Taxa de população desempregada

DES = Número de desempregados inscritos nos centros de desemprego

ATIVA = Número de indivíduos com idade entre os 15 e os 64 anos

Estabelecimentos de Ensino

$$EE = EPE + EB + ES \quad (3)$$

Onde:

EE = Estabelecimentos de Ensino

EPE = Estabelecimentos de ensino pré-escolar

EB = Estabelecimentos de ensino básico

ES = Estabelecimentos de ensino secundário

Depois da criação destes novos indicadores, são removidos os indicadores auxiliares a este processo uma vez que estão agora incorporados nos novos dados. Assim, foram removidos os indicadores referentes à diferença entre o Salário Mínimo e a Remuneração Média Mensal (DIF), ao número de desempregados inscritos nos centros de desemprego (DES), à população em idade ativa (ATIVA), e os indicadores referentes ao número de estabelecimentos dos diversos níveis de ensino (EPE, EB e ES).

Posteriormente, procedeu-se à manipulação de alguns indicadores representados por valores absolutos, transformando-os em indicadores per capita que possam, de alguma forma, traduzir a dimensionalidade populacional de cada município. Esta transformação, como referido anteriormente, tem como objetivo permitir a comparação entre municípios com elevadas disparidades nas suas dimensões. Por exemplo, é bastante justificável que municípios como o município de Lisboa ou do Porto, onde se concentra grande parte do

dinamismo de Portugal, quer em termos turísticos, populacionais ou comerciais, apresentem um número relativamente superior de estabelecimentos de ensino que municípios mais isolados como Penedono ou Alcoutim, visto que os primeiros têm uma maior dimensão populacional e, desta forma, têm de responder de acordo com as suas necessidades. Por conseguinte, torna-se necessário relacionar indicadores deste género com a dimensão populacional de cada município, o que pode ser alcançado dividindo os valores originais desses mesmos indicadores pelo número médio de indivíduos residentes no município, ou seja, pelo indicador *RESID*.

Os indicadores a sofrerem esta manipulação são os que ainda não consistiam numa taxa ou percentagem que representasse a dimensão populacional que se tem vindo a descrever, ou seja, os indicadores *MENOS5*, *MAIS65*, *UNI*, *SS*, *INCAP*, *ANC*, *HOSP*, *FARM*, e *EE*.

Para estes novos valores serem dotados de uma leitura mais direta, opta-se ainda pela sua multiplicação por mil de forma que o valor resultante desta manipulação devolva indicadores medidos numa unidade por mil habitantes, ou seja, o indicador *HOSP*, por exemplo, irá representar o número de hospitais, existentes no município, por mil habitantes. Esta relação é traduzida, em termos práticos, através da **Equação 4**.

$$Y = \frac{X}{RESID} * 1000 \quad (4)$$

Onde:

Y = Valores do indicador após a transformação

X = Valores originais do indicador

RESID = População residente média anual do município

Esta etapa é finalizada com a normalização dos indicadores. Como já foi mencionado, a normalização tem o intuito de tornar todos os indicadores comparáveis entre si, igualando a escala por onde estes se distribuem. A forma como esta normalização é realizada varia entre a literatura analisada. Uns autores optam por padronizar os valores, ou seja, subtrair o valor médio desse indicador e dividir o respetivo resultado pelo desvio padrão, obtendo uma variável com média nula e desvio padrão igual à unidade, como é indicado no estudo realizado pela University of South Carolina (2016). Outros autores optam por identificar, em

primeiro lugar, o tipo de relação estabelecida entre o indicador em causa e a vulnerabilidade social para, de seguida, aplicar o método de Normalização Min-Max, que traduz, desde início, essa relação, como se verifica nos artigos de Žurovec et al. (2017) e de Ramadas et al. (2018).

Neste estudo, optou-se por realizar a normalização dos indicadores de acordo com este último método, garantindo que os valores normalizados apresentem, no final, uma correlação positiva com a vulnerabilidade. Assim, a normalização de um indicador que estabeleça uma relação positiva com a vulnerabilidade social, isto é, que perante um aumento do seu valor, aumente a vulnerabilidade social, é realizada de acordo com a **Equação 5**, mas, caso o aumento do valor do indicador em causa se reflita numa diminuição da vulnerabilidade social, sendo estabelecida uma relação negativa, esta normalização já será feita de acordo com a **Equação 6**.

$$Y_{ij} = \frac{X_{ij} - \text{Min}(X_j)}{\text{Max}(X_j) - \text{Min}(X_j)} \quad (5)$$

Onde:

Y_{ij} = Valor normalizado do indicador j, referente ao município i

X_{ij} = Valor observado do indicador j, referente ao município i

$\text{Max}(X_j)$ = Valor máximo do indicador j, entre todos os municípios

$\text{Min}(X_j)$ = Valor mínimo do indicador j, entre todos os municípios

$$Y_{ij} = \frac{\text{Max}(X_j) - X_{ij}}{\text{Max}(X_j) - \text{Min}(X_j)} \quad (6)$$

Estando definidas ambas as fórmulas de normalização, importa agora definir quais os indicadores que apresentam uma relação negativa ou positiva com a vulnerabilidade social de um município. Desta forma, e de acordo com o que foi verificado durante a revisão de literatura, foi construída a **Tabela 7**, onde estão registados os tipos de relação que cada indicador estabelece com a vulnerabilidade, assim como um conjunto de autores que justificam essa mesma relação nos seus estudos.

Antes de dar por concluída a etapa de transformação dos dados, pode, ainda, fazer-se notar que esta etapa teve um forte foco na comparabilidade dos dados, começando pela manipulação dos indicadores representados por dados absolutos, para que estes se traduzissem em dados per capita, de forma que fosse possível estabelecer uma comparação entre municípios, e terminando com a normalização dos indicadores, tendo em vista a comparação entre variáveis, através da equiparação das suas escalas.

Tabela 7 - Relação estabelecida entre os indicadores e a vulnerabilidade social.

Indicador	Relação	Autores
MENOS5	+	Guillard-Gonçalves et al. (2014), Rabby et al. (2019), e Tasnuva et al. (2020)
MAIS65	+	Guillard-Gonçalves et al. (2014), Rabby et al. (2019), e Tasnuva et al. (2020)
DEP	+	Guillard-Gonçalves et al. (2014), Rabby et al. (2019), e Tasnuva et al. (2020)
NAT	+	Cutter et al. (2003), e Mendes (2009)
ILIT	+	Gomes (2018), Guillard-Gonçalves et al. (2014), e Tasnuva et al. (2020)
UNI	-	Gomes (2018), Guillard-Gonçalves et al. (2014), e Tasnuva et al. (2020)
TDES	+	Cumberbatch et al. (2020), Dintwa et al. (2019), e Rabby et al. (2019)
SMED	-	Cumberbatch et al. (2020), Dintwa et al. (2019), e Rabby et al. (2019)
SS	+	Chen et al. (2013), Dintwa et al. (2019), e Schmidlein et al. (2008)
INCAP	+	Guillard-Gonçalves et al. (2014), Mendes (2009), e Rabby et al. (2019)
DENS	-	Cutter et al. (2003) e de Rabby et al. (2019)
ANC	+	Gomes (2018), Guillard-Gonçalves et al. (2014), e Rabby et al. (2019)
HOSP	-	Cutter et al. (2003), Guillard-Gonçalves et al. (2014), e Mendes (2009)
FARM	-	Cutter et al. (2003), Guillard-Gonçalves et al. (2014), e Mendes (2009)
EE	-	Mendes (2009)

Fonte: Elaboração Própria.

Assim, estando todos os indicadores transformados e normalizados, dá-se por concluída a etapa de transformação dos dados, assim como o seu pré-processamento, podendo então

proceder-se à próxima etapa para a criação de um Índice de Vulnerabilidade Social a partir destes novos dados.

3.3. Índice Compósito

Para a criação de um Índice de Vulnerabilidade Social, como foi verificado anteriormente, a literatura é bastante consistente no que refere ao procedimento a efetuar, notando-se uma forte preferência pela Análise de Componentes Principais, ou ACP, como procedimento auxiliar à criação de um índice compósito.

A Análise de Componentes Principais é comumente utilizada como um método de redução de dados, uma vez que decompõe um conjunto de variáveis originais num conjunto de menor dimensão de outras variáveis (Solangaarachchi et al., 2012). Estas novas variáveis são denominadas de Componentes Principais, ou CPs, e estão relacionadas com as variáveis originais através de combinações lineares, agregando a informação transmitida pelas mesmas (Avila-Vera et al., 2012).

Inicialmente, é realizada uma análise da viabilidade dos dados, de forma a perceber se os dados em análise são, ou não, adequados para este tipo de procedimento, e, posteriormente, será determinado o número de componentes principais a serem usadas para a construção do IVS.

3.3.1. Viabilidade dos Dados

Antes de se proceder à criação do Índice de Vulnerabilidade Social, importa testar a viabilidade dos dados de forma a verificar se estes são, de facto, adequados à aplicação do processo de Análise de Componentes Principais. A existência desta viabilidade é verificada através de dois métodos: o Teste da Esfericidade de Bartlett e o Teste de Kaiser-Meyer-Olkin, ou Teste de KMO.

O Teste da Esfericidade de Bartlett testa a hipótese de a matriz de correlações ser igual à Matriz Identidade, isto é, a uma matriz onde os elementos da sua diagonal principal são iguais à unidade e qualquer elemento fora desta é nulo (Regis, 2012). Caso esta hipótese seja verificada, pode-se concluir que nenhum indicador está minimamente correlacionado com outro indicador, o que se traduz num conjunto de dados não adequado à aplicação do método ACP. A execução deste teste devolve, entre outros valores, um *p-value* que, caso

seja inferior a 0,05, leva à rejeição da hipótese mencionada, podendo assim concluir-se que o conjunto de dados é adequado (Aprilliant, 2021).

Por fim, o Teste de KMO avalia o quão adequada será a Análise de Componentes Principais, tendo em conta os dados da amostra. Esta adequação é medida pela comparação entre as correlações entre duas variáveis e as suas correlações parciais, sendo que estas correlações parciais consistem na medida da relação entre essas duas variáveis, sendo removida a influência que as outras variáveis da amostra têm nessa mesma relação (Rakotomalala, 2013). Segundo Regis (2012), quanto mais próximo da unidade for o valor devolvido pelo Teste de KMO, mais adequada será a amostra à aplicação do método ACP, no entanto, Aprilliant (2021) ainda acrescenta que o valor do teste deverá ser superior a 0,5, e, caso o seja, conclui-se a favor da adequação dos dados na realização da análise em causa.

3.3.2. Determinação do Número de Componentes

Como foi referido anteriormente, a aplicação do método ACP resulta num determinado número de componentes que tem de ser predefinido e, para esta predefinição, existem diversos critérios usados na literatura.

O critério mais comumente utilizado é o Critério de Kaiser, seguido pelo Teste de Cattell e pelo Critério de Benzécri (Norlén, 2018).

O Critério de Kaiser consiste na seleção das componentes que devolvem um autovalor superior à unidade. Por autovalor, ou valor próprio, entende-se, de uma forma geral, a variância capturada por essa componente, que corresponde aos elementos da diagonal da matriz de variâncias e covariâncias das componentes principais (Holland, 2019). Desta forma, este critério, ao descartar as CPs com um autovalor inferior à unidade, garante que cada componente selecionada explica mais variância que uma única variável (Sheytanova, 2015).

Quanto ao Teste de Cattell, este procura justificar a escolha do número de componentes principais a selecionar através de uma representação gráfica, o *Scree Plot*. Esta representação gráfica mostra, no eixo horizontal, o número correspondente à componente principal em causa e, no eixo vertical, o seu autovalor. Devido à ordenação das componentes pelo seu autovalor, do maior para o menor, o gráfico resultante assemelha-se a um penhasco, ou “*scree*”, em inglês. Assim, é obtido um gráfico semelhante ao representado na

Figura 2, onde se seleciona o número de componentes correspondente ao ponto onde a curva deste gráfico altera a sua inclinação de forma notável, também denominado de “cotovelo” (Sheytanova, 2015).

Por fim, existe ainda o Critério de Benzécri onde são selecionadas as componentes principais que expliquem uma determinada percentagem da variância dos dados. Este critério é também conhecido como a Abordagem da Variância Explicada Acumulada, ou VEA, uma vez que permite atingir um limiar predefinido de variância explicada pelo conjunto das CPs selecionadas, garantindo que a redução de dimensionalidade não resulta em significativas perdas de informação (Brems, 2017). Por norma, importa que o valor predefinido de VEA seja, pelo menos, superior a 70%, pois, desta forma, é possível garantir que é explicada grande parte da variabilidade dos dados originais, depois da sua redução de dimensionalidade (Rozalia Gabor et al., 2010).

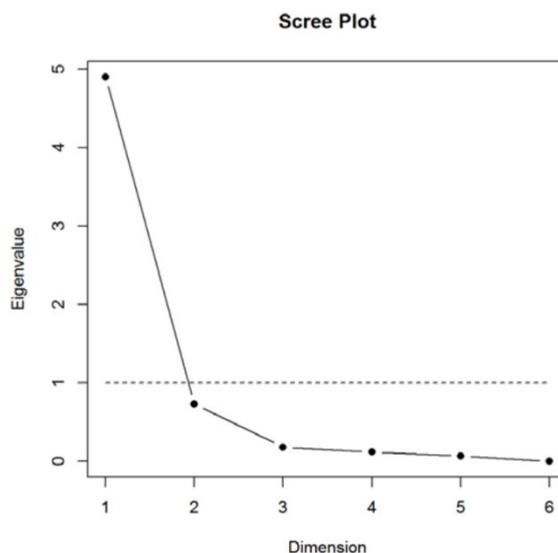


Figura 2 - Exemplo de um Scree Plot.

