UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Ferramentas de apoio à decisão alicerçadas em SIG com integração de informação em saúde: Caso de estudo da Diabetes na ULSLA

Renata Vieira Ribeiro

Mestrado em Engenharia Biomédica e Biofísica

Dissertação orientada por:
Dra. Cristina Catita
Dr. Nuno Miguel de Pinto Lobo e Matela

2024

"Tudo está relacionado com tudo o resto, mas as coisas próximas estão mais relacionadas do que as distantes" (Tobler, 1970).

“Everything is related to everything else, but near things are more related than distant things.” (Tobler,1970)

Agradecimentos

Este trabalho não seria possível sem o apoio e orientação de várias pessoas e instituições, a quem gostaria de expressar a minha gratidão.

Em primeiro lugar, quero manifestar o meu sincero agradecimento aos meus orientadores, Dra. Cristina Catita e Dr. Nuno Miguel de Pinto Lobo e Matela, cuja orientação foi crucial para o desenvolvimento deste projeto. Um agradecimento especial à Dra. Cristina Catita pela sua orientação especializada em ArcGIS Pro. A sua experiência e disponibilidade foram determinantes para a realização das análises técnicas que sustentam este trabalho.

Expresso igualmente a minha gratidão ao Prof. Rui Santos, que, quando mais precisei de orientação, me forneceu as bases essenciais e me orientou com a sua sabedoria e clareza quando mais precisei de direção.

Agradeço também ao Prof. Rui Cortes, da empresa Lean Health, por me ter proporcionado a oportunidade de realizar este projeto e pela confiança que depositou em mim.

Agradeço também à Dra. Constança Lage pelo suporte contínuo e pelo encorajamento, especialmente nas fases mais exigentes e minuciosas deste projeto. Onde a sua ajuda foi fundamental para ultrapassar os momentos menos entusiasmantes.

Por fim, mas não menos importante, quero agradecer à minha família, especialmente à minha mãe, pelo apoio incondicional, compreensão e encorajamento ao longo de todo este percurso. O vosso suporte foi fundamental para levar a bom porto trabalho.

A todos, o meu sincero agradecimento.

Resumo

Este estudo analisa a aplicação de Sistemas de Informação Geográfica (SIG) para apoiar a tomada de decisões na Unidade Local de Saúde do Litoral Alentejano (ULSLA), com particular atenção à prevalência da Diabetes Mellitus Tipo 2 (DMT2). Através da análise de padrões espaciais e da identificação dos determinantes sociais de saúde (DSS), foi possível mapear a distribuição da DMT2 e identificar as áreas de maior risco no Litoral Alentejano, região em estudo. A metodologia adotada inclui a aplicação de modelos de regressão espacial e Regressão Geograficamente Ponderada (GWR), para captar as variações espaciais importantes e determinar a influência dos diferentes fatores socioeconómicos de acordo com a localização. Os resultados mostram que a dimensão espacial desempenha um papel crucial na compreensão da prevalência da DMT2, influenciada por fatores como o nível de educação e o estado civil, com variações espaciais notáveis entre as freguesias. A análise realizada sublinha a importância de abordagens localizadas e personalizadas na gestão da saúde pública, especialmente na prevenção e controlo da DMT2. Este projeto destaca a relevância de se continuar a investigar e aplicar técnicas de análise espacial para melhorar a compreensão e a resposta aos desafios de saúde pública na região, oferecendo também um modelo que pode ser adaptado a outras áreas de Portugal.

Palavras-chave: Sistemas de Informação Geográfica, Diabetes Mellitus Tipo 2, Regressão Geograficamente Ponderada, Determinantes Sociais de Saúde.

Abstract

This study analyses the application of Geographic Information Systems (GIS) to support decision-making in the Local Health Unit of the Alentejo Coast (ULSLA), with a particular focus on the prevalence of Type 2 Diabetes Mellitus (T2DM). Through the analysis of spatial patterns and the identification of social determinants of health (SDH), it was possible to map the distribution of T2DM and identify the areas of greatest risk in the region. The methodology adopted includes applying spatial regression models and Geographically Weighted Regression (GWR), to capture significant spatial variations and determine the influence of different socioeconomic factors according to location. The results indicate that the geographic distribution of T2DM is influenced by local factors, such as demographic and socioeconomic characteristics, which vary across parishes. The analysis highlights the importance of localized and customized approaches in public health management, particularly in preventing and controlling T2DM. This project emphasizes the relevance of continued research and the application of spatial analysis techniques to improve the understanding and response to the region's public health challenges, offering a model that can be adapted to other areas of Portugal.

Keywords: Geographic Information Systems, Type 2 Diabetes Mellitus, Geographically Weighted Regression, Social Determinants of Health.

Índice

Agradecimentos i
Resumo ii
Abstract iii
Índice de Figuras vi
Índice de Tabelas ix
Lista de Abreviaturas, Siglas e Símbolos x
1. Introdução 1
1.1. Problema 1
1.2. Motivação 2
1.3. Objetivos 3
1.4. Estrutura 3
2. Conhecimentos Teóricos 4
2.1. Geografia na Saúde 4
2.1.1. Definição e Conceitos de Saúde 4
2.1.2. Determinantes de Saúde 7
2.2. SIG 9
2.3. Diabetes 10
2.3.1. Diabetes e SIG 11
3. Materiais e Métodos 13
3.1. Metodologia 13
3.2. Caracterização da região de estudo 13
3.3. Dados e metadados 16
3.3.1. Pré-processamento dos Dados dos Utentes 18
3.3.2. Segmentação dos Utentes 19
3.3.3. Visualização de Dados 20
3.3.4. Autocorrelação Espacial Local 22
3.3.5. Modelação 23
3.3.6. Modelação preditiva espácio-temporal 25
4. Resultados 26
4.1. Contextualização 26
4.2. Pré-processamento dos MCDT's 27
4.2.1. Dados Brutos 27
4.2.2. Dados Processados 31
4.2.3. Segmentação do Perfil 33
4.3. Análise Espácio-temporal 35

4.3.1. Distribuição Geográfica dos Utentes 35
4.3.2. Análise do padrão de distribuição espacial de utentes com T90 36
4.3.3. Emerging Hot Spots 39
4.4. Análise de Autocorrelação Espacial 41
4.4.1. Variáveis Explicativas 41
4.4.2. Variável a Estimar 61
4.5. Modelação 62
4.5.1. Modelo de Regressão 63
4.5.2. Modelo *Spatial Lag* 65
4.5.3. Modelo GWR 66
4.5.4. Estimativa para 2026 70
5. Discussão 72
6. Conclusão 74
7. Referências. 75
ANEXOS 81
Anexo I- Tabelas Descritivas 81
Anexo II- Visualização de Dados 84

Índice de Figuras

Figura 2.1 *Esquema de integração vertical no Sistema Nacional de Saúde (SNS), mostrando o fluxo de atendimento através da Unidade Local de Saúde (ULS), com encaminhamento para Hospitais (HSP) ou Cuidados de Saúde Primários (CSP). 6*
Figura 2.2 *Pirâmides etárias por género ao longo nos anos 1960,1992 e 2024, respetivamente, destacando a distribuição percentual da população masculina e feminina em diferentes faixas etárias. 7*
Figura 2.3 *Determinantes sociais da saúde. Adaptado de Dahlgren & Whitehead, 1991 [32]. 8*
Figura 2.4 *Esquema que ilustra SIG (GIS acrónimo inglês) [33], transformam dados do mundo real em camadas de informação organizadas, facilitando a visualização e análise espacial para melhor compreensão dos dados. 9*
Figura 3.1 *Metodologia utilizada ao longo do projeto. 13*
Figura 3.2 *À esquerda, a posição geográfica do Alentejo Litoral em Portugal continental é apresentada em destaque [54], enquanto o mapa à direita detalha a subdivisão dos concelhos na sub-região [55]. 14*
Figura 3.3 *Mapa com a distribuição dos Centros de Saúde em Portugal (2023). 15*
Figura 3.4 *Mapa com a distribuição dos Hospitais em Portugal (2023). 15*
Figura 3.5 *Workflow da metodologia aplicada, utilizando as ferramentas Python, ArcGIS Pro® e GeoDa. . 16*
Figura 3.6 *Cubo espácio-temporal que ilustra a ferramenta "Create Space Time Cube by Aggregating Points", imagem retirada da ESRI [62]. 21*
Figura 3.7 *Análise de Hot Spots Emergentes, utilizando o cubo espácio-temporal para identificar padrões de mudança ao longo do tempo, imagem retirada da ESRI [63]. 22*
Figura 4.1 *Análise da Densidade de Kernel dos Centros de Saúde em Portugal (2023). A cor vermelha corresponde a zonas com menor densidade de centros de saúde. 26*
Figura 4.2 *Análise da Densidade de Kernel dos Hospitais em Portugal (2023). A cor vermelha corresponde a zonas com menor densidade de hospitais. 27*
Figura 4.3 *Distribuição das Variáveis Clínicas, antes do processamento, por USF em 2017. A distribuição dos dados é ilustrada através de boxplots, onde cada cor representa uma USF específica, detalhada na legenda à direita. 29*
Figura 4.4 *Distribuição das Variáveis Clínicas, antes do processamento, por USF em 2023. A distribuição dos dados é ilustrada através de boxplots, onde cada cor representa uma USF específica, detalhada na legenda à direita. 30*
Figura 4.5 *Distribuição das Variáveis Clínicas Após Limpeza de Dados em 2023. Esta matriz de histogramas ilustra a distribuição de várias variáveis clínicas coletadas após o processo de limpeza de dados. 32*
Figura 4.6 *Comparação dos Indicadores Médios por Clusters na ULSLA em 2023. As barras representam o IMC, HbA1c, e o perímetro abdominal, enquanto a linha azul mostra a idade média dos pacientes em cada cluster. 34*
Figura 4.7 *Distribuição geográfica de utentes com T90 nos anos 2017, 2020 e 2023. A distribuição para os anos intermédios está disponível no Anexo II. 36*
Figura 4.8 *Densidade de Kernel do número de Pacientes com T90 nos Anos de 2017, 2020 e 2023. A cor vermelha indica maior densidade de utentes com DMT2. .....37*
Figura 4.9 *Distribuição Direcional de Pacientes com T90 em LA em 2017, 2020 e 2023. As áreas dentro das elipses indicam regiões de interesse onde há maior concentração de casos. 38*
Figura 4.10 *Relatório ANN para Pacientes com T90 de 2017, obtido através do ArcGIS Pro. 38*
Figura 4.11 *Análise da Função K para Distribuição Espacial. Os gráficos dos anos 2017-2022 encontram-se no Anexo II. 39*
Figura 4.12 *Distribuição Espácio-temporal dos Hot Spots de T90 do LA (2017-2023). O mapa ilustra a localização dos hot spots persistentes, emergentes e esporádicos de T90 ao longo do período de sete anos. 40*

Figura 4.13 *Análise Espacial da População Total por Freguesia e Autocorrelação Espacial em 2023. Da esquerda para a direita, o mapa mostra a distribuição da população por freguesia, indica a significância espacial de clusters e o índice de Moran's I.* *42*
Figura 4.14 *Distribuição Percentual da População por Sexo por Freguesia em 2023. Os mapas ilustram a população feminina (à esquerda) e masculina (à direita).* *43*
Figura 4.15 *Análise Espacial de Significância para População Feminina e Masculina em 2023. O mapa à esquerda detalha a população feminina e o mapa à direita a masculina.* *43*
Figura 4.16 *Análise de Moran's I para Autocorrelação Espacial da População Feminina e Masculina. Os gráficos mostram o índice de Moran's I para a autocorrelação espacial das populações feminina (à esquerda) e masculina (à direita).* *44*
Figura 4.17 *Distribuição Percentual da População por Faixas Etárias de 2023, da esquerda para a direita e de cima para baixo:0-14, 15-29, 30-44, 45-59 e 60+ anos.* *45*
Figura 4.18 *Análise de Autocorrelação Espacial por Faixas Etárias em 2023, da esquerda para a direita, os mapas representam as faixas etárias de 0-14, 15-29, 30-44, 45-59 e 60+ anos.* *46*
Figura 4.19 *Gráficos de Moran's I para Autocorrelação Espacial por Faixas Etárias em 2023, da esquerda para a direita e de cima para baixo: 0-14, 15-29, 30-44, 45-59 e 60+ anos.* *46*
Figura 4.20 *Distribuição das Habilitações Literárias por Freguesia em 2023. Os mapas, da esquerda para a direita, apresentam a percentagem da população que completou o 1º ciclo, 2º ciclo e 3º ciclo.* *47*
Figura 4.21 *Autocorrelação Espacial das Habilitações Literárias por Freguesia em 2023. Da esquerda para a direita os mapas representam o 1º ciclo (PopEdu1), 2º ciclo (PopEdu2), e 3º ciclo (PopEdu3).* *48*
Figura 4.22 *Gráficos de Moran's I para Autocorrelação Espacial das Habilitações Literárias em 2023. Os gráficos representam o índice de Moran's I, da esquerda para a direita, o 1º ciclo (PopEdu1), 2º ciclo (PopEdu2), e 3º ciclo (PopEdu3).* *49*
Figura 4.23 *Distribuição das Habilitações Literárias por Freguesia em 2023. Os mapas apresentam a percentagem da população que concluiu o ensino secundário, pós-secundário e graus académicos avançados (licenciatura, mestrado e doutorado), da esquerda para a direita.* *49*
Figura 4.24 *Autocorrelação Espacial das Habilitações Literárias em 2023. Os mapas ilustram a análise de autocorrelação espacial para o ensino secundário (PopEdu4), educação pós-secundária (PopEdu5) e educação superior (PopEdu6), da esquerda para a direita.* *50*
Figura 4.25 *Análise de Moran's I das Habilitações Literárias por Freguesia em 2023. Os gráficos representam o índice de Moran's I, da esquerda para a direita, o ensino secundário (PopEdu4), educação pós-secundária (PopEdu5) e educação superior (PopEdu6).* *51*
Figura 4.26 *Distribuição de Rendimentos em 2023. Os mapas apresentam a percentagem de agregados familiares em diversas faixas de rendimentos.* *52*
Figura 4.27 *Autocorrelação Espacial das Faixas de Rendimentos em 2023.* *53*
Figura 4.28 *Gráficos de Moran's I para Autocorrelação Espacial das Faixas de Rendimentos em 2023.* *54*
Figura 4.29 *Distribuição Per Capita de Despesas com Álcool e Autocorrelação Espacial por Freguesia em 2023. O mapa detalha a despesa per capita com álcool por freguesia, mostra os padrões espaciais de autocorrelação e apresenta o índice de Moran's I, da esquerda para a direita.* *55*
Figura 4.30 *Distribuição Per Capita de Despesas com Tabaco e Autocorrelação Espacial por Freguesia em 2023.* *55*
Figura 4.31 *Distribuição Percentual do Desemprego e Autocorrelação Espacial por Freguesia em 2022. Da esquerda para a direita, o mapa detalha a percentagem de desempregados por freguesia, ilustra padrões espaciais e o índice de Moran's I.* *57*
Figura 4.32 *Distribuição de Estado Civil por Freguesia em 2023. Da esquerda para a direita, os mapas mostram a percentagem de solteiros, de casados e de divorciados.* *58*
Figura 4.33 *Autocorrelação Espacial dos Estados Civis em Freguesias Selecionadas em 2023. Os mapas analisam a autocorrelação espacial de estados civis: solteiros, casados e divorciados, da esquerda para a direita.* *59*

Figura 4.34 *Análise de Moran's I para Autocorrelação Espacial de Estados Civis em 2023. Os três gráficos mostram o índice de Moran's I com base no estado civil: solteiro, casado e divorciado, da esquerda para a direita.* *59*
Figura 4.35 *Distribuição do Poder de Compra Per Capita por Freguesia e Autocorrelação Espacial em 2023. Da esquerda para a direita, o poder de compra per capita por freguesia, os padrões de autocorrelação e o gráfico do índice de Moran's I.* *60*
Figura 4.36 *Distribuição de Pacientes com Diabetes (primeiro mapa à esquerda) e Prevalência (segundo mapa) por Freguesia em 2023. O mapa à direita e o gráfico de Moran's I analisam a autocorrelação espacial.* *61*
Figura 4.37 *Matriz de Gráficos de Dispersão, que analisa as relações entre as variáveis independentes e a variável dependente.* *63*
Figura 4.38 *Distribuição espacial dos Resíduos do Modelo OLS para a prevalência de diabetes em 2023.* *65*
Figura 4.39 *Distribuição espacial dos Resíduos do Modelo Spatial Lag para a prevalência de diabetes em 2023.* *66*
Figura 4.40 *Distribuição Espacial dos Coeficientes GWR para as variáveis PER1EDU (à esquerda) e PERDIVORC (à direita).* *68*
Figura 4.41 *Resíduos do Modelo GWR para Diabetes em 2023 por Freguesia.* *69*
Figura 4.42 *Distribuição Espacial dos Condition Number do Modelo GWR.* *70*
Figura 4.43 *Estimativa de novos casos de Diabetes para 2026, através do Modelo de Séries Temporais.* *71*

Índice de Tabelas

Tabela 3.1 *Descrição dos Conjuntos de Dados Utilizados no Estudo* *17*
Tabela 3.2 *Dados Clínicos da População Diabética, incluindo Dados Demográficos, Geográficos, Estilo de Vida, Comorbilidades e Meios Complementares de Diagnóstico e Terapêutica (MCDT's).* *18*
Tabela 4.1 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2017. Contêm o número de observações (count), média (mean), desvio padrão (std), valor mínimo (min), percentis 25%, 50% (mediana), 75% e valor máximo (max) para cada variável clínica.* *31*
Tabela 4.2 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2023. Contêm o número de observações (count), média (mean), desvio padrão (std), valor mínimo (min), percentis 25%, 50% (mediana), 75% e valor máximo (max) para cada variável clínica.* *31*
Tabela 4.3 *Características demográficas e de IMC da população diabética em 2023.* *33*
Tabela 4.4 *Características clínicas da população diabética em 2023.* *33*
Tabela 4.5 *Análise dos valores de NNR, z-score e p-value para Pacientes com T90 de 2017 a 2023.* *39*
Tabela 4.6 *Resumo da análise de autocorrelação espacial para diferentes variáveis no LA e respetivos índices de Moran's I.* *62*
Tabela 4.7 *Resultados do modelo OLS. Apresenta os coeficientes estimados, com seus erros-padrão, estatísticas-t, probabilidades associadas para cada variável e o VIF.* *64*

Tabela Anexo I. 1 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2017.* *81*
Tabela Anexo I. 2 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2018.* *81*
Tabela Anexo I. 3 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2019.* *81*
Tabela Anexo I. 4 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2020.* *81*
Tabela Anexo I. 5 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2021.* *81*
Tabela Anexo I. 6 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2022.* *82*
Tabela Anexo I. 7 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2023.* *82*
Tabela Anexo I. 8 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2017.* *82*
Tabela Anexo I. 9 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2018.* *82*
Tabela Anexo I. 10 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2019.* *82*
Tabela Anexo I. 11 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2020.* *83*
Tabela Anexo I. 12 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2021.* *83*
Tabela Anexo I. 13 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2022.* *83*
Tabela Anexo I. 14 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2023.* *83*

Lista de Abreviaturas, Siglas e Símbolos

AICc	Critério de Informação de Akaike corrigido
ANN	*Average Nearest Neighbor*
API	Interface de Programação de Aplicações
BGRI	Base Geográfica de Referenciação de Informação
CSP	Cuidados de Saúde Primários
DMT2	Diabetes Mellitus Tipo 2
DSS	Determinantes Sociais de Saúde
GWR	Regressão Geograficamente Ponderada
HbA1c	Hemoglobina Glicada
HDL	Lipoproteínas de Alta Densidade
HSP	Hospitais
ICPC-2	Classificação Internacional de Assistência Primária
IMC	Índice de Massa Corporal
INE	Instituto Nacional de Estatística
INSEF	Inquérito Nacional de Saúde com Exame Físico
LDL	Lipoproteínas de Baixa Densidade
MCDT	Meios Complementares de Diagnóstico e Terapêutica
MIM@UF	Módulo de Informação e Monitorização das Unidades Funcionais
NNR	Nearest Neighbor Ratio
OLS	Método dos Mínimos Quadrados Ordinários
OMS	Organização Mundial da Saúde
Perímetro AB	Perímetro Abdominal
SIG	Sistemas de Informação Geográfica
SNS	Serviço Nacional de Saúde
T90	Diabetes não insulino-dependente
TAd	Tensão Arterial diastólica

TAs	Tensão Arterial sistólica
UFS	Unidades Funcionais de Saúde
ULS	Unidades Locais de Saúde
ULSLA	Unidade Local de Saúde do Litoral Alentejano

1. Introdução

1.1. Problema

Atualmente, o mundo enfrenta ameaças crescentes à saúde pública através de desigualdades persistentes, alterações climáticas, aumento da esperança de vida e escassez de recursos. São tempos difíceis, estes que atravessamos, e o setor da saúde está e estará cada vez mais sob pressão, enfrentando desafios exigentes e crescentes, devido a uma procura progressiva dos cuidados de saúde, num cenário de recursos limitados [1].

Em Portugal, entre 2012 e 2022, a esperança média de vida aumentou para ambos os sexos. Para os homens fixou-se em 78.8, e para as mulheres em 84.5, em 2019. Sendo que devido à pandemia de Covid, este valor apresentou um ligeiro decréscimo, retrocedendo para os 84.2, para logo em 2021 [2] recuperar e continuar a subir no entanto a quantidade não está associada à qualidade.

O aumento de doenças crónicas, agravada pelo envelhecimento populacional [3], exige uma atenção centrada em múltiplos parâmetros, que vá muito para além da mera extensão dos anos de vida, com as ULSs a desempenharem um papel estratégico na coordenação e prevenção eficaz de cuidados. Em Portugal, o acesso à saúde é um desafio. Persistem disparidades regionais, com áreas rurais enfrentando maiores dificuldades de acesso a cuidados e condições de saúde desfavoráveis [4]. A crescente competição por profissionais entre o setor público e privado, aliada à insatisfação e escassez, afeta o acesso e a sustentabilidade do serviço público.

Em Portugal, persistem disparidades regionais na saúde, contrastando com as áreas urbanas, nas regiões rurais perduram as condições precárias. Estas áreas enfrentam limitações no acesso a cuidados de saúde, escassez de profissionais médicos, determinantes de saúde menos favoráveis e riscos ambientais. Para atenuar estas desigualdades, é necessário melhorar a acessibilidade aos serviços de saúde, incentivar a fixação de profissionais no interior e investir em infraestruturas e programas de saúde pública adaptados a estas comunidades [3].

A necessidade de melhorar a prestação de serviços de saúde, a fim de otimizar a alocação de recursos, tornou-se imperativa. A identificação e compreensão dessas variações é decisiva para desenvolver estratégias direcionadas que possam mitigar as disparidades em saúde e promover uma maior equidade no sistema de saúde português.

A necessidade de assegurar sistemas de saúde eficazes e viáveis é enfatizada pela Organização Mundial de Saúde (OMS), que destaca que apenas 30% dos resultados em saúde derivam diretamente da assistência médica, enquanto a grande maioria, 70%, é influenciada por determinantes sociais e ambientais [5].

Em Portugal, as ULS são vitais na prestação de cuidados de saúde primária, atuando como pilares fundamentais do Sistema Nacional de Saúde (SNS) e servindo diretamente as comunidades locais. As ULS têm a capacidade de atuar na intersecção entre a assistência médica e os determinantes sociais da saúde, implementando estratégias que abrangem desde a prevenção de doenças até à intervenção em condições socioeconómicas que afetam a saúde das comunidades.

Assim, para reforçar o impacto positivo das ULS no SNS e na saúde da população, é essencial que estas unidades estejam equipadas com ferramentas que lhes permitam não só identificar mas também agir proactivamente diante das necessidades específicas das comunidades que servem. A adoção de algoritmos de decisão avançados pode ser um passo crucial para alcançar uma gestão de recursos mais inteligente e uma resposta mais ágil às dinâmicas de saúde em constante mudança.