Fonte: Norlén (2018).

Estando selecionado o número de componentes principais, e verificada a sua apropriação, pode-se proceder à aplicação do método ACP.

3.3.3. Análise de Componentes Principais

Apesar de ter vindo a ser referido e descrito, de uma forma geral, ao longo deste estudo, importa agora realizar uma descrição mais detalhada do método da Análise de Componentes Principais. Este método, como referido anteriormente, consiste numa técnica estatística

multivariada, frequentemente usada como um método de redução de dimensionalidade, uma vez que decompõe um conjunto original de variáveis, num conjunto de menor dimensão de variáveis lineares não correlacionadas entre si, denominadas de Componentes Principais (Solangaarachchi et al., 2012).

Dados k indicadores referentes a n unidades geográficas, pode ser construída a matriz representada na **Equação 7**, onde $I_{n,k}$ diz respeito ao valor do indicador k , referente ao município n .

$$I = \begin{bmatrix} I_{1,1} & I_{1,2} & \dots & I_{1,k} \\ I_{2,1} & I_{2,2} & \dots & I_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ I_{n,1} & I_{n,2} & \dots & I_{n,k} \end{bmatrix} \quad (7)$$

Supondo que os valores destes indicadores foram devidamente tratados e cumprem os requisitos necessários para a aplicação da Análise de Componentes Principais, e aplicando então este método, é criado o conjunto de igualdades representado pela **Equação 8**, onde cada igualdade é referente a uma Componente Principal, para cada unidade geográfica:

$$\begin{aligned} CP_1 &= \omega_{1,1}I_1 + \omega_{1,2}I_2 + \dots + \omega_{1,k}I_k \\ CP_2 &= \omega_{2,1}I_1 + \omega_{2,2}I_2 + \dots + \omega_{2,k}I_k \\ &\vdots \\ CP_k &= \omega_{k,1}I_1 + \omega_{k,2}I_2 + \dots + \omega_{k,k}I_k \end{aligned} \quad (8)$$

Matematicamente, cada componente principal é composta por um conjunto de combinações lineares que são quantificadas pelos *loadings* representados por ω , indicando o tipo e força da relação estabelecida entre a componente principal e o indicador (Sheytanova, 2015). Estes *loadings* traduzir-se-ão no autovalor da componente principal em causa que, por sua vez, será contabilizado para a porção de variância explicada por essa mesma componente (Norlén, 2018). Desta forma, a ACP procura explicar o máximo de variabilidade observada das variáveis originais, através da variância explicada do menor número possível de CPs (Avila-Vera et al., 2012).

Procedendo à aplicação deste método, após ser avaliada a adequação dos dados, e ser definido um número de componentes principais a serem calculadas, é utilizada a função “PCA” pertencente ao módulo *Decomposition*, ou “Decomposição”, em português, da

biblioteca *Scikit-learn*. Esta função permite acesso aos *loadings* das componentes principais, assim como à variância explicada por elas. É através destes dois elementos que irá ser construído o Índice de Vulnerabilidade Social.

3.3.4. Criação do Índice

Após terem sido obtidas as componentes principais, através do método de ACP, é necessário proceder à sua agregação de forma a obter um índice compósito.

Um índice compósito procura agregar um determinado número de diversos indicadores numa única componente, permitindo uma leitura mais rápida e simples das informações transmitidas por esse conjunto inicial de indicadores (Instituto Nacional de Estatística, 2013). Este tipo de índice revela inúmeras vantagens, como, por exemplo, a sumarização de realidades complexas e multidimensionais, a sua aplicabilidade a diferentes unidades geográficas, permitindo uma fácil comparação entre nações e países de diferentes dimensões, e ainda, a facilitação da transmissão de informações complexas ao público em geral (Nardo et al., 2008). No entanto, existem também algumas desvantagens, como o facto dos resultados “generalistas” provenientes destes poderem influenciar políticas e ações públicas, originando conclusões simplistas e inapropriadas à realidade, tal como referido durante a revisão de literatura. Desta forma, é importante que os índices compósitos sejam vistos como uma abertura à discussão e estimulação do interesse público (Nardo et al., 2008).

Na literatura, existem diversos métodos de criação de um índice compósito através dos resultados obtidos pela ACP, porém, facilmente se percebe que, apesar desta diversidade, não é atingido nenhum consenso sobre que forma de agregação de indicadores é a mais correta entre elas.

Múltiplos autores optam por realizar uma agregação com base em pesos igualitários onde é atribuída, a todas as variáveis em causa, igual importância para a construção de um índice. Este método, apesar de implicar que todas as variáveis têm o mesmo valor para o índice compósito, poderá disfarçar a insuficiência de conhecimento relativo às relações entre essas variáveis ou ainda, a falta de um consenso na forma de agregação destas, podendo-se traduzir num índice com uma estrutura desequilibrada (Nardo et al., 2008).

No estudo de Dintwa et al. (2019), é aplicada esta agregação, também baseada na suposição de que cada componente contribui de igual forma para a vulnerabilidade, devido

à não existência de um método consistente e confiável de atribuição de pesos a estas componentes. Estes autores justificam, ainda, esta decisão com o facto de este ter sido exatamente o procedimento realizado no trabalho de Cutter et al. (2003), no qual basearam a sua metodologia. Esta mesma justificação é também utilizada para a aplicação do método de agregação baseado em pesos iguais nos estudos de Mavhura et al. (2017) e de Schmidtlein et al. (2008).

Outra forma de agregação de indicadores num índice compósito passa pela atribuição de pesos aos indicadores baseados na percentagem de variância explicada por estes. Esta metodologia, explicada no livro de Nardo et al. (2008), é aplicada nos estudos de Sabão Popa et al. (2021), Solangaarachchi et al. (2012) e Wood et al., (2010), e segue a fórmula indicada na **Equação 9** e permite, então, que componentes com variância explicada superior tenham uma maior contribuição para o valor geral do índice a calcular. Ainda através desta equação, percebe-se que o peso atribuído a cada componente principal corresponde à divisão do valor da variância explicada por essa mesma componente pelo valor da variância explicada acumulada pela totalidade de CPs seleccionadas para a análise em causa, como se pode observar na **Equação 10**.

$$I = \frac{VE_1}{VEA} CP_1 + \frac{VE_2}{VEA} CP_2 + \dots + \frac{VE_k}{VEA} CP_k \quad (9)$$

Onde:

I = Valor do índice calculado, para determinado ano

VE_i = Variância Explicada pela Componente Principal i

CP_i = Valor da Componente Principal i

VEA = Variância Explicada Acumulada, para o ano em causa

$$\rho_i = \frac{VE_i}{VEA} \quad (10)$$

Onde:

ρ = Peso atribuído à Componente Principal i, em determinado ano

VE_i = Variância Explicada pela Componente Principal i

VEA = Variância Explicada Acumulada, para o ano em causa

Após multiplicar o valor de cada componente pelo respetivo peso, os valores resultantes são adicionados, sendo então obtido um índice compósito inicialmente pretendido.

3.3.5. Normalização do Índice

Tendo já os dados referentes ao índice compósito, torna-se necessário proceder à sua normalização de forma que seja possível estabelecer comparação dos valores desse índice em diferentes anos e, ainda, de forma a facilitar a sua interpretação, tornar os limites do índice em causa mais claros.

Nos estudos de Aung et al. (2021) e de Dintwa et al. (2019), após ter sido obtido o índice compósito, é realizada uma normalização dos seus valores através do método já mencionado anteriormente, a Normalização Min-Max, de forma que estes variem entre 0 e 1. Já no estudo de Schmidtlein et al. (2008), os valores do índice são padronizados, sendo obtida uma média nula e desvio-padrão igual à unidade, sem qualquer referência a valores fixos dos limites desse índice.

Neste estudo, como já foi referido, considera-se de elevada importância serem estabelecidos limites mínimos e máximos concretos para este índice, uma vez que facilita a sua interpretação. Desta forma, procede-se do mesmo modo que Aung et al. (2021) e Dintwa et al. (2019), onde os valores obtidos para o IVS são normalizados segundo a **Equação 11**.

$$Y_i = \frac{X_i - \text{Min}(X)}{\text{Max}(X) - \text{Min}(X)} \quad (11)$$

Onde:

Y_i = Valor normalizado do índice, referente ao ano i

X_i = Valor observado do índice, referente ao ano i

$\text{Max}(X)$ = Valor máximo do índice, entre todos os anos

$\text{Min}(X)$ = Valor mínimo do índice, entre todos os anos

Tendo sido realizada esta normalização, o índice calculado encontra-se devidamente tratado, podendo-se proceder à sua análise.

3.4. Representação Geográfica

Como último passo deste processo, pretende-se representar o índice calculado ao longo da zona geográfica definida inicialmente.

As representações geográficas e cartogramas consistem nas formas mais simples e intuitivas de observar a distribuição de indicadores e índices ao longo de determinado espaço geográfico.

Desta forma, considerou-se oportuno proceder à representação gráfica da distribuição do índice calculado neste estudo. No entanto, antes de ser efetuada essa representação, realizou-se uma classificação dos valores do índice que, conciliados com o seu mapeamento, permitirão uma interpretação mais intuitiva e uma recolha de informação e tomada de decisões mais facilitadas.

3.4.1. Classificação

Após os valores do Índice de Vulnerabilidade Social terem sido calculados e normalizados, importa proceder à criação de classes de acordo com os valores obtidos, permitindo uma melhor interpretação da distribuição da vulnerabilidade no espaço geográfico analisado.

Assim, e de forma análoga ao processo realizado nos estudos de Cumberbatch et al. (2020) e de Žurovec et al. (2017), são criadas cinco classes às quais são atribuídas um indicador qualitativo referente ao nível de vulnerabilidade verificado na unidade geográfica em causa, segundo o seguinte modo:

- Muito baixa (entre 0,0 e 0,2);
- Baixa (entre 0,2 e 0,4);
- Moderada (entre 0,4 e 0,6);
- Elevada (entre 0,6 e 0,8);
- Muito elevada (entre 0,8 e 1,0).

3.4.2. Cartogramas

De forma a conseguir representar o IVS num cartograma em Python, recorreu-se à biblioteca *GeoPandas*. O uso desta biblioteca implica que seja utilizado um ficheiro do tipo *shapefile* onde esteja armazenada toda a informação geográfica a ser utilizada no estudo em causa.

Neste estudo, foi utilizado o *shapefile* da Direção-Geral do Território (2022), denominado de Carta Administrativa Oficial de Portugal que transmite informação relativa a áreas e perímetros das entidades administrativas de Portugal, quer continental quer das Regiões Autónomas, e ainda informações relativas às freguesias e aos municípios das mesmas zonas. Deste *shapefile*, foram extraídas apenas as informações referentes aos limites de cada município de Portugal Continental, as quais foram conciliadas com os valores do índice calculado e as classificações desses municípios.

Tendo esta informação devidamente organizada, é então possível proceder à criação e manipulação de cartogramas, de forma que sejam visualizadas as componentes pretendidas. Observando a literatura já existente, verifica-se que os dados do índice calculado são mapeados para a área geográfica e período temporal selecionados no início do presente trabalho. No entanto, neste estudo, para além da área geográfica ser bastante extensa e detalhada, foram analisados múltiplos anos e, assim, justifica-se a criação de diversos cartogramas de forma a permitir uma mais completa análise.

Desta forma, são mapeados os valores do IVS e os níveis de vulnerabilidade, quer em vista nacional quer em vista distrital, para determinado ano, sendo assim possível observar

a distribuição da vulnerabilidade ao longo de Portugal e ao longo do período temporal analisado.

3.5. Análise de Robustez

Um passo adicional que se pretende incluir ainda neste capítulo consiste numa análise de robustez ao processo metodológico efetuado, mais especificamente, aos indicadores selecionados para a criação do IVS.

Importa, primeiramente, efetuar uma breve exploração do conceito de robustez. Vincke (1999) classifica a robustez como sendo uma propriedade de um modelo ou sistema onde, quando verificada, este não apresenta resultados contraditórios caso sejam utilizadas diferentes fontes de informação ou enunciações. Desta forma, pode-se afirmar que uma análise de robustez procura perceber a adequação dos dados e metodologia utilizadas, estando assim relacionada com os conceitos sensibilidade e de estabilidade.

Nesta Análise de Robustez, procura-se comparar os valores obtidos, para o ano de 2019, seguindo toda a metodologia descrita nos subcapítulos antecedentes, com novos valores obtidos, também para o ano de 2019, com leves alterações no conjunto selecionado de indicadores. Assim, em primeiro lugar, mencionam-se as alterações efetuadas ao nível da seleção dos indicadores. São removidos os indicadores com periodicidade não anual, isto é, os indicadores que apenas são recolhidos aquando das recolhas censitárias:

- Indicador referente à taxa de iliteracia da população;
- Indicador referente ao número de indivíduos com nível de ensino superior; e,
- Indicador referente ao número de alojamentos não clássicos.

Depois desta remoção, foram acrescentados outros indicadores que estejam, de alguma forma relacionados com a vulnerabilidade social de uma população. Estes indicadores, assim como a sua relação com a vulnerabilidade social, e uma breve justificação dessa relação são apresentados na **Tabela 8**.

Tabela 8 - Indicadores adicionados para análise de robustez.