Além disso, o compromisso com a melhoria contínua dos serviços prestados pelas ULS, alinhado com as diretrizes da OMS, pode impulsionar Portugal em direção a um sistema de saúde mais resiliente e adaptável. A integração de dados e tecnologias analíticas avançadas permitirá a personalização dos cuidados de saúde, garantindo que cada indivíduo receba a assistência mais eficaz, ao mesmo tempo em que se promove a equidade em todo o sistema.

Este estudo propõe o desenvolvimento de ferramentas de apoio à decisão para a Unidade Local de Saúde do Litoral Alentejano (ULSLA), com o objetivo de otimizar o planeamento e a gestão dos recursos de saúde. Com uma metodologia baseada na análise de dados, incluindo a espacialização de informação e o uso de plataformas digitais, pretende-se elevar a eficiência nos cuidados prestados, promovendo a equidade no acesso à saúde. Este esforço de pesquisa aponta para a melhoria contínua da qualidade assistencial na ULSLA, através de uma gestão de recursos mais eficaz e da identificação precisa de necessidades comunitárias, contribuindo para a redução das desigualdades geográficas em saúde.

1.2. Motivação

Os Sistemas de Informação Geográfica (SIG) são essenciais para a observação, análise, interpretação e gestão de uma vasta gama de dados. Indicadores sociais, económicos e ambientais estão sempre ligados ao território, e o seu uso para o planeamento, monitorização e avaliação do estado de um município, região ou país é uma vantagem. Transportá-la para a área da saúde foi o estímulo deste estudo, dando um contributo positivo para facilitar a gestão e alocação de recursos médicos, melhorando o acesso e a eficiência dos cuidados de saúde.

Ao analisar exemplos internacionais, constata-se que países como o Canadá, Austrália e o Reino Unido, que já incorporaram tecnologias de SIG nas suas estratégias de saúde pública, conseguiram atenuar desigualdades geográficas e otimizar a distribuição de serviços de saúde [6]. Esta investigação procura explorar como tais práticas podem ser adaptadas ao contexto português, tendo em conta as especificidades regionais e os desafios únicos enfrentados pelo SNS.

Através da análise comparativa com países que já observam resultados positivos decorrentes da aplicação de SIG em saúde, pretende-se identificar estratégias eficazes em Portugal, com o intuito de promover uma distribuição mais justa de recursos de saúde e melhorar o acesso aos cuidados médicos, especialmente em zonas rurais e comunidades menos servidas, como o Litoral Alentejano.

A escolha do Litoral Alentejano como foco deste estudo justifica-se pela sua configuração geográfica e demográfica singular, que representa desafios particulares na prestação de cuidados de saúde. É um território com 5 308 Km² e um total de 96 442 habitantes (Censos 2021). A região caracteriza-se por áreas rurais extensas que acentuam a dispersão dos habitantes e dificultam o acesso aos serviços de saúde, com uma heterogeneidade populacional devido à migração de pessoas para trabalhar na agricultura. Estes fatores contribuem para disparidades no acesso aos cuidados médicos, tornam o Litoral Alentejano uma zona de interesse para investigação e estudo de como o SIG pode ajudar a superar barreiras geográficas e melhorar a distribuição de recursos de saúde, respondendo assim às necessidades tão variadas quanto a diversidade da sua população.

Neste âmbito, propõe-se aprofundar a análise para identificar os fatores que têm maior impacto na população estudada em relação a cada doença específica. Utilizando a capacidade analítica dos SIG, esta abordagem permite não só mapear a incidência de doenças e a distribuição de recursos de saúde mas também correlacionar esses dados com variáveis socioeconómicas, ambientais e de acesso aos cuidados médicos.

Com isso, espera-se obter perspetivas detalhadas sobre os determinantes de saúde que mais influenciam a prevalência de doenças no Litoral Alentejano, possibilitando a formulação de intervenções

mais precisas e direcionadas. Esta análise detalhada contribuirá para uma compreensão mais profunda dos desafios de saúde enfrentados pela população da região, facilitando a implementação de estratégias de saúde pública mais eficazes e equitativas, especialmente projetadas para mitigar as disparidades identificadas e melhorar o bem-estar geral dos residentes.

1.3. Objetivos

O objetivo geral deste estudo é implementar ferramentas de apoio à decisão baseados em dados espacializados para a ULSLA. Esta implementação procura melhorar a decisão do planeamento e gestão na alocação de recursos de saúde, otimizar a eficiência na prestação de cuidados, e consequentemente, atenuar as desigualdades geográficas entre os utentes, com as seguintes etapas:

- Analisar o padrão espacial de utentes e prevalência da diabetes (variável dependente) nas freguesias geridas pela ULSLA.
- Identificar os principais indicadores socioeconómicos que influenciam a variável dependente.
- Identificar os padrões geográficos para cada um dos indicadores encontrados.
- Estimar um modelo de regressão linear com um bom ajuste e uma boa capacidade preditiva.
- Estimar um modelo de regressão espacial local para investigar a influência de cada indicador socioeconómico em cada freguesia.
- Estimar a incidência da diabetes na região.

Com esta análise, pretende-se dar resposta às seguintes questões de investigação:

1. Quais são os fatores socioeconómicos que mais afetam as taxas de prevalência da diabetes?
2. Quais são as regiões em que cada fator exerce maior influência?

Para atingir o propósito deste estudo, é essencial cumprir uma série de objetivos específicos, usando um processo lógico e progressivo.

1.4. Estrutura

Este trabalho divide-se em 6 capítulos. No primeiro capítulo, apresenta-se o problema de investigação relacionado com a prevalência da DMT2 no Litoral Alentejano e definem-se os objetivos do estudo. A seguir vem um capítulo dedicado à revisão teórica, onde são explorados os conceitos chave, como a relação entre geografia e saúde e a aplicação dos Sistemas de Informação Geográfica (SIG) na área da saúde.

O capítulo de materiais e métodos detalha as técnicas utilizadas para a recolha e análise dos dados. No quarto capítulo, são apresentados os resultados das metodologias aplicadas. A discussão crítica dos resultados e as suas implicações para a formulação de políticas de saúde pública são abordadas no capítulo seguinte.

Finalmente, a tese conclui com um resumo dos principais resultados e a apresentação de recomendações práticas para futuras políticas de saúde, bem como sugestões para investigações futuras.

2. Conhecimentos Teóricos

2.1. Geografia na Saúde

É inegável que o estado de saúde da população e a geografia de uma região se influenciam mutuamente, dado que o território onde vivemos afeta diretamente o nosso bem-estar [7]. O ambiente envolvente que inclui fatores modificáveis, ambientais, comportamentais e sociais (relações comunitárias e condições de vida) [8], desempenha um papel crucial na determinação da nossa saúde.

Um dos marcos iniciais da ligação entre Geografia e Saúde foi o estudo do Dr. John Snow em 1850. Ao mapear as mortes causadas por um surto de cólera em Londres, Snow identificou a fonte de contaminação nos poços de água, ajudando a conter a propagação da doença e demonstrando a transmissão hídrica da cólera. Esta ação pioneira abriu caminho para que outros combinassem geografia e saúde para esclarecer problemas até então desconhecidos [9], [10].

Em Portugal, a Geografia na Saúde ainda se encontra numa fase inicial, tanto na investigação como no ensino [11]. No entanto, já existem diversos estudos que exploram a relação entre o território e os determinantes de saúde [12], [13], [14].

2.1.1. Definição e Conceitos de Saúde

A saúde é um direito fundamental, valorizado por todos e estudado por diversas áreas do conhecimento, o que resulta em múltiplas definições. Atualmente, uma das definições amplamente aceite é a da Organização Mundial da Saúde (OMS), que descreve a saúde como um estado completo de bem-estar físico, mental e social, e não apenas como ausência de doença, conforme estabelecido no Preâmbulo da sua constituição em 1946, com o objetivo de criar uma definição universal e clara [15].

No entanto, esta visão mais abrangente nem sempre foi a norma. Inicialmente, saúde era definida simplesmente como "ausência de doença" [16]. Com o tempo, essa definição evoluiu para incluir outras dimensões. Atualmente, a saúde é compreendida não apenas como um conceito objetivo ou um estado biológico, mas como um modelo mais complexo que incorpora componentes psíquicas, sociais, qualidade e estilos de vida da população [11] onde o local e o ambiente desempenham um papel fundamental.

Desde a década de 90, a integração de cuidados em saúde tem sido um conceito em evolução. Inicialmente focado na eficiência organizacional, jurídica e financeira, [3] e posteriormente orientado para a melhoria dos resultados em saúde, redução da fragmentação e aumento da eficiência geral do sistema. A diversidade de definições para integração de cuidados, apesar de não haver consenso [4], [17], [18], destaca a complexidade e a natureza multifacetada deste conceito, que procura proporcionar um contínuo de serviços de saúde ajustados às necessidades dos pacientes a diferentes níveis [19].

Aplicado ao sector da saúde, o termo integração vertical é utilizado para designar um conjunto relativamente amplo de movimentos e mudanças [20]. A integração vertical consiste na criação de uma única entidade gestora (propriedade e controlo) de duas ou mais entidades que prestam serviços em diferentes níveis de cuidados, com o intuito de melhorar o estado de saúde geral num determinado contexto geodemográfico [20], [21].

O Serviço Nacional de Saúde (SNS) português é um sistema de saúde universal, predominantemente financiado por impostos, que abrange todos os residentes, incluindo os requerentes de asilo e migrantes que se encontram em processo de regularização do seu estatuto [22].

O SNS é definido, segundo o Decreto-Lei n.º 11/93, como um "conjunto ordenado e hierarquizado de instituições e serviços oficiais prestadores de cuidados de saúde, funcionando sob a tutela do Ministério da Saúde". Segundo o mesmo Decreto-Lei, o SNS, tem como objetivo a "efetivação, por parte do estado, da responsabilidade que lhe cabe na proteção da saúde individual e coletiva". A integração vertical constitui um modelo operacional no qual uma única entidade assume a responsabilidade por todos os elementos do contínuo de cuidados [23]. Este modelo abrange os vários níveis de assistência à saúde, incluindo os cuidados primários, hospitalares e continuados [4], [24]. Segundo a definição da Organização Mundial de Saúde (OMS) [4], a integração vertical compreende a agregação de recursos, bem como a prestação e gestão de serviços que vão desde a prevenção e promoção da saúde até ao diagnóstico, tratamento e reabilitação. Esta abordagem abrangente está diretamente associada à procura de serviços de saúde caracterizados pela sua qualidade, eficiência, acessibilidade e capacidade de satisfazer as necessidades dos utentes.

No Decreto-Lei nº 11/93 de 15 de janeiro já se salientava que a dicotomia entre cuidados de saúde primários e cuidados hospitalares não promovia a eficiência da prestação, sendo necessária a criação de unidades de saúde integradas que se articulassem entre si, com vista a uma gestão de recursos mais eficiente. Pode ler-se: "A tradicional dicotomia entre cuidados primários e cuidados diferenciados revelou-se não só incorreta do ponto de vista médico, mas também causadora de disfunções sob o ponto de vista organizativo. Daí a criação de unidades integradas de cuidados de saúde - unidades de saúde - que hão-de viabilizar a imprescindível articulação entre grupos personalizados de centros de saúde e hospitais. A indivisibilidade da saúde, por um lado, e a criteriosa gestão de recursos, por outro, impõem a consagração de tal modelo, em que radica um dos aspetos essenciais da nova orgânica do Serviço Nacional de Saúde".

As Unidades Locais de Saúde (ULS) surgem como uma resposta à necessidade de integrar de forma mais eficaz os cuidados de saúde prestados pelos centros de saúde e hospitais, reunindo sob uma única entidade os recursos humanos, financeiros e materiais. Criadas no âmbito da reforma do sistema de saúde português, as ULS visam melhorar a eficiência, coesão e qualidade dos cuidados de saúde, facilitando o acesso da população aos serviços e a circulação entre diferentes níveis de cuidados, desde os primários aos secundários e continuados. Além disso, desempenham um papel decisivo na promoção da saúde e na prevenção da doença, adaptando-se continuamente às mudanças demográficas, económicas e sociais, bem como aos novos desafios de saúde pública.

Os sistemas integrados de saúde, centrados nas necessidades reais dos utentes, promovem uma cultura de segurança do paciente, reduzem riscos clínicos e melhoram a qualidade dos cuidados através da colaboração multidisciplinar [23] e do acesso centralizado a dados clínicos [24].

No caso do doente diabético, estudos demonstraram que a continuidade de cuidados está relacionada com a melhoria do bem-estar a vários níveis: físico, social, mental e perceção da dor [25]. O melhor controlo glicémico também está relacionado com a continuidade de cuidados [26] bem como a redução de complicações [27].

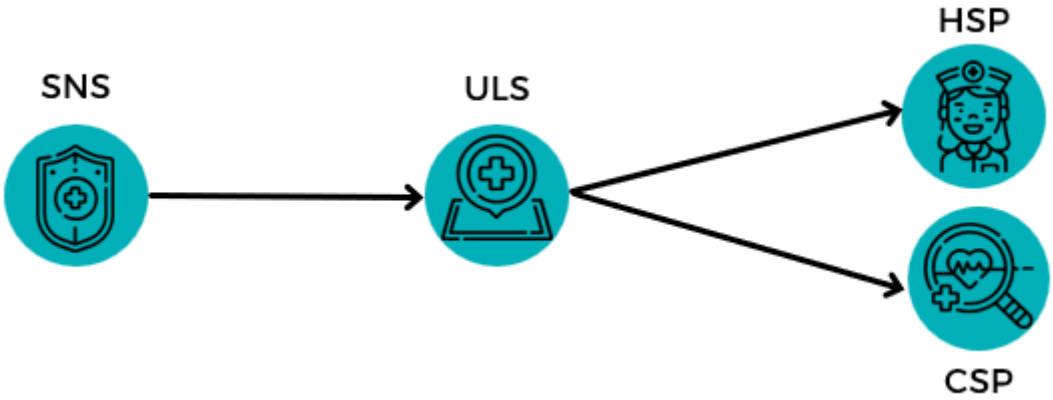


Figura 2.1 *Esquema de integração vertical no Sistema Nacional de Saúde (SNS), mostrando o fluxo de atendimento através da Unidade Local de Saúde (ULS), com encaminhamento para Hospitais (HSP) ou Cuidados de Saúde Primários (CSP).*

A integração dos cuidados prestados pelos hospitais e centros de saúde nas ULS contribui para a eficácia do percurso dos utentes dentro do sistema de saúde, como está representado na Figura 2.1. Esta abordagem facilita a continuidade da assistência, promovendo a saúde da população de forma holística, ao longo do ciclo de vida – desde a prevenção da doença e promoção da saúde até ao diagnóstico, tratamento e reabilitação. As ULS representam um importante passo em direção a uma atenção na saúde mais eficaz, coordenada e centrada no paciente, onde a promoção da saúde e a prevenção de doenças desempenham um papel central.

As ULS em Portugal exemplificam a integração vertical no sector da saúde, promovendo uma gestão unificada dos cuidados a diferentes níveis, com o objetivo de melhorar a saúde da população. Esta abordagem facilita a interligação entre diferentes níveis de prestação de cuidados de saúde, reconhecida pela Comissão Europeia como vital para a sustentabilidade do sistema de proteção social na União Europeia [28].

Os cuidados de saúde primários desempenham um papel fundamental no sistema de saúde, sendo o pilar fundamental da prestação eficaz de cuidados de saúde. Esta abordagem visa fornecer serviços integrados, acessíveis e contínuos, que abrangem a grande maioria das necessidades em saúde de indivíduos e comunidades.

Estes cuidados de saúde envolvem diversas dimensões que se estendem para além do simples tratamento de doenças, incluindo acesso, continuidade, coordenação, abrangência, qualidade, eficiência dos cuidados e equidade na saúde. Cada uma destas dimensões desempenha um papel determinante na promoção de um desempenho cada vez melhor do sistema nacional de saúde e na saúde em geral. Promovendo a prevenção de doenças, estes cuidados garantem a continuidade da assistência ao longo da vida, desempenhando um papel fundamental na deteção precoce de problemas de saúde, na gestão de doenças crónicas e na redução da necessidade de cuidados de saúde secundários e terciários, que são muito dispendiosos.

Como foi mencionado anteriormente, o setor da saúde enfrenta desafios difíceis e crescentes, em que a procura de cuidados de saúde continua a aumentar num cenário de recursos limitados [1].

A esperança média de vida, em Portugal, tem vindo a aumentar gradualmente, o que acarreta mudanças nas necessidades da população. Na Figura 2.2, observa-se a distribuição da população por grupos etários e por sexo para os anos de 1960, 1992 e 2024 [29], representando a evolução da população ao longo deste período. Em 64 anos, Portugal passou de um país 'jovem' a um país envelhecido. Esta análise revela que a natalidade e a mortalidade têm vindo a abrandar e ou estabilizar. A baixa natalidade

pode ser atribuída à generalização do planeamento familiar, mas principalmente a degradação das condições económicas e financeiras das famílias, enquanto a redução da mortalidade é resultado do aumento dos cuidados médicos e da enorme evolução na medicina.

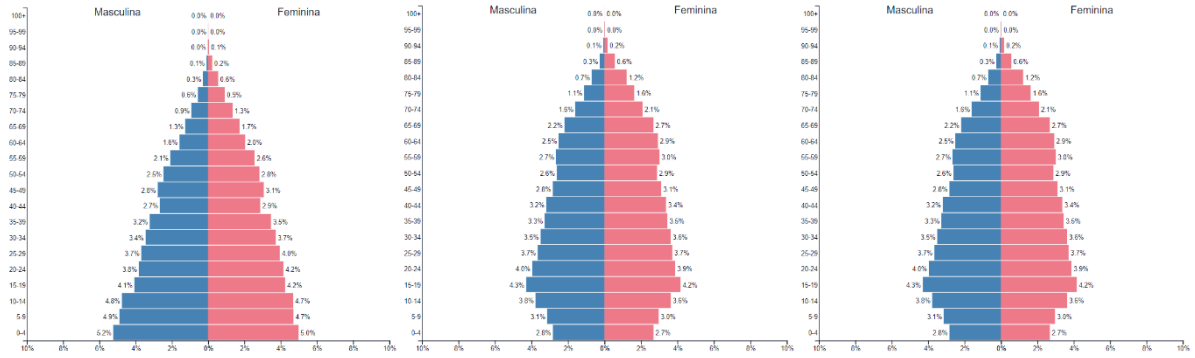


Figura 2.2 *Pirâmides etárias por género ao longo nos anos 1960,1992 e 2024, respetivamente, destacando a distribuição percentual da população masculina e feminina em diferentes faixas etárias.*

Enfrentamos o desafio do acesso, com as famílias portuguesas a financiarem 29% da despesa em saúde [22], refletindo as barreiras financeiras que impedem especialmente os mais necessitados de acederem aos cuidados necessários.

A globalização e a disponibilidade de informação estão a alterar o perfil das comunidades, elevando as suas expectativas quanto à qualidade e exigindo uma adaptação às novas necessidades, o que requer do sistema de saúde uma resposta mais proativa e seletiva na escolha e alocação dos profissionais de saúde.

2.1.2. Determinantes de Saúde

É necessário refletir sobre algumas questões como o que torna uma comunidade mais ou menos saudável? Ou quais são os fatores que influenciam a saúde das pessoas ou, inversamente, contribuem para a sua falta? As características da população tanto podem atuar como barreiras à propagação de doenças, quanto criar situações de risco que favorecem a transmissão de certas patologias. Estes fatores populacionais, juntamente com as condições de vida, os hábitos culturais e comportamentais que impactam a saúde, são conhecidos como determinantes sociais de saúde [30].

O termo “determinantes de saúde” surgiu na década de 70 [31]. De acordo com o Relatório de Lalonde [32], os quatro fatores principais que desempenham um papel central na saúde da população são, de forma resumida, os seguintes: (i) genética e fisiologia humana; (ii) estilos de vida; (iii) qualidade do ambiente; (iv) cuidados de saúde (organização e políticas de saúde). Estes determinantes estão intrinsecamente ligados e, em conjunto, moldam o quadro geral da saúde de uma comunidade.

O modelo dos Determinantes Sociais da Saúde (DSS), representado na Figura 2.3, apresenta-se como uma estrutura em camadas concêntricas, onde cada nível exerce influência sobre os subsequentes.

Figura 2.3 *Determinantes sociais da saúde. Adaptado de Dahlgren & Whitehead, 1991 [32].*

No núcleo central encontram-se as características individuais, como idade, sexo e fatores genéticos, sobre os quais há pouca ou nenhuma possibilidade de intervenção.

A primeira camada envolvente representa os comportamentos e estilos de vida individuais, que são influenciados tanto por escolhas pessoais quanto pelo contexto sociocultural. Neste nível, podemos considerar hábitos alimentares, prática de exercícios físicos ou consumo de tabaco. Uma pessoa que cresce numa família que valoriza a alimentação saudável e a atividade física terá mais probabilidade de adotar esses hábitos.

O próximo nível abrange as redes sociais e comunitárias, cuja qualidade e extensão podem impactar a saúde dos indivíduos. Por exemplo, um idoso que participa regularmente em atividades comunitárias e mantém fortes laços sociais tende a ter melhor saúde mental e física do que um idoso isolado.

A camada seguinte engloba as condições de vida e de trabalho, incluindo aspetos como ambiente laboral, educação, habitação, acesso a serviços de saúde e saneamento. Estes fatores estão intimamente ligados às políticas públicas e às condições socioeconómicas. Por exemplo, um trabalhador exposto a condições precárias no local de trabalho tem maior risco de desenvolver doenças ocupacionais, enquanto uma pessoa que vive numa área com saneamento adequado tem menor risco de contrair doenças infeciosas.

A camada mais externa representa o contexto macro em que a sociedade opera, incluindo políticas económicas e sociais, normas culturais e condições ambientais gerais. Por exemplo, políticas que promovem a igualdade de género podem melhorar o acesso das mulheres à educação e aos cuidados de saúde, impactando positivamente a saúde da população em geral. Da mesma forma, normas culturais que estigmatizam certas condições de saúde mental podem dificultar o acesso ao tratamento para os indivíduos afetados.

No geral, os determinantes de saúde estão interligados e exigem uma abordagem abrangente para melhorar os resultados em saúde. Ao abordar os fatores relacionados à biologia humana, estilo de vida, ambiente e organização dos cuidados de saúde, é possível promover a saúde, prevenir doenças e impulsionar o bem-estar geral, refletindo-se no aumento da esperança de vida.

2.2. SIG

A análise espacial de dados em saúde é uma ferramenta poderosa para compreender as complexas interações entre a geografia e a saúde da população. Os métodos aplicados na análise espacial estão na sua maioria implementados em algoritmos específicos nas ferramentas de Sistemas de Informação Geográfica (SIG), os quais permitem desvendar padrões geográficos de saúde e promover tomadas de decisão mais informadas [33].

Os SIG são ferramentas que permitem visualizar, inquirir, analisar e interpretar dados para compreender relações, padrões e tendências de dados geoespaciais. Através da organização dos dados, como está representado na Figura 2.4, em camadas de informação devidamente georreferenciadas, é possível produzir visualizações de dados e análises integrada dos mesmos, cujos resultados são expressos em mapas, gráficos, tabelas de dados ou cenários 3D, que facilitam a compreensão dos dados e das análises produzidas.