Indicador	Relação	Justificação
Taxa de retenção do ensino secundário	+	Uma maior taxa de retenção a nível do ensino secundário refletirá um menor nível de escolaridade do município.
Nº de edifícios construídos	-	A construção de novos edifícios traduz-se num aumento da rede de infraestruturas.
Saldo migratório	-	Um elevado saldo migratório traduz-se numa elevada entrada de migrantes no município, representando o elevado nível de qualidade de vida desse mesmo município.

Fonte: Elaboração Própria.

Tendo sido extraídos estes indicadores, apenas para o ano de 2019, aplica-se o procedimento metodológico anterior de forma a ser obtido um IVS que possa ser comparado ao índice principal e, desta forma, ser estabelecida uma análise de robustez.

Dá-se, assim, por concluído o capítulo Metodologia, procedendo-se com os Resultados consequentes desta etapa.

4. Resultados

Após definida a metodologia a ser seguida neste estudo, procede-se com a sua implementação com o objetivo de definir e calcular um índice que classifique a vulnerabilidade social de Portugal a nível municipal.

Durante este capítulo são, então, apresentados e discutidos os respetivos resultados consequentes da execução de cada etapa apresentada durante o capítulo da metodologia.

Este capítulo é iniciado com a análise da viabilidade dos novos dados e à determinação do número de componentes a serem selecionadas aquando do cálculo do IVS. Após esta etapa, são analisados os valores obtidos para o índice calculado e algumas medidas de dispersão e centralidade do mesmo. Por fim, este capítulo termina com o mapeamento do IVS ao longo da área geográfica selecionada para este estudo, assim como para os diversos anos analisados.

Relembra-se que a implementação de toda a metodologia descrita no capítulo anterior, assim como a obtenção dos resultados apresentados neste capítulo, foram realizadas com recurso ao software *Visual Studio Code* e à linguagem de programação *Python*, assim como todas as bibliotecas mencionadas durante esse mesmo capítulo.

4.1. Pré-Processamentos dos Dados

4.1.1. Limpeza dos Dados

Perante o conjunto de dados selecionados, não se verifica a necessidade da realização deste passo: não são verificados valores em falta, e os poucos *outliers* que se possam verificar consistem em observações cuja remoção não faz sentido. Desta forma, procede-se ao passo seguinte, a transformação dos dados.

4.2. Índice de Vulnerabilidade Social

4.2.1. Viabilidade dos Dados

Para se analisar a viabilidade dos dados, foram examinados os indicadores de cada ano, no entanto apenas são representados, neste subcapítulo, os dados referentes ao ano de 2009. Os resultados obtidos para os restantes anos são apresentados no **Anexo 1**.

Esta etapa é iniciada com a aplicação dos testes da Esfericidade de Bartlett, e de KMO aos dados de 2009. Os resultados daqui consequentes são registados na **Tabela 9**, onde se pode observar que da aplicação do Teste da Esfericidade de Bartlett é devolvido um *p-value* nulo, podendo-se confirmar a não existência de multicolinearidade. Quanto ao Teste de KMO, é obtido um valor de 0.851, o que permite concluir que os dados são adequados ao estudo em causa.

Tabela 9 - Resultados dos Testes de Viabilidade dos Dados para o ano de 2009.

Teste de Esfericidade de Bartlett	0,000
Teste KMO	0,851

Fonte: Elaboração Própria.

Sabendo assim que é possível e viável a aplicação da Análise de Componentes Principais, importa agora decidir quanto ao número de componentes a selecionar.

4.2.2. Determinação do Número de Componentes

Após a aplicação das três abordagens referidas durante o capítulo Metodologia, verifica-se que os resultados obtidos a partir destas, como se verifica de seguida, são bastante divergentes.

Com o Critério de Kaiser, onde se selecionam as componentes com um autovalor superior à unidade, para o ano de 2009, seriam selecionadas 8 componentes, o que, do ponto de vista de redução de dimensionalidade deste modelo, não é muito favorável, visto que apenas se passou de 15 variáveis para 8 componentes principais, como se observa na **Tabela 10**. No entanto, e por consequência desta fraca redução de dimensão, a variância explicada por estas 8 componentes iguala 91,35 pontos percentuais.

Tabela 10 - Autovalores das 15 Componentes Principais para o ano de 2009.

Componente	Autovalor
Componente 1	21,337
Componente 2	5,825
Componente 3	3,618
Componente 4	2,673
Componente 5	2,013
Componente 6	1,721
Componente 7	1,384
Componente 8	1,177
Componente 9	0,984
Componente 10	0,702
Componente 11	0,652
Componente 12	0,607
Componente 13	0,407
Componente 14	0,326
Componente 15	0,0850

Fonte: Elaboração Própria.

Quanto ao Teste de Cattell, deste resultaram os gráficos representados na **Figura 3**. Pode-se observar que, para o ano de 2009, o “cotovelo” do gráfico é bastante definido, sendo, de acordo com este, apenas selecionadas 2 componentes principais, e obtida uma variância explicada acumulada de 62,43%, o que não representa um valor suficientemente elevado para proceder com esta abordagem. Adicionalmente, para outros anos da análise, como o ano de 2019, por exemplo, este cotovelo já não se encontra tão definido, dificultando assim a seleção de um número de componentes a selecionar. Por estas razões, o Teste de Cattell é descartado como método de seleção de componentes principais.

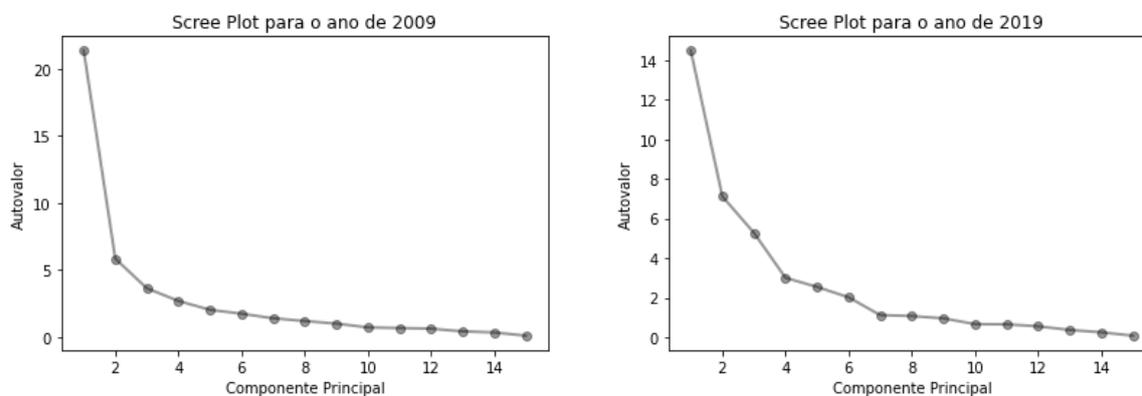


Figura 3 - Scree Plots dos anos de 2009 e 2019.

Fonte: Elaboração Própria.

Resta ainda o Critério de Benzécri, a partir do qual são selecionadas as 4 primeiras componentes principais, para a maioria dos anos, com exceção dos anos 2009 e 2010, onde são selecionadas apenas 3 componentes. Por coerência, opta-se por se selecionarem as 4 componentes principais para todos os anos analisados, mesmo nos dois anos excepcionais. No ano de 2009, com as 4 primeiras CPs selecionadas, atingem-se 76,88 pontos percentuais de variância explicada acumulada, como se pode observar na **Tabela 11**.

Tabela 11 - VEA das 15 Componentes Principais para o ano de 2009.

Componente	Variância Explicada Acumulada
Componente 1	49,04%
Componente 2	62,43%
Componente 3	70,74%
Componente 4	76,88%
Componente 5	81,51%
Componente 6	85,47%
Componente 7	88,65%
Componente 8	91,35%
Componente 9	93,61%
Componente 10	95,23%
Componente 11	96,72%
Componente 12	98,12%
Componente 13	99,05%
Componente 14	99,80%
Componente 15	100,00%

Fonte: Elaboração Própria.

Comparando este critério com o critério de Kaiser, opta-se por selecionar o critério de Benzécri como método de seleção do número de componentes principais durante este estudo, uma vez que se atinge uma quantidade superior de variância explicada acumulada com um menor número de componentes principais selecionadas.

Aplica-se este mesmo critério aos restantes anos, e apresentam-se os resultados daí consequentes no **Anexo 2**.

Dando por terminada esta etapa, conclui-se que serão selecionadas 4 componentes principais para cada ano analisado, sendo então obtidas as Variâncias Explicadas Acumuladas descritas na **Tabela 12**.

Tabela 12 - VEA das 4 Componentes Principais selecionadas.

Ano	Variância Explicada Acumulada
2009	76,88%
2010	76,26%
2011	75,55%
2012	74,43%
2013	74,14%
2014	73,84%
2015	74,17%
2016	73,62%
2017	74,96%
2018	75,49%
2019	74,41%

Fonte: Elaboração Própria.

4.2.3. Obtenção do Índice

Para ser possível a obtenção de um índice, foi aplicado o método de Análise de Componentes Principais com 4 CPs e foram obtidos os valores registados na **Tabela 13**. Estes valores são referentes aos autovalores e às variâncias explicadas de cada componente, assim como o total de variância explicada acumulada pelo conjunto selecionado. Ainda nesta tabela, é registado o peso atribuído a cada índice, para o ano de 2009, calculado através da fórmula presente na **Equação 10**, como referido anteriormente.

Tabela 13 - Autovalores e Variâncias das CPs para o ano de 2009.

Componente	Autovalor	Variância	VAE	Peso
1	21,337	49,04%	49,04%	0,638
2	5,825	13,39%	62,43%	0,174
3	3,618	8,31%	70,74%	0,108
4	2,673	6,14%	76,88%	0,080

Fonte: Elaboração Própria.

Com estes dados, é obtida a fórmula a ser utilizada para calcular o IVS, no ano de 2009, descrita na **Equação 12**.

$$IVS_{2009} = 0,638CP_1 + 0,174CP_2 + 0,108CP_3 + 0,089CP_4 \quad (12)$$

As fórmulas de cálculo do IVS para os restantes anos são obtidas de forma análoga a esta e são apresentadas no **Anexo 3**.

Com estas fórmulas é calculado o valor do Índice de Vulnerabilidade Social de cada município de Portugal Continental para todo o conjunto de anos analisados e, tendo estes valores sido obtidos pode-se então proceder às suas descrições e representações geográficas.

4.2.4. Descrição do Índice

Após os valores para o IVS terem sido calculados, considera-se necessário realizar uma análise das suas medidas de centralidade e dispersão, para os anos analisados, depois da sua normalização. Estas medidas, apresentadas na **Tabela 14**, permitem observar que os extremos mínimos e máximos deste índice – 0,000 e 1,000, respetivamente – ocorrem, ambos, no ano 2009, tendo-se, a partir deste momento, verificado um aumento dos valores mínimos e uma diminuição dos valores máximos, aproximando-se, em simultâneo, do valor médio, durante o período temporal analisado.

Tabela 14 - Medidas de Centralidade e Dispersão do IVS.

Ano	Média	Desvio-Padrão	Mínimo	Máximo
2009	0,358	0,191	0,000	1,000
2010	0,358	0,185	0,006	0,996
2011	0,358	0,161	0,008	0,937
2012	0,358	0,158	0,035	0,923
2013	0,358	0,152	0,042	0,916
2014	0,358	0,146	0,024	0,914
2015	0,358	0,132	0,055	0,835
2016	0,358	0,128	0,051	0,862
2017	0,358	0,128	0,096	0,797
2018	0,358	0,125	0,123	0,824
2019	0,358	0,128	0,129	0,810

Fonte: Elaboração Própria.

Ainda através desta tabela, e consequente do referido ao longo do parágrafo anterior, observa-se uma diminuição do valor verificado para o desvio-padrão, significando que os valores evoluem no sentido de uma menor dispersão ao longo dos anos.

Uma simples justificação para esta diminuição de amplitude da quantificação da vulnerabilidade social passa, por um lado, pela crescente preocupação governamental com a vulnerabilidade social e a consequente criação de medidas de combate à pobreza como, o aumento dos salários mínimos, a criação do Rendimento Social de Inserção, o Complemento Solidário para Idosos, o Abono de família, entre outras medidas, que provocam uma diminuição dos valores máximos verificados (Marques & Matos, 2016); e, por outro lado, o acumular de vulnerabilidades das sociedades, normalmente das regiões menos desenvolvidas onde as fragilidades das sociedades são sobrepostas dificultando os esforços e medidas de combate que possam ser tomadas, contribuindo para o aumento dos valores mínimos verificados (Costa et al., 2021).

4.2.5. Índice de Vulnerabilidade Social

Neste subcapítulo explora-se, de forma mais detalhada, a distribuição e evolução do IVS, quer ao longo da zona geográfica, quer ao longo do período temporal.

Tabela 15 - Medidas de centralidade do IVS distrital.