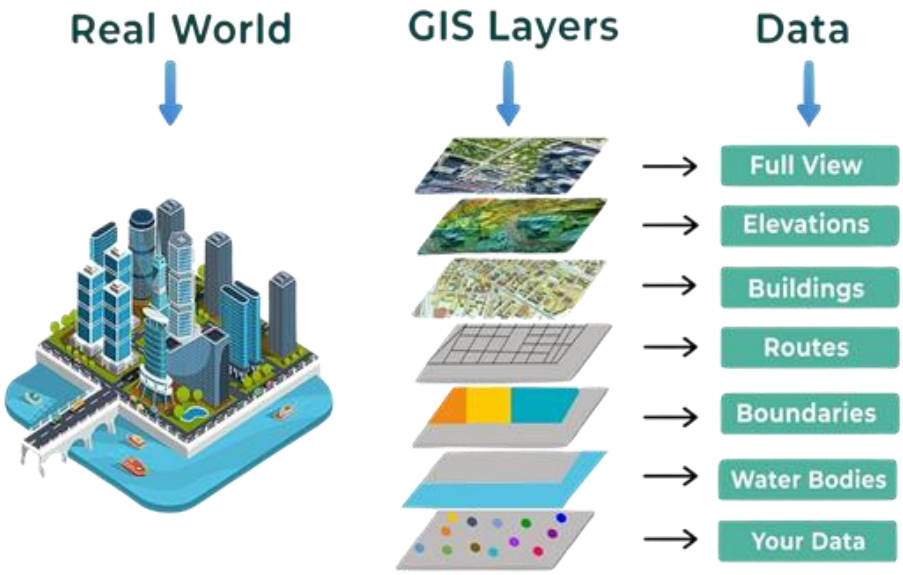


Figura 2.4 *Esquema que ilustra SIG (GIS acrónimo inglês) [33], transformam dados do mundo real em camadas de informação organizadas, facilitando a visualização e análise espacial para melhor compreensão dos dados.*

Através da análise espacial é possível reconhecer que a saúde da população não é uniforme, e que diferentes áreas geográficas podem apresentar desafios de saúde distintos [34]. Partindo do pressuposto de que a localização desempenha um papel fundamental na determinação dos resultados em saúde, fatores geográficos, como proximidade a fontes de poluição, acesso a serviços de saúde e características da envolvente ambiental, têm um impacto direto na saúde das comunidades.

Uma das principais contribuições da análise espacial na saúde é a capacidade de identificar padrões geográficos de saúde. Isto permite visualizar onde certas doenças são mais prevalentes, quais as áreas afetadas por certos determinantes de saúde e como a distribuição geográfica da população se relaciona com resultados de saúde específicos. Através de mapas e outras representações geográficas, estes padrões tornam-se evidentes, destacando necessidades prioritárias.

Além disso, a análise espacial de dados em saúde suporta a tomada de decisões baseada na localização. Esta abordagem é essencial para identificar áreas de alto risco, alocar recursos de forma eficiente e priorizar intervenções nas comunidades que mais necessitam. É indispensável para o

planeamento e avaliação de programas de saúde pública, permitindo que gestores e profissionais de saúde adotem abordagens mais direcionadas e eficazes, melhorando assim a prevenção e promoção da saúde.

As aplicações da análise espacial de dados em saúde são variadas. Alguns exemplos, entre muitos outros, são: monitorização de doenças, rastreamento de surtos de doenças e a identificação de áreas de maior incidência; mapeamento de comportamentos de risco, como tabagismo e práticas alimentares incorretas, a nível geográfico; avaliação do impacto de fatores ambientais na saúde, como a exposição a poluição e o planeamento da localização de centros de atendimento e hospitais [35].

Em síntese, a análise espacial de dados em saúde é uma ferramenta útil e vantajosa para profissionais de saúde pública. Ao revelar padrões geográficos e fornecer informações baseadas na localização, contribui para uma tomada de decisões mais informada, resultando em estratégias de saúde mais eficientes e eficazes, e em abordagens personalizadas para responder às necessidades específicas de comunidades e em larga escala, como aconteceu recentemente com a análise da propagação do Covid.

2.3. Diabetes

A diabetes é uma doença crónica que afeta globalmente mais de 500 milhões de adultos entre os 20 e os 79 anos, dos quais 61,4 milhões se encontram na Europa [36]. Em Portugal, os dados do Inquérito Nacional de Saúde com Exame Físico (INSEF 2015) indica uma prevalência maior em homens (12,1%) em comparação com as mulheres (7,8%) [37].

Esta diferença entre os sexos é apenas uma das várias complexidades da diabetes tipo 2, uma patologia influenciada por múltiplos fatores de risco, incluindo a idade, histórico familiar e excesso de peso. Em Portugal, 68% dos diabéticos referem ter um familiar de primeiro grau com a patologia [37].

No nosso país, a diabetes tem um impacto expressivo, sendo responsável por 2,8% de todas as mortes em 2021 [38]. Entre 2020 e 2022, foram registadas mais de 2,5 milhões de avaliações de risco para esta patologia, nos cuidados de saúde primários, abrangendo 46% da população alvo. Durante o mesmo período, os custos com antidiabéticos não insulínicos e insulinas atingiram 540 milhões de euros [38], evidenciando o peso económico substancial da doença no sistema de saúde português.

A prevenção e gestão da Diabetes Mellitus tipo 2 (DMT2) baseiam-se fortemente na adoção e manutenção de um estilo de vida saudável desde cedo [39]. Um controlo rigoroso dos fatores de risco e intervenções precoces nas complicações podem evitar hospitalizações, reduzir custos e melhorar a qualidade de vida dos pacientes [36].

O baixo estatuto socioeconómico tem um impacto considerável, aumentando o risco de desenvolvimento da DMT2 devido à sua influência em vários aspetos da vida [40]. Fatores como o nível de educação, rendimentos, taxas de pobreza, hábitos como o tabagismo e a dinâmica familiar, têm uma influência significativa na prevalência e progressão da doença [41]. Além disso, estilos de vida pouco saudáveis, caracterizados por uma dieta rica em açúcares, gorduras e calorias, e a inatividade física, têm contribuído para o aumento da prevalência e das complicações associadas à diabetes [36].

Além disso, o ambiente de residência, incluindo a localização rural, condições habitacionais, níveis de segurança e criminalidade, também afetam o acesso a cuidados de saúde e a adoção de estilos de vida saudáveis. A exposição a poluentes atmosféricos, como as partículas finas PM2.5, está igualmente associada a um aumento do risco desta doença [42], [43], [44].

Fatores como a presença de endocrinologistas nos hospitais, a realização regular do teste de hemoglobina A1c, a disponibilidade de programas de autogestão da diabetes e a existência de grupos de apoio são essenciais para uma gestão eficaz da doença [42]. Além disso, a alocação eficiente de médicos de família a pacientes com diabetes tipo 2 desempenha um papel crucial na gestão da condição [44].

2.3.1. Diabetes e SIG

O facto de se usar ferramentas de SIG com aplicação a dados de saúde tem sido possível demonstrar que as capacidades de análise são ampliadas exponencialmente. A análise espacial permite mapear e compreender a distribuição da diabetes, diversas técnicas podem ser aplicadas para abordar vários aspetos da diabetes tipo 2, como por exemplo a sua prevalência, fatores de risco e mortalidade associada.

No estudo efetuado na Índia por Krishnamoorthy et al. [45], os autores utilizaram técnicas de análise espacial para identificar *clusters* geográficos de alta prevalência de diabetes, encontrando uma concentração significativa de casos nos estados do sul da Índia. O estudo também destacou que a deteção de diabetes é mais baixa em áreas rurais e com baixo nível socioeconómico. Esta descoberta sublinha a importância de considerar não apenas a prevalência da doença, mas também as disparidades no acesso ao diagnóstico. O estudo de Krishnamoorthy et al. demonstra claramente a necessidade de considerar variações geográficas em grande escala ao planear intervenções de saúde pública, permitindo uma alocação mais eficiente de recursos para áreas de maior necessidade.

O estudo espácio-temporal da mortalidade por diabetes, realizado na China por Zhou et al. entre 2007 e 2014 [46], usando técnicas de regressão geograficamente ponderada, revelou variações na prevalência e deteção da diabetes ao longo do tempo e entre diferentes regiões. Demostrou que fatores como urbanização e nível socioeconómico têm associações variáveis com a diabetes, dependendo da área geográfica. Em algumas regiões urbanas, a prevalência era maior, enquanto em áreas rurais essa relação era menos evidente ou inversa. Estas descobertas destacam a complexidade da epidemiologia da diabetes e sublinham a necessidade de políticas de saúde pública que sejam adaptadas às realidades locais, em vez de abordagens uniformes a nível nacional.

Nurjannah et al. adotaram uma abordagem inovadora ao examinar padrões espácio-temporais de mortalidade por diabetes em distritos escolares de Michigan, nos Estados Unidos, entre 2007 e 2014 [47]. O estudo destacou-se por focar-se em unidades geográficas menores, em contraste com investigações anteriores que abordaram padrões a nível nacional ou regional. Os autores utilizaram técnicas de análise espacial, incluindo a análise de autocorrelação espacial e a identificação de padrões espácio-temporais. Esta metodologia permitiu-lhes identificar não apenas *clusters* geográficos de mortes por diabetes, mas também caracterizar a evolução temporal destes *clusters*. Os resultados revelaram a existência de "hotspots" persistentes e intensificadores de mortes por diabetes em distritos escolares específicos. Esta abordagem tem implicações práticas significativas, pois permite direcionar programas de prevenção do diabetes baseados em escolas para as áreas de maior necessidade. Além disso, a análise temporal pode informar o planeamento futuro de programas de saúde pública.

Concluindo, a espacialização de dados em saúde evidencia a complexa interação entre fatores geográficos, demográficos, socioeconómicos e comportamentais na DMT2, sublinhando a importância de considerar o contexto ambiental e social na prevenção e controlo da doença e não unicamente os fatores de risco individuais. A utilização de SIG e análises espaciais permite desenvolver estratégias de prevenção e tratamento adaptadas a diferentes regiões e populações, promovendo intervenções de saúde pública mais eficazes e personalizadas.

A complexidade e o volume de dados gerados pelos estudos espaciais e temporais sobre a DMT2, como os mencionados anteriormente, evidenciam a necessidade de ferramentas avançadas para processar e interpretar esta informação de forma eficaz. O cérebro humano, por si só, não tem capacidade para processar toda esta informação em tempo útil. Esta realidade enquadra-se no contexto da 4ª Revolução Industrial, onde a inteligência artificial e a integração de dados surgem como soluções indispensáveis para o processamento eficiente de uma vasta quantidade de informação [48].

Neste cenário, emerge a necessidade de plataformas agregadoras de informação de saúde. No entanto, o desenvolvimento e implementação destas plataformas enfrentam desafios significativos. A falta de interoperabilidade técnica e a dispersão de dados em silos dificultam a análise e implementação eficaz da integração de cuidados, em particular no tratamento da diabetes tipo 2, uma patologia sistémica que requer uma abordagem holística. Os dados de saúde, relevantes para a gestão da diabetes, encontram-se dispersos por hospitais, clínicas, laboratórios e empresas farmacêuticas. Isolados, estes dados não são analisados nem interligados, perdendo-se assim informação valiosa que poderia resultar da sua consolidação.

Além da dispersão de dados, a ausência de uma semântica comum universalmente adotada nos sistemas de informação em saúde e a falta de implementação sólida de standards de comunicação de dados representam obstáculos adicionais. A superação destes desafios é essencial para potenciar a precisão dos cuidados prestados em ambiente clínico aos pacientes com DMT2 e implicará a colaboração dos diversos parceiros do setor da saúde.

O Plano Nacional de Saúde 2020-2030 destaca a importância da equidade no acesso aos cuidados de saúde, a necessidade de uma alocação eficiente de recursos, e a promoção de uma participação ativa dos cidadãos nas decisões de saúde. Estas prioridades são especialmente pertinentes no contexto da gestão da DMT2, uma doença crónica que requer cuidados contínuos e um envolvimento ativo do paciente.

Assim, a criação de plataformas que reúnam dados de saúde, capazes de processar e analisar a vasta quantidade de informação necessária para estudos espaciais e temporais apresentam-se como necessidade e solução. Estas ferramentas não só facilitarão a tomada de decisão clínica, mas também contribuirão para uma gestão mais eficaz dos recursos de saúde e para a melhoria da qualidade dos cuidados prestados aos pacientes em Portugal.

3. Materiais e Métodos

Este capítulo apresenta todas as informações sobre as análises realizadas ao longo do projeto. Foi conduzida uma análise inicial exploratória da distribuição espacial dos atributos, seguida do estudo de autocorrelação espacial da variável DMT2 usando o índice local de Moran [49] que é uma técnica estatística utilizada para analisar a dependência espacial dos dados numa determinada área. Em seguida, e com o objetivo de determinar quais os fatores que mais influenciam a DMT2 de acordo com a localização, foram aplicados métodos de regressão espacial e regressão espacial ponderada (*Geographically Weighted Regression* GWR), os quais permitiram determinar modelos que permitem estimar a DMT2 em função do valor e variabilidade espacial das variáveis que mais contribuem para a prevalênica da DMT2. Neste capítulo são feitas ainda algumas considerações sobre os resultados de ambos os métodos, demonstrando como os mesmos podem atuar de forma complementar na caracterização da população em termos de saúde.

3.1. Metodologia

O processo teve início com uma análise nacional dos centros de saúde e hospitais, para entender como as instituições de saúde estavam distribuídas na região em estudo, em comparação ao resto do país. Com essa contextualização geográfica procedeu-se a uma análise focada exclusivamente na área de estudo, repartida por cinco fases, conforme ilustrado na Figura 3.1. Após a análise exploratória dos dados, a etapa seguinte consistiu no pré-processamento dos MCDT's, na análise exploratória e análise Espácio-temporal, às quais se seguiu a Modelação através do métodos de regressão (OLS) e GWR.

Figura 3.1 *Metodologia utilizada ao longo do projeto.*

3.2. Caracterização da região de estudo

O Litoral Alentejano, situado no sudoeste de Portugal como representado na Figura 3.2, engloba os concelhos de Alcácer do Sal, Grândola, Santiago do Cacém, Sines e Odemira, com uma área aproximada de 5.308 km². A região caracteriza-se pela sua rica biodiversidade e combinação de paisagens únicas, que variam desde planícies extensas e áreas montanhosas, como a Área Protegida do Sudoeste Alentejano e Costa Vicentina [50]. O clima predominante é o mediterrânico, com verões quentes e secos e invernos suaves e húmidos, influenciado pela proximidade do Oceano Atlântico. A precipitação anual varia entre 600 e 800 mm, com temperaturas médias anuais que oscilam entre 15ºC e 18ºC [51].

A população do Litoral Alentejano é superior a 100 mil habitantes (96 442 segundo o Censos de 2021), distribuída de forma desigual pelos cinco concelhos. A densidade populacional é baixa, variando entre 19 e 30 habitantes por km², refletindo o despovoamento típico das áreas rurais alentejanas [52]. A região apresenta uma população maioritariamente envelhecida, com uma grande percentagem de residentes com mais de 65 anos, embora nos últimos anos Sines tenha registado um aumento de jovens, atraídos pelas oportunidades de emprego na área industrial [53].

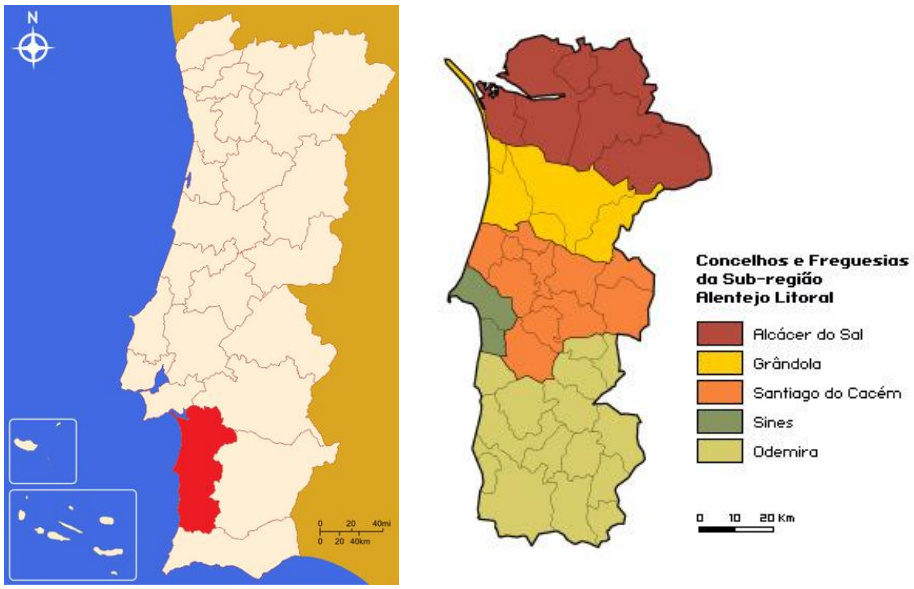


Figura 3.2 *À esquerda, a posição geográfica do Alentejo Litoral em Portugal continental é apresentada em destaque [54], enquanto o mapa à direita detalha a subdivisão dos concelhos na sub-região [55].*

Economicamente, a região depende da agricultura, pecuária e pesca, atividades tradicionais que continuam a ter um peso significativo [53]. Contudo, o setor do turismo tem vindo a crescer, especialmente devido à atratividade das praias e da paisagem natural. O porto industrial de Sines é um dos principais motores económicos da região, sendo um importante polo de atividades industriais e comerciais, com destaque para a petroquímica, a produção de energia e o transporte marítimo.

Para comparar a cobertura de saúde pública no Litoral Alentejano com o restante território de Portugal, foi realizada uma análise detalhada, em que se considerou a distribuição espacial dos centros de saúde e hospitais.

A partir do ficheiro que contém todas as unidades de saúde e hospitais, com as respetivas coordenadas geográficas, procedeu-se à sua representação através de um ficheiro vetorial de pontos usando a ferramenta "*Create Points From Table*" e "*XY Table To Point*", permitindo visualizar espacialmente a distribuição das instituições prestadoras de cuidados de saúde.

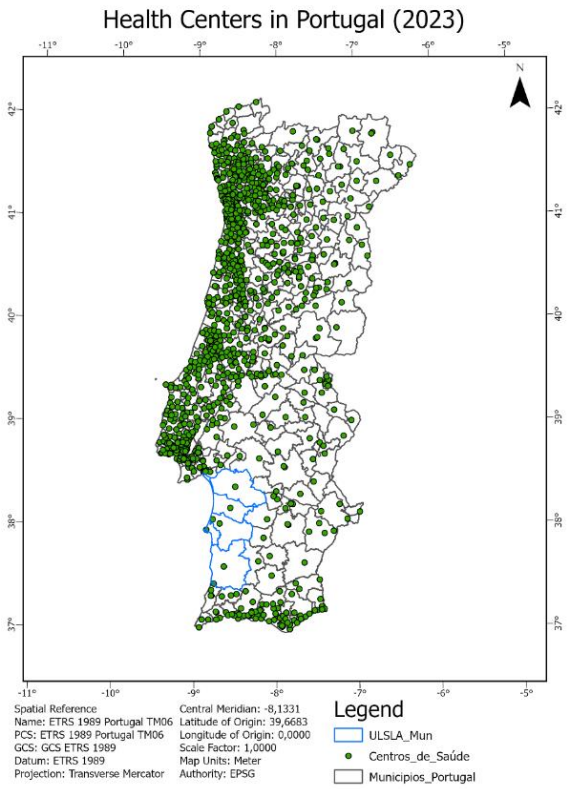


Figura 3.3 *Mapa com a distribuição dos Centros de Saúde em Portugal (2023).*

O Litoral Alentejano, destacado a azul na Figura 3.3, apresenta uma distribuição de centros de saúde que, comparativamente com restante território de Portugal, é bastante dispersa. Enquanto, que nas regiões do Litoral Norte e Centro e Algarve, há uma grande concentração de centros de saúde, o Litoral Alentejano revela-se menos equipado em termos de infraestruturas de saúde primária.

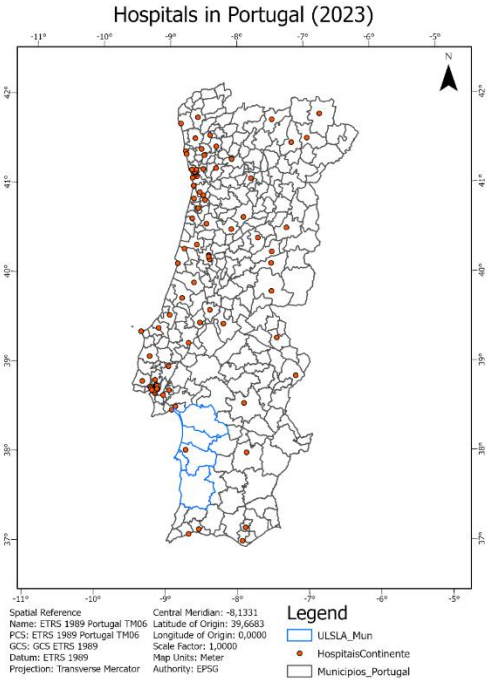


Figura 3.4 *Mapa com a distribuição dos Hospitais em Portugal (2023).*

Relativamente à distribuição de hospitais na Figura 3.4, o Litoral Alentejano apresenta uma escassez ainda mais pronunciada em comparação com outras regiões de Portugal. Enquanto as áreas urbanas e metropolitanas, como Lisboa, Porto e áreas limítrofes, beneficiam de uma maior concentração de hospitais, o Litoral Alentejano conta apenas com um hospital, o que obriga os residentes a longas deslocações para aceder a cuidados hospitalares.

3.3. Dados e metadados

Para facilitar a compreensão da metodologia aplicada, a Figura 3.5 apresenta o *workflow* das várias etapas. Este diagrama esquematiza o processo, incluindo a recolha de dados, o pré-processamento, a análise espacial e as técnicas de modelação aplicadas.

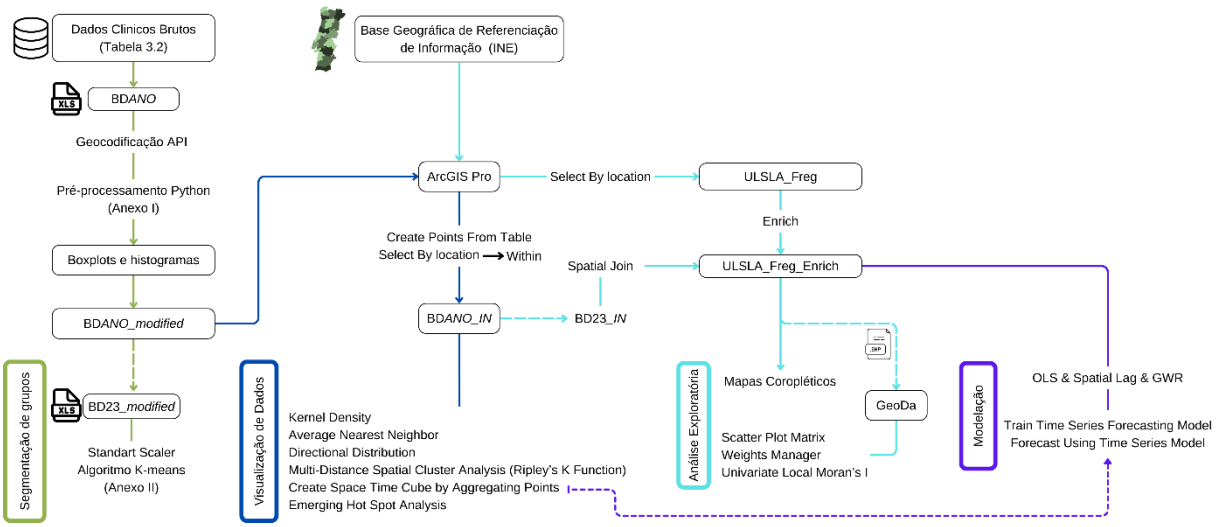


Figura 3.5 *Workflow da metodologia aplicada, utilizando as ferramentas Python, ArcGIS Pro® e GeoDa.*

A unidade espacial usada neste estudo foi a Base Geográfica de Referenciação de Informação (BGRI), apresentada num ficheiro de estrutura vetorial que contêm as subsecções estatísticas referenciadas cartograficamente no sistema de referência ETRS 1989 TM06-Portugal. Nesta representação, o território nacional é discretizado em pequenos polígonos contíguos, que representam as unidades territoriais para as quais são disponibilizados os dados do último Censos de 2021, pelo Instituto Nacional de Estatística (INE).

Para a contextualização das instituições prestadoras de cuidados de saúde da região em comparação ao território nacional, foram reunidos dois conjuntos de instituições que se dividiram em, centros de saúde e hospitais. A cada unidade de saúde e hospital, foi adicionado o concelho e ULS a que pertence, assim como a morada e respetivo código-postal. A fim de serem obtidas as coordenadas de latitude e longitude, foi realizada a geocodificação, utilizando a morada e o código postal, através da ferramenta "*Geocode Addresses"* disponibilizada pelo ArcGIS Pro®.