Distrito	Índice 2009			Índice 2019		
	Min	Méd	Máx	Min	Méd	Máx
Aveiro	0,091	0,205	0,369	0,208	0,252	0,357
Beja	0,296	0,484	0,692	0,303	0,452	0,597
Braga	0,079	0,193	0,407	0,174	0,274	0,465
Bragança	0,349	0,536	0,705	0,307	0,491	0,626
Castelo Branco	0,351	0,595	0,945	0,341	0,485	0,810
Coimbra	0,141	0,383	0,773	0,224	0,344	0,590
Faro	0,053	0,298	1,000	0,202	0,331	0,775
Guarda	0,304	0,529	0,779	0,301	0,444	0,534
Leiria	0,171	0,342	0,676	0,210	0,322	0,512
Lisboa	0,043	0,183	0,359	0,136	0,205	0,353
Portalegre	0,277	0,596	0,817	0,357	0,534	0,753
Porto	0,000	0,084	0,262	0,152	0,243	0,388
Santarém	0,110	0,367	0,752	0,141	0,327	0,514
Setúbal	0,104	0,217	0,465	0,129	0,255	0,406
Viana do Castelo	0,183	0,359	0,617	0,203	0,299	0,408
Vila Real	0,252	0,425	0,619	0,324	0,466	0,574
Viseu	0,242	0,385	0,553	0,284	0,398	0,532
Évora	0,296	0,461	0,675	0,267	0,431	0,604

Fonte: Elaboração Própria.

Em primeiro lugar, são analisadas as medidas de centralidade do índice por distrito, tanto para o primeiro ano de análise, como para o último ano. Estas medidas são registadas na **Tabela 15**, através da qual se observa, mais uma vez, e apesar de existirem exceções à regra, um aumento dos valores mínimos e uma diminuição dos valores máximos, de 2009 para 2019. Por outro lado, é também possível observar o local de ocorrência dos extremos do IVS: o limite mínimo – 0,000 – ocorre no distrito do Porto; e o limite máximo – 1,000 – ocorre no distrito de Faro, ambos em 2009, como já referido anteriormente.

Desta forma, procede-se, de seguida, à apresentação do IVS dos municípios dos distritos do Porto, e de Faro, na **Tabela 16**, onde se torna possível delimitar, ainda mais, a ocorrência destes extremos mínimo e máximo aos municípios de Paredes e de Alcoutim, respetivamente.

Tabela 16 - IVS para os distritos do Porto e de Faro.

Porto		Faro	
Município	Índice 2009	Município	Índice 2009
Gondomar	0,072	Albufeira	0,053
Maia	0,040	Alcoutim	1,000
Matosinhos	0,116	Aljezur	0,496
Paredes	0,000	Castro Marim	0,404
Porto	0,262	Faro	0,181
Póvoa de Varzim	0,107	Lagoa	0,144
Santo Tirso	0,160	Lagos	0,182
Trofa	0,058	Loulé	0,140
Valongo	0,003	Monchique	0,553
Vila do Conde	0,084	Olhão	0,113
Vila Nova de Gaia	0,095	Portimão	0,072
Amarante	0,127	São Brás de Alportel	0,233
Baião	0,216	Silves	0,256
Felgueiras	0,010	Tavira	0,313
Lousada	0,004	Vila do Bispo	0,458
Marco de Canaveses	0,073	Vila Real de Santo António	0,170
Paços de Ferreira	0,002		
Penafiel	0,078		

Fonte: Elaboração Própria.

De forma contrária ao que se poderia pensar, o centro urbanístico do distrito do Porto, o município Porto, não é o município onde se verifica a menor quantificação da vulnerabilidade social. Esta ocorrência deve-se ao facto dos municípios mais desenvolvidos, geralmente localizados no litoral do país, apresentarem um forte crescimento da população, e um

consequente aumento da densidade populacional, que se traduzem na degradação das condições das infraestruturas e serviços, quer pelo uso excessivo, quer por sobrelotação (Mendes, 2009). Assim, é comum que as regiões circundantes a grandes centros urbanos apresentem menor vulnerabilidade social que os próprios centros pois, apesar de também usufruírem do desenvolvimento e dinâmica industrial, comercial e territorial, não sofrem com as consequências geralmente associadas a estas.

Por outro lado, é compreensível que o valor de vulnerabilidade social máximo seja atingido em Alcútem, visto que este município pertence à região interior do país, uma zona caracterizada por um fraco dinamismo económico e menor desenvolvimento (Mendes et al., 2011).

No passo seguinte, foram criadas e atribuídas categorias de classificação a cada município de forma a facilitar a interpretação da distribuição geográfica da vulnerabilidade, assim como possibilitar uma perceção mais geral dessa mesma distribuição. Na **Tabela 17** observa-se essa distribuição e, apesar de, em ambos os anos, 2009 e 2019, a classe com mais ocorrências ser a classe “Baixa”, esta mostrou um significativo aumento ao longo dos anos, aumentando de 101 municípios, em 2009, para 154 municípios, em 2019. Desta tabela, destaca-se também a diminuição do número de municípios pertencentes à classe “Muito Baixa” que passou de 65 municípios, em 2009, para 19 municípios, em 2019, algo que era de esperar devido ao aumento dos valores mínimos do Índice de Vulnerabilidade Social calculado. Por último, refere-se ainda as diminuições simultâneas das classes referentes aos níveis mais elevados de vulnerabilidade que corroboram a diminuição dos valores máximos do IVS, verificada anteriormente.

Tabela 17 - Distribuição das classes pelos municípios.

	2009	%	2019	%
Muito Baixa	65	23,4%	19	6,8%
Baixa	101	36,3%	154	55,4%
Moderada	83	29,9%	96	34,5%
Elevada	25	9,0%	8	2,9%
Muito Elevada	4	1,4%	1	0,4%

Fonte: Elaboração Própria.

Estes resultados são semelhantes aos resultados obtidos no estudo de Frigerio et al. (2018) onde, com o passar do tempo se verifica uma maior homogeneidade de resultados e de distribuição da vulnerabilidade social.

4.2.6. Cartogramas

Como forma de facilitar a visualização dos resultados obtidos, foram criados diversos cartogramas, ou mapeamentos desses resultados. Em primeiro lugar, são representados os valores do IVS e as respetivas classes, na **Figura 4** e **Figura 5**, respetivamente, a partir das quais se observa uma forte heterogeneidade de resultados.

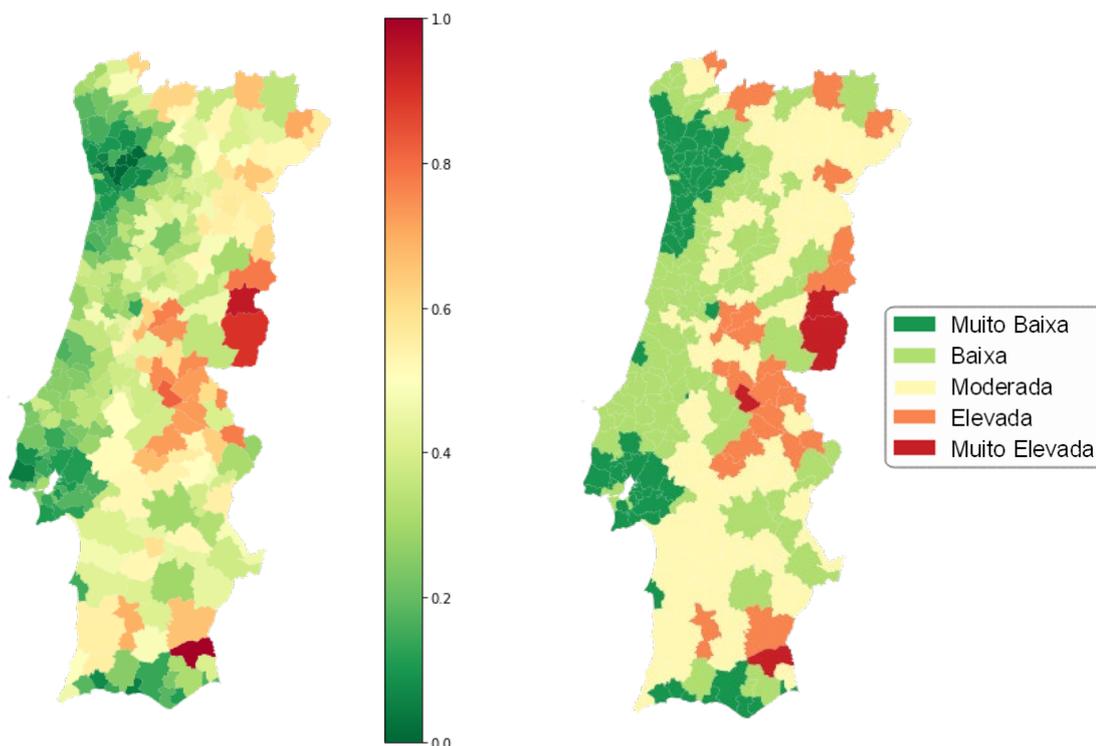


Figura 4 - Mapa nacional do IVS para 2009.

Fonte: Elaboração Própria.

Figura 5 - Mapa nacional das classes do IVS para 2009.

Fonte: Elaboração Própria.

Através destas figuras, facilmente se apercebe da forte disparidade notada entre o litoral e o interior do país: o litoral marcado por menores valores de vulnerabilidade social, e o interior com distintas manchas de valores elevados deste indicador. A **Figura 4**, permite identificar, de forma bastante clara, os extremos mencionados no subcapítulo anterior: Paredes, no distrito do Porto, identificado com verde-escuro, e Alcoutim, no distrito de Faro, representado a mancha a vermelho mais escuro no sul do mapa.

É, ainda, possível observar, na **Figura 5**, como os grandes centros urbanos, como Porto e Lisboa, se encontram rodeados por municípios com um menor IVS que os próprios, facto já descrito anteriormente. Ainda a partir desta figura, observa-se uma mancha representada

a amarelo a sul do país, mesmo nos municípios litorais. Esta aspeto deve-se ao facto de desta região ser uma região mais dependente, no geral, do setor primário e apresentar uma consequente menor dinâmica económica que o resto do país (Mendes, 2009).

Para o ano 2019, representado na **Figura 6** e **Figura 7**, nota-se uma maior homogeneidade dos dados, confirmada pela diminuição dos valores obtidos para o desvio-padrão ao longo dos anos analisados, verificada e justificada também no capítulo anterior.

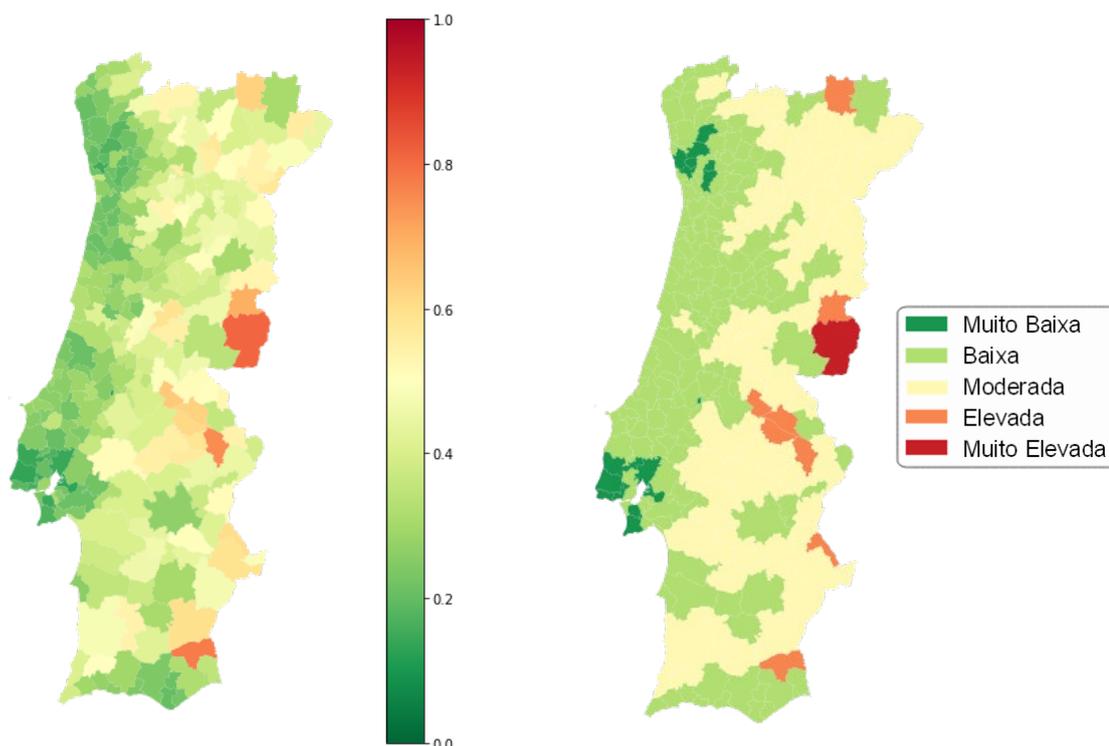


Figura 6 - Mapa nacional do IVS para 2019.

Fonte: Elaboração Própria.

Figura 7 - Mapa nacional das classes do IVS para 2019.

Fonte: Elaboração Própria.

Passando a uma exploração dos cartogramas dos distritos onde se verificaram os limites mínimo e máximo, em 2009: Porto, representado na **Figura 8** e **Figura 9**, e Faro, na **Figura 10** e **Figura 11**.

Começando pelo distrito do Porto, na **Figura 8** verifica-se que Paredes – que passou de um IVS de 0,000 para 0,189 – e os municípios a ele adjacentes, em 2009, constituem a zona com menor vulnerabilidade social deste distrito, e, no lado contrário, o município do Porto com o maior valor de IVS verificado. Na **Figura 9**, confirma-se, uma vez mais, o aumento dos valores mínimos de IVS, tendo grande parte dos municípios do distrito analisado

passado, de 2009 a 2019, da classe de vulnerabilidade social “Muito Baixa” para a classe “Baixa”.