Para criar em SIG as camadas correspondentes aos DSS's, foi utilizada uma ferramenta de geoprocessamento denominada "*Enrich*" [56]. Esta ferramenta permite adicionar dados demográficos a áreas geográficas presentes nas camadas de informação, utilizando como fonte de dados a consultora Michael Bauer Research [57]. A "*Enrich*" enriquece as camadas geográficas com diversos atributos, considerando a localização dos dados, desde que estejam disponíveis no nível geográfico pretendido. A base de dados está organizada em oito pilares principais: População, Rendimentos, Agregados Familiares, Educação, Emprego, Estado Civil, Idade e Gastos, havendo a hipótese de incluir outros indicadores importantes [58].

A fim de realizar a caracterização da população residente, foram selecionados dados para complementar as informações provenientes da BGRI, utilizando a ferramenta "*Enrich*". Foram incluídos os seguintes indicadores: **População Total Feminina e Masculina**; **Densidade Populacional**; **Poder de Compra per capita**, que representa o rendimento médio disponível por pessoa, refletindo a capacidade económica média dos residentes da região; **Gasto em álcool per capita**, que indica o valor médio em euros gasto por pessoa em álcool, considerando apenas o consumo doméstico e **Gasto em tabaco per capita**, que abrange todas as compras de tabaco, seja em restaurantes, bares, cafés, estações de serviço, entre outros. Os **Rendimentos**, foram classificados em cinco quintis: o 1º quintil inclui famílias com rendimentos abaixo de 13.296 €, o 2º quintil, com rendimentos entre 13,296 € e 19,250 €, o 3º quintil, com rendimentos entre 19,251 € e 26,418 €, o 4º quintil, com rendimentos entre 26,419€ e 38,608€, e o 5º quintil, famílias com rendimentos superiores a 38,609€. Quanto à **Faixa etária** foram considerados os números de residentes nos grupos etários de 0 a 14 anos, 15 a 29 anos, 30 a 44 anos, 45 a 59 anos, e 60 anos ou mais. Relativamente à **Educação**, foram incluídos dados sobre residentes com o 1º ciclo, 2º ciclo, 3º ciclo, ensino secundário, ensino pós-secundário e ensino superior, concluídos. Também foram analisados dados sobre **a População Desempregada,** ou seja, o número de residentes desempregados que estão à procura de emprego.

Tabela 3.1 *Descrição dos Conjuntos de Dados Utilizados no Estudo*

Conjunto de Dados	Fonte	Tipo de Dados	Unidade Espacial	Período de Referência dos Dados	Período de Acesso /Recolha	Observações
BGRI	INE	Dados geográficos e censitários	Subsecção estatística	2021	Fevereiro de 2024	-
Centros de Saúde e Hospitais	Lean Health	Lista com a localização de centros de saúde e hospitais	Pontos geográficos	2023	Fevereiro de 2024	Geocodificados a partir de moradas e códigos postais
Camadas dos DSS's	"Enrich" (Michael Bauer Research)	Demográficos	Freguesia	2023	Março de 2024	População e Densidade Populacional
						Idade
						Estado Civil
						Habilitações Literárias
		Socioeconómicos				Rendimentos
						Poder de Compra
						Desemprego

No contexto do estudo foi fundamental garantir que o acesso a informações clínicas sensíveis fosse feita de forma ética e segura. Para o efeito foi necessário elaborar um protocolo para submissão à Comissão de Ética da ULSLA, que exigiu uma análise detalhada e bem fundamentada dos dados necessários.

Inicialmente, procedeu-se à revisão de literatura científica para identificar os indicadores clínicos mais relevantes para o estudo da DMT2. Esta revisão foi indispensável para definir quais os dados essenciais para a investigação e as metodologias a serem aplicadas. No protocolo submetido à Comissão de Ética, foram detalhados os objetivos da recolha de dados e a sua aplicação no contexto do estudo. Foi dada especial atenção às medidas que visavam garantir a confidencialidade e a segurança dos dados, incluindo procedimentos de anonimização e proteções contra acessos não autorizados. A transparência e o rigor ético na manipulação de dados sensíveis foram fundamentais para a aprovação do estudo pela Comissão de Ética, assegurando que a investigação pudesse avançar com um sólido compromisso para com os direitos e a privacidade dos participantes.

Após a aprovação da Comissão de Ética, deu-se início à recolha dos dados, detalhados na Tabela 3.2, na plataforma Módulo de Informação e Monitorização das Unidades Funcionais (MIM@UF). Esta

ferramenta, implementada em todas as Unidades de Saúde da Região de Saúde do Alentejo, foi projetada para otimizar a gestão da informação clínica. O sistema permite a integração e o acesso a dados detalhados dos utentes, facilitando tanto a análise clínica quanto a monitorização da prescrição de medicamentos e dos meios complementares de diagnóstico e terapêutica (MCDT). A extração foi realizada: por ano, por Unidade de Saúde e por médico de medicina geral e familiar, incluindo os utentes sem médico.

Tabela 3.2 *Dados Clínicos da População Diabética, incluindo Dados Demográficos, Geográficos, Estilo de Vida, Comorbilidades e Meios Complementares de Diagnóstico e Terapêutica (MCDT's).*

Conjunto de Dados	Fonte	Período de Referência dos Dados	Período de Acesso /Recolha	Formato dos Dados	Observações
Dados demográficos	MIM@UF	2017-2023	Abril a Junho de 2024	Numérico	Idade
				Categórico	Nacionalidade
				Binário	Sexo
					Desemprego nos últimos 12 meses
					Insuficiência Económica (≤720,65€)
Dados geográficos				Texto	Morada utilizada para georreferenciação
Atribuição de Médico de Família				Binário	Com ou sem Médico de Família
Estilo de Vida					Abuso de Tabaco (P17)
					Abuso Crónico de Álcool (P15)
Comorbidades					Neuropatia (N94)
					Retinopatia (F83)
					Glomerulonefrite (U88)
					Pielonefrite (U70)
					Doença Urinaria - Outra (U99)
					AVC (K90)
					EAM (K75)
					Doença Vascular Cardiaca (K83)
					Doença Vascular Cerebral (K91)
					Aterosclerose (K92)
MCDT's				Numérico	IMC
					Tensão Distólica
					Tensão Sistólica
					HDL
					LDL
					Trigliceridos
					Hemoglobina A1c (Glicada) - (Sangue)
					Perímetro Abdominal
					Glicémia Capilar
		2019-2023			Creatinina

Após a recolha de dados, procedeu-se à sua agregação, com o propósito de reunir toda a informação, por ano e por utente num só ficheiro, resultando em bases de dados anuais, para permitir posteriormente uma análise temporal. Através das moradas, freguesias e localidades fornecidas, foi realizada a geocodificação utilizando uma API disponível online [59], a fim de obter as coordenadas de latitude e longitude de cada utente, essenciais para a georreferenciação na base de dados SIG.

3.3.1. Pré-processamento dos Dados dos Utentes

Para assegurar a qualidade e a consistência dos dados utilizados nas análises subsequentes, foi realizado um pré-processamento dos ficheiros de Excel referentes aos utentes com diabetes do código de Python, com o objetivo de eliminar inconsistências, formatando os dados afim de assegurar a sua qualidade, preparando-os para as análises subsequentes.

Cada ficheiro foi lido e convertido em *dataframes* utilizando a Biblioteca Pandas, o que facilitou tratamento posterior dos dados. Durante esta fase, foram verificadas e corrigidas as tipologias de dados de cada coluna, garantindo que estivessem no formato adequado para a análise. Além disso, foram eliminadas colunas desnecessárias para simplificar o conjunto de dados, e procedeu-se à anonimização dos utentes. A coluna referente ao nome do médico de família foi transformada de uma variável

categórica para uma variável binária, onde 0 indicava a ausência de médico de família e 1 a presença de um. Este procedimento visou a padronização da variável para uma análise mais eficiente.

Foram criadas tabelas de estatísticas descritivas para os indicadores dos MCDT's, foram gerados boxplots para identificar possíveis outliers e dados incoerentes. Após esta análise, foram definidos limites para remover dados considerados redundantes ou inconsistentes.

Os valores em falta (NaNs) foram estimados e preenchidos, segmentando os dados por sexo, idade e unidade funcional de saúde, de modo a manter a integridade das análises subsequentes. É importante perceber, que esta segmentação não garante um preenchimento total de todos os valores em falta, havendo a possibilidade de alguns grupos de segmentação serem vazios.

Após o preenchimento dos valores em falta, foram novamente gerados boxplots e matrizes de histogramas para verificar a distribuição dos dados, garantindo que as correções realizadas não introduziram novas inconsistências. Tabelas de estatísticas descritivas foram recriadas para assegurar a conformidade dos dados. Após a conclusão do pré-processamento, os *dataframes* finais, contendo as correções e normalizações aplicadas, foram exportados com o nome 'BD*ANO*modified'. Estes ficheiros foram então preparados para serem integrados nas ferramentas de análise espacial em ambiente SIG.

3.3.2. Segmentação dos Utentes

A segmentação permite identificar subgrupos homogéneos dentro da população, facilitando a implementação de estratégias de tratamento e prevenção mais direcionadas. A abordagem adotada nesta secção envolveu a aplicação do algoritmo de *clustering K-means* [60] para a segmentação da população diabética do ano de 2023, uma vez que os dados já tinham sido processados anteriormente.

Os dados foram importados do ficheiro CSV denominado 'BD23modified.csv', que contém informações clínicas e demográficas de pacientes diabéticos referentes ao ano de 2023. Foram selecionadas variáveis numéricas relevantes para a segmentação, incluindo: Idade, Índice de Massa Corporal (IMC), Hemoglobina Glicada (HbA1c), Perímetro Abdominal, indicadores de comorbidades e indicadores socioeconómicos. As variáveis de identificação pessoal, como o número de anonimização e as coordenadas geográficas (latitude e longitude) foram excluídas para focar nas características clínicas e demográficas dos pacientes.

Aplicou-se o método de *StandardScaler* para normalizar os dados. Este processo ajusta os valores das variáveis para que tenham uma média de zero e um desvio padrão de um. Para garantir que todas as variáveis contribuam igualmente para o processo de *clustering*, evitando que variáveis com escalas maiores dominem a segmentação. Após a preparação dos dados, procedeu-se à segmentação da população diabética através do algoritmo de *clustering K-means*.

O algoritmo *K-means* foi escolhido devido à sua eficácia e simplicidade na segmentação de dados em *clusters* distintos. O *K-means* agrupa os dados em k *clusters*, onde cada observação pertence ao *cluster* com o centróide mais próximo, minimizando a variabilidade *intra-cluster*.

Optou-se por definir k=4 *clusters*. Esta escolha foi baseada em considerações práticas e na necessidade de dividir a população em subgrupos suficientemente distintos para uma análise detalhada. Embora, tenha sido aplicado o método do cotovelo para determinar o número ideal de *clusters*, que indicou ser entre 3 e 4.

Os dados são iterativamente atribuídos aos *clusters* com base na proximidade ao centróide mais próximo. Os centróides são recalculados até que a atribuição dos dados aos *clusters* se estabilizasse,

indicando a convergência do algoritmo. No fim do algoritmo, todos os utentes foram distribuídos a um dos quatro *clusters* identificados pelo algoritmo. Sendo possível criar uma caracterização detalhada, e diferencial, de cada grupo existente dentro da população diabética na ULSLA.

3.3.3. Visualização de Dados

A partir do ficheiro com todos os utentes diabéticos, classificados como T90 através do ICPC-2, procedeu-se à representação por pontos usando as ferramentas "*Create Points From Table*" e "*XY Table To Point*", o que permite visualizar, através de pontos, observar num mapa a distribuição dos utentes na região, por ano.

Como as moradas registadas são fornecidas pelos próprios utentes, podendo alguns estar registados num centro de saúde do Litoral Alentejano, mas com moradas fora da região, alguns pontos foram excluídos para não comprometer as análises. Para isto, utilizou-se a ferramenta "*Select Layer by Location*", utilizando como *input feature* a camada de pontos BD*ANOmodified*, (ex. BD17; BD18...), a relação "*Within*" e como *Selecting Feature* a camada das freguesias ULSLA_Freg. Depois de selecionados os pontos interiores à região, usou-se a ferramenta "Export Features" para criar uma camada com o nome de BD*ANO*_IN.