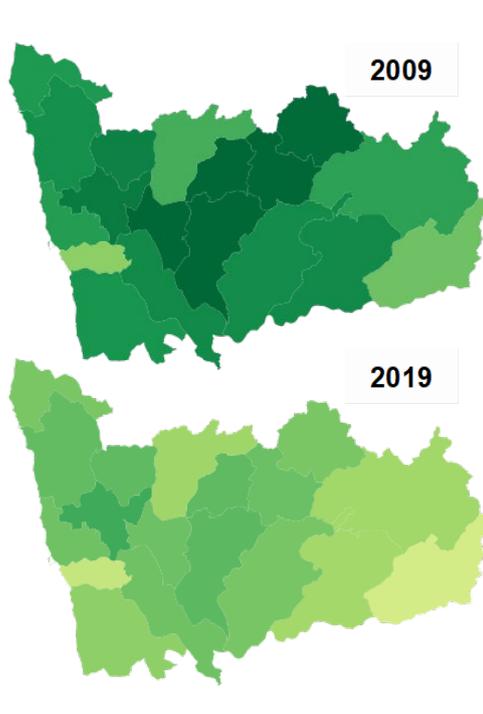


Figura 8 - Valores do IVS para o distrito do Porto em 2009 e 2019.

Fonte: Elaboração Própria.

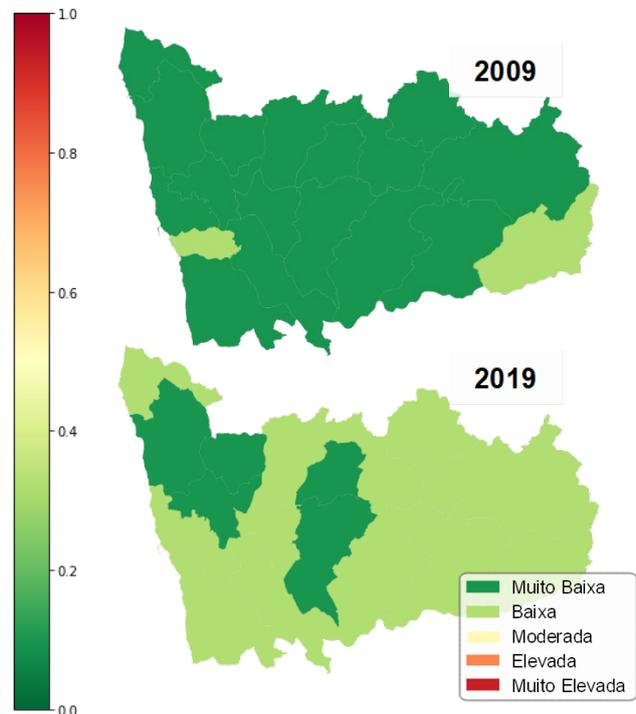


Figura 9 - Classes do IVS para o distrito do Porto em 2009 e 2019.

Fonte: Elaboração Própria.

Quanto ao distrito de Faro, observa-se também, na **Figura 10**, uma clara diminuição das zonas representadas a verde mais escuro, conseqüente do aumento dos valores mínimos do IVS. No entanto, é ainda verificada uma diminuição das zonas representadas a amarelo e, o município de Alcoutim, representado a vermelho mais escuro em 2009 – com um IVS de 1,000 –, passa a ser representado a laranja em 2019 – diminuindo para um IVS de 0,775 –, o que se traduz numa diminuição dos níveis mais elevados de vulnerabilidade social verificado nos municípios. Pela **Figura 11**, confirma-se a homogeneidade e convergência dos valores do IVS para o seu valor médio, já verificadas anteriormente.

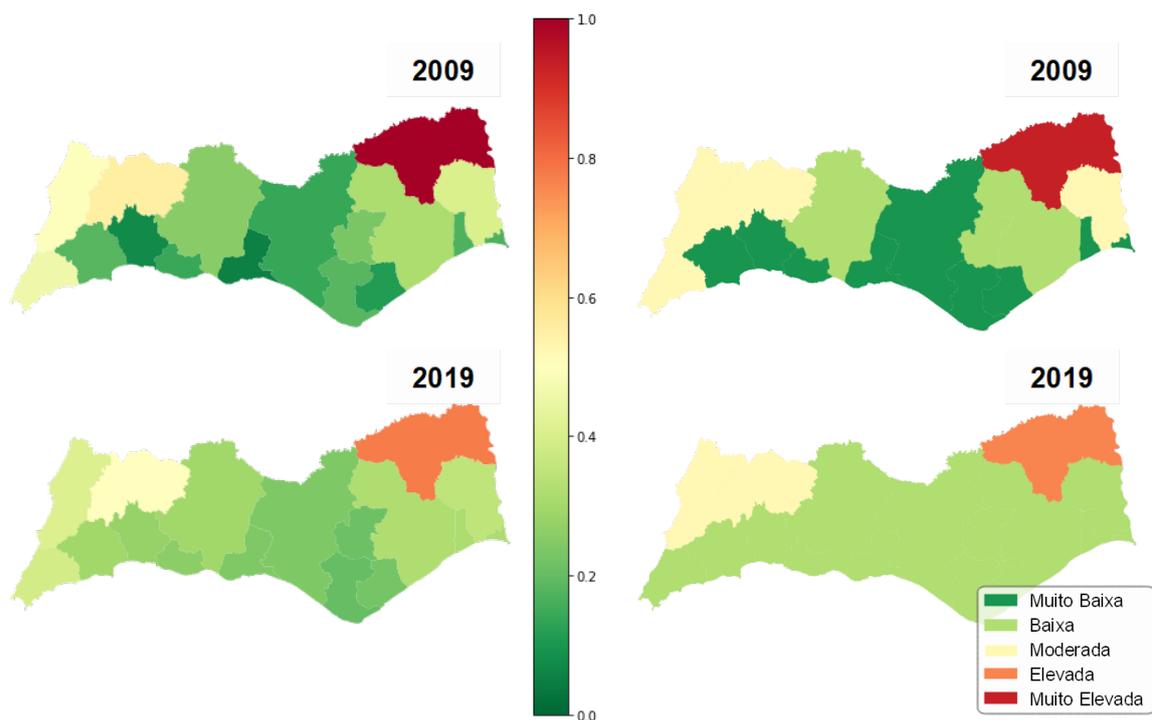


Figura 10 - Valores do IVS para o distrito de Faro em 2009 e 2019.

Fonte: Elaboração Própria.

Figura 11 - Classes do IVS para o distrito de Faro em 2009 e 2019.

Fonte: Elaboração Própria.

Assim, de uma forma geral, e em jeito de conclusão deste subcapítulo, pode-se constatar que os municípios localizados no litoral do país apresentam menores quantificações de vulnerabilidade social que os municípios localizados no interior. Este facto, devido ao referido durante a Revisão de Literatura, já era esperado, no entanto, com a observação dos cartogramas, quer nacionais, quer distritais, confirma-se tal facto e atribui-se a sua justificação à dinâmica económica e social sentida a litoral que se sobrepõe à sentida no interior do país (Mendes et al., 2011).

Para além disto, também se destaca o facto destes resultados mostraram que, em geral, as capitais distritais demonstram uma menor vulnerabilidade social que os restantes municípios envolventes, devido, em parte, à concentração de recursos e infraestruturas nessas capitais (Mendes et al., 2011). No estudo de Davino et al. (2021), foi concluído que as zonas periféricas em Roma são as mais afetadas pela condição de vulnerabilidade quando comparadas com a capital, podendo-se observar, assim, uma segmentação espacial

elevada caracterizada por uma bipartição na cidade, entre centro e periferia: baixa vulnerabilidade no centro e alta vulnerabilidade nas zonas periféricas orientais e ocidentais.

4.3. Análise de Robustez

O processo metodológico de obtenção do IVS, aqui realizado, é fortemente semelhante ao efetuado anteriormente, apenas se identifica uma disparidade: a existência de dados em falta.

Ao analisar a existência de dados em falta, verifica-se que o único indicador onde ocorre esta situação corresponde ao indicador referente à taxa de retenção do ensino secundário, e deve-se ao facto de, nos municípios em que ocorre esta falta de valores, não existirem estabelecimentos escolares deste nível de ensino.

Desta forma, procedeu-se à substituição desses dados pela média distrital do indicador em causa, para todos os anos analisados onde foram, então, substituídos 23 valores em falta pela respetiva média distrital. De seguida, verificou-se que já não existiam dados em falta, tendo este procedimento sido aplicado corretamente.

A metodologia efetuada posteriormente não apresenta mais divergências, pelo que se procede à apresentação dos principais resultados, consequentes desta, para o novo IVS.

Em primeiro lugar, é analisada a viabilidade dos dados, através do Teste de Esfericidade de Bartlett e do Teste KMO, para os quais são apresentados os respetivos resultados na **Tabela 18**. Os resultados obtidos permitem retirar a conclusão de que os dados em causa são viáveis à aplicação do método ACP, sendo assim possível prosseguir com esta análise.

Tabela 18 - Viabilidade dos dados de análise de robustez.

Teste de Esfericidade de Bartlett	0,000
Teste KMO	0,732

Fonte: Elaboração Própria.

De seguida, é aplicado o Critério de Benzécri de forma a ser determinado o número de componentes principais a ser selecionado. Pela observação dos valores representados na **Tabela 19**, percebe-se que a 5ª componente principal atinge o limiar de variância explicada acumulada definido pelo critério referido anteriormente (70%) pelo que se conclui pela seleção de 5 CPs. Estas 5 componentes permitem atingir uma VEA de 73,97%.

Tabela 19 - VEAs das CPs da análise de robustez.

Componente	Variância Explicada Acumulada
Componente 1	28,94%
Componente 2	45,28%
Componente 3	59,41%
Componente 4	66,75%
Componente 5	73,97%
Componente 6	79,16%
Componente 7	83,98%
Componente 8	88,14%
Componente 9	91,23%
Componente 10	93,53%
Componente 11	95,60%
Componente 12	97,57%
Componente 13	99,14%
Componente 14	99,81%
Componente 15	100,00%

Fonte: Elaboração Própria.

Com estes resultados, podemos perceber que apesar de ser selecionada uma componente extraordinária, o total de variância explicada acumulada obtida é menor ao valor obtido com os dados originais, dando, desde já, a entender que talvez estes dados não sejam tão adequados ao estudo a ser efetuado quanto os originais.

Posteriormente, é apresentada a fórmula de obtenção do IVS, representada na **Equação 13**, depois de aplicado o método ACP com 5 componentes principais selecionadas.

$$IVS'_{2019} = 0,391CP_1 + 0,223CP_2 + 0,189CP_3 + 0,099CP_4 + 0,098CP_5 \quad (13)$$

De seguida são apresentadas, na **Tabela 20**, algumas medidas de centralidade e de dispersão do novo IVS, assim como do IVS original, de forma que seja possível estabelecer uma comparação entre estas.

Tabela 20 - Medidas de centralidade e dispersão para a análise de robustez.

IVS	Média	Desvio-Padrão
Original	0,358	0,128
Análise de Robustez	0,318	0,149

Fonte: Elaboração Própria.

Com estes valores, percebe-se que o novo IVS apresenta uma menor média das quantificações da vulnerabilidade social dos municípios, no entanto, estas encontram-se mais dispersas ao longo da sua amplitude, como indicado pelo desvio-padrão superior ao verificado com o IVS original.

Tabela 21 - Medidas de centralidade distritais para a análise de robustez.

Distrito	Índice Original			Índice Análise de Robustez		
	Min	Méd	Máx	Min	Méd	Máx
Aveiro	0,208	0,252	0,357	0,130	0,217	0,395
Beja	0,303	0,452	0,597	0,222	0,391	0,619
Braga	0,174	0,274	0,465	0,104	0,221	0,449
Bragança	0,307	0,491	0,626	0,267	0,430	0,580
Castelo Branco	0,341	0,485	0,810	0,295	0,451	0,843
Coimbra	0,224	0,344	0,590	0,148	0,324	0,558
Faro	0,202	0,331	0,775	0,141	0,325	1,000
Guarda	0,301	0,444	0,534	0,255	0,435	0,584
Leiria	0,210	0,322	0,512	0,175	0,307	0,548
Lisboa	0,136	0,205	0,353	0,036	0,127	0,354
Portalegre	0,357	0,534	0,753	0,285	0,526	0,684
Porto	0,152	0,243	0,388	0,057	0,166	0,330
Santarém	0,141	0,327	0,514	0,030	0,292	0,488
Setúbal	0,129	0,255	0,406	0,000	0,181	0,351
Viana do Castelo	0,203	0,299	0,408	0,166	0,263	0,376
Vila Real	0,324	0,466	0,574	0,238	0,381	0,521
Viseu	0,284	0,398	0,532	0,255	0,366	0,528
Évora	0,267	0,431	0,604	0,232	0,413	0,589

Fonte: Elaboração Própria.

Posteriormente, procede-se à comparação das medidas de centralidade de cada distrito para os IVS calculados. Estas medidas, representadas na **Tabela 21**, permitem observar que, apesar do valor mínimo verificado por cada Índice de Vulnerabilidade Social ter ocorrido no mesmo distrito, em Setúbal – tendo registado um valor de 0,129 com o IVS original e um valor de 0,000 com o novo IVS –, o mesmo não acontece para o valor máximo. No IVS

original, o extremo máximo tem ocorrência no distrito de Castelo Branco – com um valor de 0,810 –, enquanto, no novo IVS, este tem ocorrência no distrito de Faro – atingindo os 1,000.

Com isto, procede-se ao mapeamento dos valores obtidos para as quantificações da vulnerabilidade social com cada IVS, no distrito de Setúbal, sendo obtidas as **Figura 12** e a **Figura 13**. Comparando estas duas figuras, facilmente se nota que o índice criado para a análise de robustez retribui manchas mais “verdes” que o IVS original, ou seja, os valores resultantes do novo IVS são inferiores aos valores obtidos através do IVS original. No entanto, com ambos os índices, o valor mínimo registado no distrito (e, neste caso, também no país) ocorre no município de Alcochete – 0,129 para o IVS original e 0,000 para o novo IVS –, e o valor máximo – 0,406 para o IVS original e 0,351 para o novo IVS – no município de Alcácer do Sal.

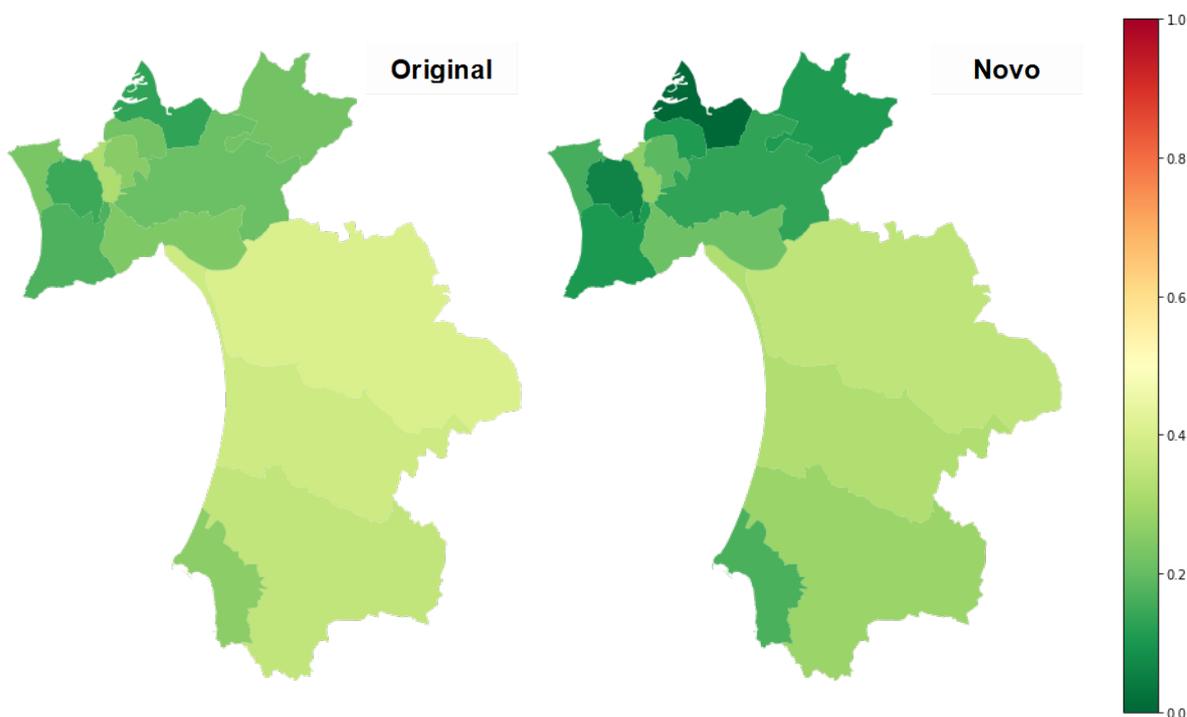


Figura 12 - Valores do IVS original para o distrito de Setúbal.

Fonte: Elaboração Própria.

Figura 13 - Valores do IVS de análise de robustez para o distrito de Setúbal.

Fonte: Elaboração Própria.

Procedendo à análise das classes resultantes de cada índice, através da **Tabela 22** é possível notar que a alteração de classes mais notável, entre estas duas ferramentas, é referente ao aumento do número de classes do tipo “Muito Baixa” do IVS original para o IVS criado neste subcapítulo que passa de 6,8 pontos percentuais para 23 pontos percentuais.

Esta alteração também pode ser notada pela comparação da **Figura 14** com a **Figura 15**, onde se nota a substituição das manchas verde-escuras por um tom mais claro.

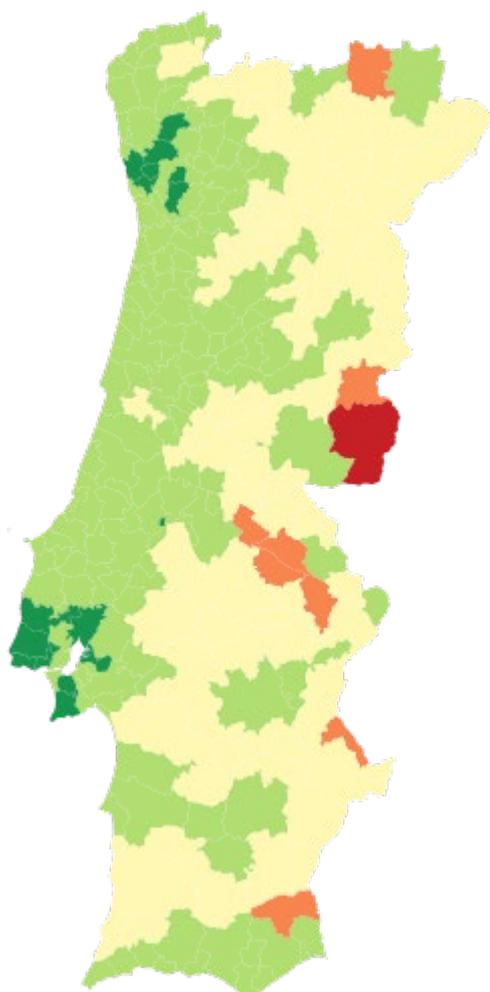


Figura 14 - Mapa nacional das classes do IVS original.

Fonte: Elaboração Própria.

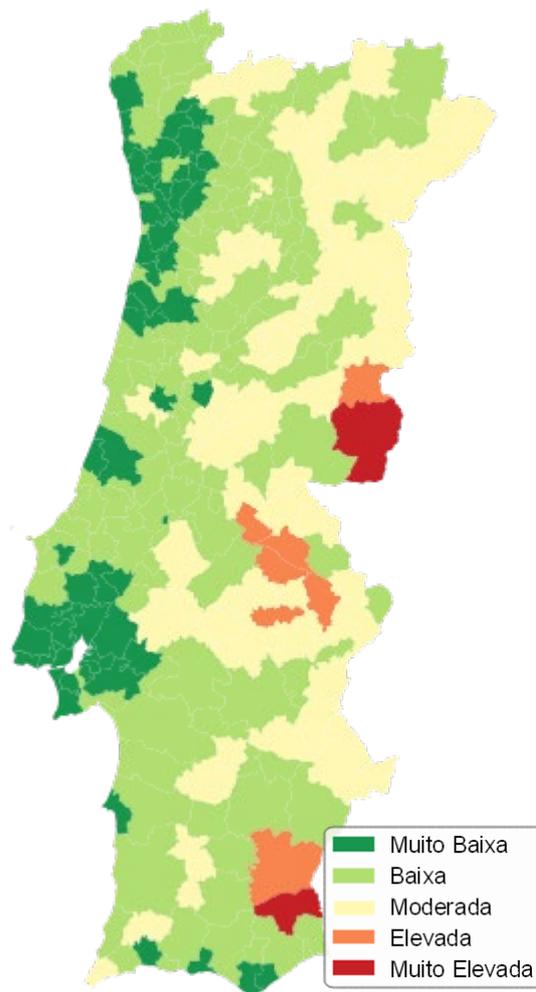


Figura 15 - Mapa nacional das classes do IVS de análise de robustez.

Fonte: Elaboração Própria.

Tabela 22 - Distribuição das classes pelos municípios para a análise de robustez.

	Original	%	Análise de Robustez	%
Muito Baixa	19	6,8%	64	23,0%
Baixa	154	55,4%	134	48,2%
Moderada	96	34,5%	71	25,5%
Elevada	8	2,9%	7	2,5%
Muito Elevada	1	0,4%	2	0,7%

Fonte: Elaboração Própria.

Apesar das diferenças e disparidades entre estes dois Índices de Vulnerabilidade Social não serem extremamente notórias, é possível concluir que pequenas alterações da informação que os origina causa alguns distúrbios nos resultados produzidos por estes.

5. Conclusão

Durante esta secção são apresentadas as principais conclusões resultantes do trabalho desenvolvido ao longo deste estudo, após uma breve recapitulação das etapas realizadas ao longo do mesmo. De seguida, são também apresentadas algumas limitações encontradas aquando da realização do estudo e processo metodológico usado, e, por último, são enunciados alguns passos e motivações para investigações futuras que possam dar continuidade à análise efetuada durante esta dissertação.

5.1. Principais Conclusões

Este estudo teve como objetivo a criação de um índice que permitisse a quantificação da vulnerabilidade social dos municípios portugueses, ao longo de um período temporal de 11 anos, de 2009 a 2019, e que, desta forma, auxiliasse a criação e orientação de políticas públicas de combate à vulnerabilidade social.

O primeiro passo deste estudo passou por realizar uma análise da literatura existente relacionada com este tema, com dois principais objetivos. Primeiramente, pretendeu-se proporcionar, ao leitor, a informação e conhecimentos necessários à leitura deste estudo e, em segundo lugar, identificar os indicadores que afetam e alteram a vulnerabilidade social de determinada sociedade, e aferir de que forma é possível quantificar essa vulnerabilidade. Assim, a Revisão de Literatura reúne, então, um conjunto de definições, conceitos e métodos através dos quais foi concluir que os índices de vulnerabilidade social, apesar de possuírem algumas inconveniências, consistem, de forma geral, numa ferramenta bastante atrativa e útil. Através da literatura analisada, rapidamente se apercebeu que o método mais utilizado para a criação deste índice passa pela Análise de Componentes Principais, tendo este sido o processo selecionado a aplicar durante a execução deste estudo.

Para a ACP ser aplicada, foram extraídos, durante o capítulo Dados e Metodologia, diversos indicadores relacionados com determinadas características da sociedade, para a dimensão selecionada nesta análise e, posteriormente, procedeu-se ao pré-processamento desses indicadores. Este passo consistiu na limpeza e transformação dos dados extraídos, como a remoção de dados em falta e a normalização dos dados, de forma que fosse possível avançar com o processo metodológico. De seguida, foi analisada a viabilidade do conjunto de dados e, quando verificada, foi aplicada a Análise de Componentes Principais a partir da qual foram extraídas 4 componentes principais, para cada ano. Estas 4 componentes foram

agregadas, dando origem a um índice compósito de quantificação da vulnerabilidade social de cada município, o IVS. Os resultados obtidos foram mapeados em cartogramas, juntamente com categorias de classificação da vulnerabilidade.

Com estes mapeamentos facilmente se concluiu que os municípios localizados no litoral do país apresentam, no geral, menores níveis e valores de vulnerabilidade que os municípios localizados no interior.

Concluiu-se, também, que os grandes centros urbanos, nomeadamente o município de Lisboa e o município do Porto, apresentam um maior IVS quando comparados com as respetivas regiões circundantes. Este facto deve-se à forte concentração demográfica acompanhada de recursos e serviços que não conseguiram acompanhar esse crescimento populacional. Por outras palavras, nestes centros urbanos, os elevados valores do Índice de Vulnerabilidade Social verificados decorrem da forte dinâmica demográfica e da consequente degradação e sobrelotação das infraestruturas e serviços (Mendes, 2009).

Em geral, as capitais distritais (com exceção dos municípios de Lisboa e do Porto) apresentam uma menor vulnerabilidade social que os restantes municípios do distrito, uma conclusão concordante com as resultantes dos estudos de Mendes et al. (2011) e Davino et al. (2021).

É, ainda, retirada a conclusão de que, com a atribuição de classes categóricas aos municípios, de acordo com a sua quantificação de vulnerabilidade social, concluiu-se, tal como o estudo de Frigerio et al. (2018), que os resultados obtidos evoluem, ao longo do período analisado, para valores mais homogéneos.

Por fim, e quanto à análise de robustez realizada, concluiu-se que a metodologia aplicada durante este estudo é, de facto, adequada e que, conforme era de esperar, pequenas alterações da informação originária ao IVS provocam distúrbios significativos nos resultados produzidos por este, remetendo para as inconveniências resultantes dos índices sociais mencionadas durante a revisão de literatura.

Com isto, pode-se concluir que a ferramenta criada durante esta dissertação, apesar das suas limitações e espaço para melhorias, a serem referidas nos subcapítulos seguintes, poderá consistir num ponto de partida de inclusão da vulnerabilidade social na definição e implementação de políticas públicas por parte dos decisores, com vista à redução dos níveis de vulnerabilidade social. Isto é, quando identificado um município com um elevado valor de IVS, este deve ser analisado, assim como as suas características, de forma a ser identificada a fonte dessa vulnerabilidade. Caso esse valor elevado seja proveniente de uma quantidade

significativa de população desempregada, poderão ser tomadas medidas que proporcionem uma maior quantidade e qualidade de emprego, ou até qualificação profissional da população, por exemplo (Marques et al., 2017). Caso a vulnerabilidade seja consequente de um reduzido número de infraestruturas, deverão ser realizados investimentos de melhoria ou criação dessas infraestruturas e prestado o devido apoio social (Mendes et al., 2011). Estes dois exemplos representam apenas uma fração do estudo e análise que deve ser realmente efetuada aquando da definição das políticas públicas, no entanto, é claro o apoio e orientação que este índice poderá oferecer nesse processo.

5.2. Limitações

A primeira limitação encontrada durante a realização deste estudo surgiu durante a recolha e extração de dados a serem utilizados na criação do IVS, pois os dados de acesso público disponíveis nem sempre são recolhidos ou organizados para a divisão geográfica pretendida, estando, muitos deles, apenas disponíveis para as NUTS II. Desta forma, apesar ser pretendida a recolha de dados mais específicos a certas características da sociedade, estes nem sempre estavam disponíveis, impossibilitando, assim, a inclusão de outros indicadores relevantes e identificados na literatura.

Outra limitação deste índice passa pelo valor resultante deste apenas transmitir a informação de quais os municípios se encontram mais ou menos vulneráveis, em comparação com os restantes, não fornecendo informação sobre qual a área em que essa vulnerabilidade se traduz.

5.3. Desenvolvimentos Futuros

Um primeiro desenvolvimento futuro a aplicar este estudo poderá passar pela modificação dos indicadores selecionados, selecionando outros mais orientados para um tipo específico de dano ou risco como, por exemplo, alterações climáticas, incêndios, secas, etc.

Outro desenvolvimento que poderá resultar deste estudo consiste na utilização e adaptação deste índice de forma a possibilitar a comparação de vulnerabilidades entre diferentes países, no entanto, este desenvolvimento requer bases de dados complexas e coerentes entre esses países que, geralmente, não são disponibilizadas.

Por último, uma outra abordagem de um possível desenvolvimento futuro deste estudo passa por ultrapassar a última limitação referida no subcapítulo anterior: a construção de um índice que transmitisse não só o nível de vulnerabilidade social do território geográfico em causa, mas que também indicasse qual a área mais responsável por essa vulnerabilidade, permitindo, desta forma, uma melhor orientação e priorização de medidas de prevenção e políticas públicas.

Referências

- Adger, W. N. (2006). Vulnerability. *Global Environmental Change*, 16(3), 268–281. <https://doi.org/10.1016/J.GLOENVCHA.2006.02.006>
- Afonso, A. (2021). *Avaliação da suscetibilidade, exposição e vulnerabilidade aos ravinamentos na cidade de Nacala em Moçambique* [Tese de Doutoramento, Universidade de Lisboa]. <https://repositorio.ul.pt/handle/10451/47780>
- Alwang, J., Siegel, P. B., & Jorgensen, S. L. (2001). Vulnerability: a view from different disciplines. *Social Protection Discussion Papers and Notes - The World Bank*, 23304. <https://econpapers.repec.org/paper/wbkhdnsu/23304.htm>
- Aprilliant, A. (2021). *The Factor Analysis for Constructing a Composite Index*. Acedido a 03 de março de 2022. <https://medium.com/analytics-vidhya/the-factor-analysis-for-constructing-a-composite-index-2496686fc54c>
- Aung, T. S., Fischer, T. B., & Wang, Y. (2021). Conceptualization of Health and Social Vulnerability of Marginalized Populations During Covid-19 Using Quantitative Scoring Approach. *Journal of Immigrant & Refugee Studies*, 20(1), 1–16. <https://doi.org/10.1080/15562948.2021.1882023>
- Avila-Vera, M., Rangel-Blanco, L., & Picazzo-Palencia, E. (2012). Application of Principal Component Analysis as a Technique to Obtain a Social Vulnerability Index for the Design of Public Policies in Mexico. *Open Journal of Social Sciences*, 8, 130–145. <https://doi.org/10.4236/jss.2020.89009>
- Balica, S. F., Wright, N. G., & van der Meulen, F. (2012). A flood vulnerability index for coastal cities and its use in assessing climate change impacts. *Natural Hazards*, 64(1), 73–105. <https://doi.org/10.1007/S11069-012-0234-1/TABLES/6>
- Bohle, H. G. (2001). Vulnerability and Criticality: Perspectives from Social Geography. *International Human Dimensions Programme*, 2(1), 3–5. https://www.researchgate.net/publication/281020579_Vulnerability_and_Criticality_Perspectives_from_Social_Geography
- Bohle, H. G., Downing, T. E., & Watts, M. J. (1994). Climate change and social vulnerability: Toward a sociology and geography of food insecurity. *Global Environmental Change*, 4(1), 37–48. [https://doi.org/10.1016/0959-3780\(94\)90020-5](https://doi.org/10.1016/0959-3780(94)90020-5)
- Brems, M. (2017). *A One-Stop Shop for Principal Component Analysis*. Towards Data Science. Acedido a 29 de abril de 2022. <https://towardsdatascience.com/a-one-stop-shop-for-principal-component-analysis-5582fb7e0a9c>
- Cambridge Dictionary. (2022). *Index - Definition in the Cambridge English Dictionary*. Acedido a 14 de setembro de 2022. <https://dictionary.cambridge.org/us/dictionary/english/index>
- Chakraborty, J., Tobin, G. A., & Montz, B. E. (2005). Population Evacuation: Assessing Spatial Variability in Geophysical Risk and Social Vulnerability to Natural Hazards.

Natural Hazards Review, 6(1), 23–33. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2005\)6:1\(23\)](https://doi.org/10.1061/(ASCE)1527-6988(2005)6:1(23))

- Chen, W., Cutter, S. L., Emrich, C. T., & Shi, P. (2013). Measuring social vulnerability to natural hazards in the Yangtze River Delta region, China. *International Journal of Disaster Risk Science*, 4(4), 169–181. <https://doi.org/10.1007/s13753-013-0018-6>
- Comfort, L., Wisner, B., Cutter, S. L., Pulwarty, R. S., Hewitt, K., Oliver-Smith, A., Wiener, J., Fordham, M., Peacock, W., & Krimgold, F. (1999). Reframing disaster policy: the global evolution of vulnerable communities. *Environmental Hazards*, 1(1), 39–44. <https://doi.org/10.3763/EHAZ.1999.0105>
- Costa, S., Santos, M., & Guerra, I. (2021). *Uma imensidão de vidas – dez anos de percursos de pobreza em Lisboa*. Acedido a 08 de novembro de 2022. <https://observatorio-lisboa.eapn.pt/lisboa-em-analise/barometro/>
- Cumberbatch, J., Drakes, C., Mackey, T., Nagdee, M., Wood, J., Degia, A. K., & Hinds, C. (2020). Social Vulnerability Index: Barbados – A Case Study. *Coastal Management*, 48(5), 505–526. <https://doi.org/10.1080/08920753.2020.1796193>
- Cutter, S. L. (2001). The Changing Nature of Risks and Hazards. Em S. L. Cutter (Ed.), *American Hazardscapes: The Regionalization of Hazards and Disasters* (pp. 1–12). Joseph Henry Press. <https://doi.org/10.17226/10132>
- Cutter, S. L. (2003). The Vulnerability of Science and the Science of Vulnerability. *Annals of the Association of American Geographers*, 93(1), 1–12. <https://doi.org/10.1111/1467-8306.93101>
- Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social Vulnerability to Environmental Hazards. *Social Science Quarterly*, 84(2), 242–261. <https://doi.org/10.1111/1540-6237.8402002>
- Davies, S. (1996). *Adaptable livelihoods: coping with food insecurity in the Malian Sahel*. Macmillan Press.
- Davino, C., Gherghi, M., Sorana, S., & Vistocco, D. (2021). Measuring Social Vulnerability in an Urban Space Through Multivariate Methods and Models. *Social Indicators Research*, 157(3), 1179–1201. <https://doi.org/10.1007/S11205-021-02680-0/FIGURES/9>
- Dintwa, K. F., Letamo, G., & Navaneetham, K. (2019). Quantifying social vulnerability to natural hazards in Botswana: An application of cutter model. *International Journal of Disaster Risk Reduction*, 37, 101189. <https://doi.org/10.1016/J.IJDRR.2019.101189>
- Direção-Geral do Território. (2022). *Carta Administrativa Oficial de Portugal*. Acedido a 24 de maio de 2022. <https://www.dgterritorio.gov.pt/cartografia/cartografia-tematica/caop>
- Flanagan, B. E., Gregory, E. W., Hallisey, E. J., Heitgerd, J. L., & Lewis, B. (2011). A Social Vulnerability Index for Disaster Management. *Journal of Homeland Security and Emergency Management*, 8. <https://doi.org/10.2202/1547-7355.1792>

- Fraser, T. (2021). Japanese social capital and social vulnerability indices: Measuring drivers of community resilience 2000–2017. *International Journal of Disaster Risk Reduction*, 52. <https://doi.org/10.1016/J.IJDRR.2020.101965>
- Frigerio, I., Carnelli, F., Cabinio, M., & de Amicis, M. (2018). Spatiotemporal Pattern of Social Vulnerability in Italy. *International Journal of Disaster Risk Science*, 9(2), 249–262. <https://doi.org/10.1007/S13753-018-0168-7/FIGURES/6>
- Frischen, J., Meza, I., Rupp, D., Wietler, K., & Hagenlocher, M. (2020). Drought Risk to Agricultural Systems in Zimbabwe: A Spatial Analysis of Hazard, Exposure, and Vulnerability. *Sustainability*, 12(3), 775. <https://doi.org/10.3390/SU12030752>
- Gallopin, G. C. (2006). Linkages between vulnerability, resilience, and adaptive capacity. *Global Environmental Change*, 16(3), 293–303. <https://doi.org/10.1016/J.GLOENVCHA.2006.02.004>
- García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., & Herrera, F. (2016). Big data preprocessing: methods and prospects. *Big Data Analytics*, 1(1). <https://doi.org/10.1186/s41044-016-0014-0>
- Gomes, R. F. C. (2018). *Social vulnerability assessment in Madeira Island* [NOVA Information Management School]. <http://hdl.handle.net/10362/40908>
- Guillard-Gonçalves, C., Cutter, S. L., Emrich, C. T., & Zêzere, J. L. (2014). Application of Social Vulnerability Index (SoVI) and delineation of natural risk zones in Greater Lisbon, Portugal. <Http://Dx.Doi.Org/10.1080/13669877.2014.910689>, 18(5), 651–674. <https://doi.org/10.1080/13669877.2014.910689>
- Han, J., Kamber, M., & Pei, J. (2012). Data Preprocessing. Em *Data Mining: Concepts and Techniques* (3.^a ed., pp. 83–124). British Library. <http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>
- Holland, S. M. (2019). *Principal Components Analysis (PCA)*. Department of Geology, University of Georgia. Acedido a 29 de abril de 2022. <https://strata.uga.edu/software/pdf/pcaTutorial.pdf>
- Hufschmidt, G. (2011). A comparative analysis of several vulnerability concepts. *Natural Hazards*, 58(2), 621–643. <https://doi.org/10.1007/S11069-011-9823-7/FIGURES/1>
- Instituto Nacional de Estatística. (2013). *O que é o Índice de Bem-estar (IBE)*. Acedido a 06 de maio de 2022. https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indbemestar&xlang=pt
- Karaye, I. M., & Horney, J. A. (2020). The Impact of Social Vulnerability on COVID-19 in the U.S.: An Analysis of Spatially Varying Relationships. *American Journal of Preventive Medicine*, 59(3), 317–325. <https://doi.org/10.1016/j.amepre.2020.06.006>
- Krause, H. (2019). *Real World Data Dispatch: Tools and Stories You Need to See - Message from Heather*. Datassist. Acedido a 15 de setembro de 2022.

<https://mailchi.mp/idadassist/lets-bring-big-data-to-the-non-profit-world-3558453?e=74c12dde04>

- Marques, T. S., & Matos, F. L. de. (2016). Crise e vulnerabilidade social: uma leitura territorial. Em *Geografia, Paisagens e Riscos* (pp. 189–214). Imprensa da Universidade de Coimbra. https://doi.org/10.14195/978-989-26-1233-1_9
- Marques, T. S., Soares, L., Teixeira, J., Abrantes, P., Sá, T., José, M., & Fernandes, A. R. (2017). *Perfis de vulnerabilidade social em Portugal*. 339–343. <https://repositorio-aberto.up.pt/handle/10216/111474>
- Mavhura, E., Manyena, B., & Collins, A. E. (2017). An approach for measuring social vulnerability in context: The case of flood hazards in Muzarabani district, Zimbabwe. *Geoforum*, 86, 103–117. <https://doi.org/10.1016/J.GEOFORUM.2017.09.008>
- Melidis, M., & Tzagkarakis, S. I. (2021). The evolution of social vulnerability in Greece during the economic crisis (2008-2017). <https://doi.org/10.1080/14616696.2021.2007973>, 24(2), 229–250. <https://doi.org/10.1080/14616696.2021.2007973>
- Mendes, J. M. (2009). Social vulnerability indexes as planning tools: beyond the preparedness paradigm. *Journal of Risk Research*, 12(1), 43–58. <https://doi.org/10.1080/13669870802447962>
- Mendes, J. M., Tavares, A., Cunha, L., & Freiria, S. (2011). A vulnerabilidade social aos perigos naturais e tecnológicos em Portugal. *Revista Crítica de Ciências Sociais*, 93(93), 95–128. <https://estudogeral.sib.uc.pt/handle/10316/20395>
- Mendes, J. M., & Tavares, A. (2011). Risco, vulnerabilidade social e cidadania. *Revista Crítica de Ciências Sociais*, 93, 05–08. <https://doi.org/10.4000/RCCS.173>
- Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffmann, A., & Giovannini, E. (2008). *Handbook on Constructing Composite Indicators: methodology and user guide* (C. Stevens, G. Baygan, K. Olsen, & S. Moore, Eds.). OECD, JRC. <https://www.oecd.org/els/soc/handbookonconstructingcompositeindicatorsmethodologyanduserguide.htm>
- National Research Council. (2001). American Hazardscapes: The Regionalization of Hazards and Disasters. Em S. L. Cutter (Ed.), *American Hazardscapes*. Joseph Henry Press. <https://doi.org/10.17226/10132>
- Ngo, E. B. (2001). When Disasters and Age Collide: Reviewing Vulnerability of the Elderly. *Natural Hazards Review*, 2(2), 80–89. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2001\)2:2\(80\)](https://doi.org/10.1061/(ASCE)1527-6988(2001)2:2(80))
- Norlén, H. (2018). Step 7: Statistical coherence (II) - Principal Component Analysis and Reliability Analysis. Em *16th JRC Annual Training on Composite Indicators & Scoreboards*. European Commission - Joint Research Centre. https://composite-indicators.jrc.ec.europa.eu/sites/default/files/COIN%202018%20Step%207%20Statistical%20coherence%20%28II%29_PCA.pdf

- Organization for Economic Cooperation and Development. (2019). *What is child vulnerability and how can it be overcome?* Acedido a 07 de setembro de 2022. <https://www.oecd-ilibrary.org/sites/23101e74-en/index.html?itemId=/content/component/23101e74-en>
- Paul, S. K. (2013). Vulnerability Concepts and its Application in Various Fields: A Review on Geographical Perspective. *Journal of Life and Earth Science*, 8, 63–81. <https://doi.org/10.3329/JLES.V8I0.20150>
- Pelling, M. (2003). *The vulnerability of cities: natural disasters and social resilience*. Earthscan Publications.
- PORDATA. (2022). *Base de Dados dos Municípios*. Acedido a 27 de dezembro de 2021. <https://www.pordata.pt/municipios>
- Pratt, M. K., & Bernstein, C. (2022). *What is Data Transformation? What Is Data Preparation? An in-Depth Guide to Data Prep*. Acedido a 21 de abril de 2022. <https://www.techtarget.com/searchdatamanagement/definition/data-transformation>
- Rabby, Y. W., Hossain, M. B., & Hasan, M. U. (2019). Social vulnerability in the coastal region of Bangladesh: An investigation of social vulnerability index and scalar change effects. *International Journal of Disaster Risk Reduction*, 41, 101329. <https://doi.org/10.1016/j.ijdr.2019.101329>
- Rajagopalan, G. (2020). Descriptive Data Analysis Basics. Em *A Python Data Analyst's Toolkit* (pp. 101–116). Apress. https://doi.org/10.1007/978-1-4842-6399-0_4
- Rakotomalala, R. (2013). *Bartlett's sphericity test and the KMO index (Kaiser-Mayer-Olkin)*. Acedido a 28 de março de 2022. https://eric.univ-lyon2.fr/~ricco/tanagra/fichiers/en_Tanagra_KMO_Bartlett.pdf
- Ramadas, S., Jha, A., Kumar, A., & Singh, S. (2018). Extent of vulnerability in wheat producing agro-ecologies of India: Tracking from indicators of cross-section and multi-dimension data. *Ecological Indicators*, 89, 771–780. <https://doi.org/10.1016/j.ecolind.2018.02.053>
- Regis, M. N. O. (2012). *Impacto do planejamento estratégico sobre o desempenho de bancos no Brasil* [Dissertação de Mestrado, Pontifícia Universidade Católica do Rio de Janeiro]. <https://doi.org/https://doi.org/10.17771/PUCRio.acad.20504>
- Rozalia Gabor, M., Ștefănescu, D., & Conțiu, L. C. (2010). The Application of Main Component Analysis Method on Indicators of Romanian National Authority for Consumers Protection Activities. *The AMFITEATRU ECONOMIC Journal*, 12(28), 314–331. https://www.amfiteatruconomic.ro/temp/Article_966.pdf
- Sabău Popa, C. D., Popa, D. N., Bogdan, V., & Simut, R. (2021). Composite financial performance index prediction – a neural networks approach. *Journal of Business Economics and Management*, 22(2), 277–296. <https://doi.org/10.3846/JBEM.2021.14000>

- Schmidtlein, M. C., Deutsch, R. C., Piegorsch, W. W., & Cutter, S. L. (2008). A Sensitivity Analysis of the Social Vulnerability Index. *Risk Analysis*, 28(4), 1099–1114. <https://doi.org/10.1111/J.1539-6924.2008.01072.X>
- Sharma, M., Burton, I., Aalst, M. van, Dilley, M., & Achrya, G. (2000). *Reducing Vulnerability to Environmental Variability: Background Paper for the Bank's Environment Strategy*. Acedido a 19 de julho de 2022. <http://web.worldbank.org/archive/website00346F/WEB/PDF/REDUCING.PDF>
- Sheytanova, T. (2015). *A Monte Carlo Study Comparing Three Methods for Determining the Number of Principal Components and Factors* [Dissertação de Mestrado, Universidade de Örebro]. <http://oru.diva-portal.org/smash/record.jsf?pid=diva2%3A896127&dswid=-6466>
- Sigorro, M. M. M. de G. (2018). *A definição de um índice de vulnerabilidade para a priorização de projetos e políticas públicas em Portugal Engenharia e Gestão Industrial* [Dissertação de Mestrado, Instituto Superior Técnico]. <https://fenix.tecnico.ulisboa.pt/cursos/megi/dissertacao/1972678479054395>
- Silva, D. F. B. F. (2021). *Pré-processamento de Dados e Comparação entre Algoritmos de Machine Learning para a Análise Preditiva de Falhas em Linhas de Produção para o Controlo* [Dissertação de Mestrado, Instituto Superior de Engenharia do Porto]. <https://recipp.ipp.pt/handle/10400.22/18266>
- Sivakumar, A., & Gunasundari, R. (2017). A Survey on Data Preprocessing Techniques for Bioinformatics and Web Usage Mining. *International Journal of Pure and Applied Mathematics*, 117(20), 785–794. <https://acadpubl.eu/jsi/2017-117-20-22/articles/20/68.pdf>
- Solangaarachchi, D., Griffin, A. L., & Doherty, M. D. (2012). Social vulnerability in the context of bushfire risk at the urban-bush interface in Sydney: a case study of the Blue Mountains and Ku-ring-gai local council areas. *Natural Hazards*, 64(2), 1873–1898. <https://doi.org/10.1007/s11069-012-0334-y>
- Spielman, S. E., Tuccillo, J., Folch, D. C., Schweikert, A., Davies, R., Wood, N., & Tate, E. (2020). Evaluating social vulnerability indicators: criteria and their application to the Social Vulnerability Index. *Natural Hazards*, 100(1), 417–436. <https://doi.org/10.1007/S11069-019-03820-Z/FIGURES/2>
- Tarling, H. A. (2017). *Comparative Analysis of Social Vulnerability Indices: CDC's SVI and SoVI®* [Dissertação de Mestrado, Lund University]. <https://lup.lub.lu.se/student-papers/search/publication/8928519>
- Tasnuva, A., Hossain, Md. R., Salam, R., Islam, A. R. Md. T., Patwary, M. M., & Ibrahim, S. M. (2020). Employing social vulnerability index to assess household social vulnerability of natural hazards: an evidence from southwest coastal Bangladesh. *Environment, Development and Sustainability*, 23(7), 10223–10245. <https://doi.org/10.1007/s10668-020-01054-9>

- University of South Carolina. (2016). *The SoVI Recipe*. Acedido a 22 de março de 2022. https://sc.edu/study/colleges_schools/artsandsciences/centers_and_institutes/hvri/documents/sovi/sovi_recipe_2016.pdf
- U.S. Department of Transportation. (2015). Acedido a 21 de julho de 2022. *Components of Vulnerability*. <https://toolkit.climate.gov/image/985>
- Vincent, K. (2004). Creating an index of social vulnerability to climate change for Africa. *Tyndall Center for Climate Change Research*, 56. https://www.researchgate.net/publication/228809913_Creating_an_Index_of_Social_Vulnerability_to_Climate_Change_in_Africa
- Vincke, P. (1999). Robust and neutral methods for aggregating preferences into an outranking relation. *European Journal of Operational Research*, 112(2), 405–412. [https://doi.org/10.1016/S0377-2217\(97\)00439-6](https://doi.org/10.1016/S0377-2217(97)00439-6)
- Watts, M. J., & Bohle, H. G. (1993). Hunger, famine and the space of vulnerability. *GeoJournal*, 30(2), 117–125. <https://doi.org/10.1007/BF00808128>
- Wisner, B., Blaikie, P., Cannon, T., & Davis, I. (1994). The Challenge of Disasters and Our Approach. Em *At Risk: natural hazards, people's vulnerability and disasters* (2.^a ed., pp. 3–48). Routledge.
- Wolf, S., Hinkel, J., Hallier, M., Bisaro, A., Lincke, D., Ionescu, C., & Klein, R. J. T. (2013). Clarifying vulnerability definitions and assessments using formalisation. *International Journal of Climate Change Strategies and Management*, 5(1), 54–70. <https://doi.org/10.1108/17568691311299363/FULL/PDF>
- Wood, N., Burton, C. G., & Cutter, S. L. (2010). Community variations in social vulnerability to Cascadia-related tsunamis in the U.S. Pacific Northwest. *Natural Hazards*, 52(2), 369–389. <https://doi.org/10.1007/S11069-009-9376-1/FIGURES/6>
- Zhao, A. (2021). *Data Pre-processing: Introduction and Integration*. Acedido a 22 de março de 2022. <https://ukdataservice.ac.uk/app/uploads/2021-01-14datapreprocessing1.pdf>
- Žurovec, O., Čadro, S., & Sitaula, B. (2017). Quantitative Assessment of Vulnerability to Climate Change in Rural Municipalities of Bosnia and Herzegovina. *Sustainability*, 9(7), 1208. <https://doi.org/10.3390/su9071208>

Anexos

Anexo 1 - Resultados dos Testes de Viabilidade dos Dados.

Ano	Teste de Esfericidade de Bartlett	Teste KMO
2010	0,000	0,841
2011	0,000	0,838
2012	0,000	0,831
2013	0,000	0,799
2014	0,000	0,809
2015	0,000	0,784
2016	0,000	0,791
2017	0,000	0,783
2018	0,000	0,769
2019	0,000	0,776

Fonte: Elaboração Própria.

Anexo 2 - VEA das 15 Componentes Principais.

CPs	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
CP1	48,30%	44,14%	42,52%	40,57%	39,81%	36,90%	36,47%	37,75%	36,23%	36,06%
CP2	61,73%	60,40%	58,05%	57,41%	56,21%	54,05%	53,53%	55,21%	54,68%	53,78%
CP3	70,15%	69,54%	68,33%	68,02%	67,24%	66,44%	66,43%	67,22%	67,56%	66,93%
CP4	76,26%	75,55%	74,43%	74,13%	73,83%	74,16%	73,62%	74,96%	75,48%	74,41%
CP5	80,82%	80,44%	80,05%	79,74%	79,51%	79,79%	79,48%	80,48%	81,38%	80,74%
CP6	84,67%	84,18%	84,45%	84,25%	84,12%	84,18%	83,97%	84,97%	86,05%	85,77%
CP7	88,09%	87,25%	87,74%	87,41%	87,31%	87,37%	87,35%	87,91%	88,75%	88,55%
CP8	90,75%	89,84%	90,35%	90,13%	89,92%	90,07%	90,22%	90,51%	91,23%	91,22%
CP9	93,02%	92,35%	92,72%	92,40%	92,14%	92,48%	92,69%	92,96%	93,44%	93,60%
CP10	94,87%	94,51%	94,87%	94,45%	94,27%	94,44%	94,62%	94,74%	95,16%	95,26%
CP11	96,53%	96,11%	96,37%	96,40%	96,12%	96,21%	96,24%	96,36%	96,71%	96,88%
CP12	97,99%	97,60%	97,79%	97,93%	97,63%	97,77%	97,81%	97,84%	98,09%	98,26%
CP13	99,00%	98,77%	98,89%	98,99%	98,89%	98,94%	98,85%	98,93%	99,13%	99,17%
CP14	99,83%	99,81%	99,83%	99,80%	99,82%	99,81%	99,79%	99,78%	99,81%	99,80%
CP15	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Fonte: Elaboração Própria.

Anexo 3 - Fórmulas de cálculo do IVS.

$$IVS_{2010} = 0,633CP_1 + 0,176CP_2 + 0,110CP_3 + 0,080CP_4$$

$$IVS_{2011} = 0,584CP_1 + 0,215CP_2 + 0,121CP_3 + 0,080CP_4$$

$$IVS_{2012} = 0,571CP_1 + 0,209CP_2 + 0,138CP_3 + 0,082CP_4$$

$$IVS_{2013} = 0,547CP_1 + 0,227CP_2 + 0,143CP_3 + 0,082CP_4$$

$$IVS_{2014} = 0,539CP_1 + 0,222CP_2 + 0,149CP_3 + 0,089CP_4$$

$$IVS_{2015} = 0,498CP_1 + 0,231CP_2 + 0,167CP_3 + 0,104CP_4$$

$$IVS_{2016} = 0,495CP_1 + 0,232CP_2 + 0,175CP_3 + 0,098CP_4$$

$$IVS_{2017} = 0,504CP_1 + 0,233CP_2 + 0,160CP_3 + 0,103CP_4$$

$$IVS_{2018} = 0,480CP_1 + 0,244CP_2 + 0,171CP_3 + 0,105CP_4$$

$$IVS_{2019} = 0,485CP_1 + 0,238CP_2 + 0,177CP_3 + 0,101CP_4$$