De forma semelhante à secção 3.2., foi utilizada a ferramenta *"Kernel Density"*. Complementarmente, recorreu-se à ferramenta "*Average Nearest Neighbor*" (ANN), que calcula a distância média entre cada utente diabético e o utente mais próximo, comparando-a com uma distribuição espacial aleatória. Um índice abaixo de 1 indica um padrão de agrupamento, sugerindo que os utentes estão mais próximos entre si do que seria de esperar num padrão aleatório. Esta análise é essencial para visualizar e determinar o tipo de distribuição dos utentes ao longo da região do Litoral Alentejano, identificando se a distribuição é mais agrupada ou dispersa.

Adicionalmente, foi aplicada a ferramenta "*Directional Distribution*", também conhecida como Elipse de Desvio Padrão, no ArcGIS Pro®, para analisar a orientação e dispersão geográfica dos dados. Esta ferramenta calcula uma elipse que abrange a distribuição espacial dos pontos de dados, indicando a orientação e a dispersão através dos eixos maior e menor, respetivamente. O ângulo de orientação da elipse, calculado em relação ao norte geográfico, auxilia na visualização e quantificação do sentido da dispersão dos dados. Esta análise é útil para compreender não apenas onde os eventos ocorrem, mas também como eles estão distribuídos no espaço.

Uma vez verificada a existência de *clusters*, com a ferramenta ANN, foi realizada uma análise utilizando a função "*Multi-Distance Spatial Cluster Analysis (Ripley's K Function)*", no ArcGIS Pro®, tendo como objetivo determinar qual a dimensão média desses clusters. Para o efeito, foi necessário reprojectar os dados da camada vetorial dos pontos representando a localização de cada utente, usando a ferramenta "*Project*", para um sistema de referencia cartográfico, uma vez que os dados foram inicialmente georreferenciados através das suas coordenadas geográficas (latitude e longitude). Para se aplicar esta ferramenta, definiu-se um polígono através do método *"Minimum enclosing rectangle"*, que estabelece uma região definida pelo menor retângulo capaz de englobar todos os pontos. Para a aplicação da função K Ripley definiram-se 10 bandas de distância de dimensão de 3 175 metros, e optou-se por aplicar 9 permutações. Esta abordagem permite quantificar a significância dos padrões de *clustering* observados, oferecendo uma medida de confiança maior, para as conclusões derivadas da análise espacial. As permutações são utilizadas para testar a hipótese nula de que os padrões observados poderiam ser resultado de uma distribuição espacial aleatória [61], fornecendo assim uma base para interpretação dos padrões de *clustering* como estatisticamente significativos ou não. A escolha de realizar 9 permutações equilibra eficácia computacional com rigor estatístico, embora ofereça uma

estimativa menos robusta da variabilidade dos dados em comparação com um maior número de permutações. Esta abordagem é útil para obter resultados rápidos e uma visão inicial dos padrões espaciais, mas pode ser menos sensível a pequenas variações nos dados. Esta análise permitiu observar os padrões de *clustering* de maneira geral e rápida. Embora a escolha de um menor número de permutações possa limitar a granularidade dos resultados, ainda assim proporcionou uma dimensão espacial útil dos *clusters* encontrados.

Adicionalmente, foram utilizadas as ferramentas "*Create Space Time Cube by Aggregating Points*" e "*Emerging Hot Spot Analysis*" com o objetivo de identificar e analisar os padrões espaciais e temporais dos casos de T90 no Litoral Alentejano. Este processo permitiu uma compreensão mais detalhada das alterações ao longo do tempo, identificando *hot spots* persistentes, emergentes e em contração.

A ferramenta "*Create Space Time Cube by Aggregating Points*" é utilizada para converter dados pontuais num formato que permite uma análise integrada tanto espacial como temporal. Esta ferramenta agrupa os dados em intervalos de tempo e espaço, criando um "cubo espácio-temporal" que facilita a identificação de padrões ao longo do tempo e em diferentes locais, permite entender não apenas onde, mas também quando os eventos ocorrem. Neste contexto, conforme ilustrado na Figura 3.6 abaixo, a estrutura de dados resultante pode ser visualizada como um cubo tridimensional composto por intervalos de espaço-tempo, com as dimensões x e y representando o espaço e a dimensão t representando o tempo. Isto permite, por exemplo, criar uma serie temporal de uma posição fixa (x,y), destacada a amarelo, ou analisar uma fatia temporal de toda a região, indicada a verde.

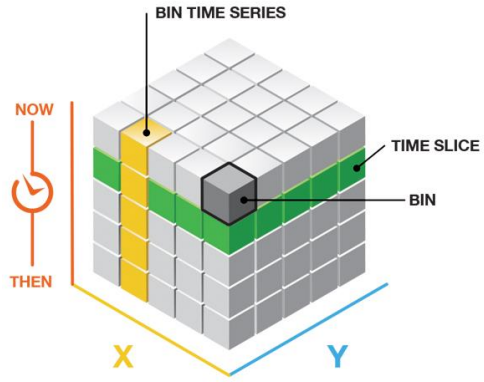


Figura 3.6 *Cubo espácio-temporal que ilustra a ferramenta "Create Space Time Cube by Aggregating Points", imagem retirada da ESRI [62].*

Para utilizar esta ferramenta, foi necessário agregar todas as bases de dados anuais, usando a ferramenta "*Merge*". Foram integradas as bases de dados desde BD17_IN até BD23_IN, resultando numa numa base de dados consolidada denominada de BD_IN_Merge. Esta base unificada foi então utilizada para criar o cubo tridimensional espácio-temporal.

A ferramenta "*Emerging Hot Spot Analysis*" é utilizada para identificar padrões espaciais e temporais de forma mais detalhada, destacando áreas que apresentam aumentos ou diminuições significativas na incidência de eventos ao longo do tempo. Esta ferramenta trabalha com o cubo espácio-temporal previamente criado para detetar *hot spots* (áreas de alta incidência) e *cold spots* (áreas de baixa incidência) que sejam emergentes, persistentes e decrescentes. Ou seja, após a da criação do cubo tridimensional, analisa as alterações ocorridas ao longo do tempo e classifica cada posição fixa (x,y) com base nas mudanças de prevalência observadas, resultando numa representação bidimensional,

conforme ilustrado na Figura 3.7. Esta análise permite investigar as mudanças que ocorreram em *hot spots* decrescentes, para que possam ser aplicadas estratégias em *hot spots* emergentes, permitindo desenvolver estratégias, promover intervenções e monitorização contínua da doença no Litoral Alentejano.

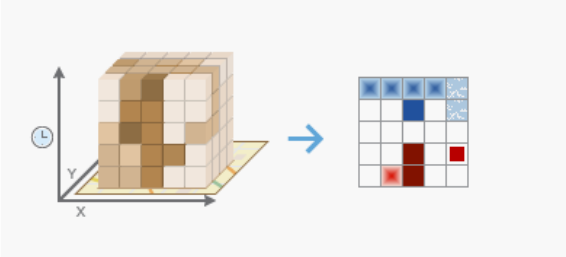


Figura 3.7 *Análise de Hot Spots Emergentes, utilizando o cubo espácio-temporal para identificar padrões de mudança ao longo do tempo, imagem retirada da ESRI [63].*

3.3.4. Autocorrelação Espacial Local

Nesta subsecção do projeto, o objetivo é compreender com maior precisão como é que os atributos usados se distribuem ou se agrupam no espaço e utilizar essa informação para interpretar, por exemplo quais os DSS's que estão correlacionados com a prevalência da Diabetes, especificamente na região do Litoral Alentejano, estabelecendo assim uma relação entre os DSS's e a prevalência da Diabetes Tipo 2 em cada freguesia.

Para isso, foi realizada uma análise exploratória da camada de informação enriquecida com os Metadados mencionados anteriormente. Com a camada de informação "ULSLA_Freg_Enrich", criaram-se mapas coropléticos e gráficos de dispersão, e foi obtido o respetivo índice de Moran [64], de modo a compreender a autocorrelação espacial local.

Para realizar uma análise de autocorrelação espacial e elaborar ~~os~~ gráficos de dispersão dos atributos em estudo recorreu-se ao índice de Moran Local disponível no GeoDa [65]. Para tal, foi necessário adicionar à camada "ULSLA_Freg_Enrich" a coluna "*Join_Count*", que contém o número de utentes T90 por freguesia, através da ferramenta "*Spatial Join*", e converter a camada do tipo geodatabase para shapefile. Durante este processo, houve o cuidado de converter a tipologia das colunas para 32-bits e ajustar o *alias* das colunas para 8 caracteres, o que resulta num nome alternativo que fornece uma terminologia mais amigável do conteúdo.

Com o ficheiro em formato shapefile, a camada foi importada para o GeoDa. Para os gráficos de dispersão, foram usadas as ferramentas "*Explore*" e "*Scatter Plot Matrix*", com o objetivo de analisar quais os indicadores que têm maior correlação com a prevalência de utentes T90 na região em estudo. As variáveis foram agrupadas de forma a facilitar a compreensão de cada gráfico. Os resultados ficam apresentados na forma normalizada, para apresentar os valores de correlação entre +1 e -1. Um valor próximo de +1 indica que à medida que uma variável aumenta, a prevalência também aumenta. Um valor próximo de -1 sugere que à medida que uma variável aumenta, a prevalência diminui. Já um valor de 0 indica uma correlação linear fraca ou inexistente entre as variáveis. O GeoDa foi escolhido porque usa asteriscos (*) para indicar a significância estatística da correlação, um asterisco (*) normalmente indica significância ao nível de 0.05, o que significa que há menos de 5% de probabilidade de que a correlação observada seja devido ao acaso, enquanto dois asteriscos (**) indicam um nível de significância ainda mais baixo de 0.01. A seleção das variáveis foi baseada num limiar de valor de correlação superior a 0.9 e na presença dos dois asteriscos.

Para calcular o índice de Moran, foi necessário criar uma relação de vizinhança entre as unidades espaciais (freguesias) para as quais se pretende obter os resultados de autocorrelação, e para o efeito utilizaram-se as ferramentas "*Tools*" e "*Weights Manager*". Isto porque a autocorrelação espacial mede a semelhança ou influência das características de um local em relação às características dos locais vizinhos, por isso é importante definir a relação de vizinhança (baseada na distância ou contiguidade dos polígonos que representam as freguesias). Neste caso, foi usada a recomendação do uso de "*Queen Contiguity*" de primeira contiguidade, que considera unidades como vizinhas se compartilharem um vértice ou uma borda, de forma a lidar com potenciais incoerências. Este método permite uma análise mais inclusiva e detalhada das relações espaciais.

Depois de definir a matriz de vizinhança espacial, procede-se à análise de "*Univariate Local Moran's I*" através da ferramenta "*Space*", resultando um gráfico de dispersão de Moran e mapas de *cluster* e significância. Esta análise produz um mapa que categoriza regiões (freguesias) em cinco classes: "*High-High Cluster*", que identifica freguesias onde o valor do atributo é elevado e que estão rodeadas por freguesias com valores igualmente elevados; "*Low-Low Cluster*" que correspondente a freguesias onde o valor do atributo é baixo e as áreas adjacentes também apresentam valores baixos; "*High-Low Outliers*", que são freguesias com valores elevados do atributo cercadas por freguesias com valores baixos; e "*Low-High Outliers*", que são o oposto [66]. Há ainda freguesias classificadas como "não significantes". Para cada variável, são calculados *z-score* e *p-value*, que indicam a significância estatística e são usados para testar a hipótese de aleatoriedade espacial na autocorrelação espacial dos atributos estudados.

3.3.5. Modelação

Com a modelação pretende-se estimar modelos que estimem uma determinada variável a partir da contribuição de um conjunto de variáveis explicativas. Os modelos mais adequados para o efeito, são os modelos de regressão que são amplamente aplicados em análises para avaliar a relação entre variáveis dependentes e independentes, permitindo estimar a influência das variáveis explicativas, fazer previsões e testar hipóteses [67].

Quando as variáveis têm uma natureza espacial e se pressupõe que a relação de uma variável que se pretende estimar o valor médio da variável, depende de forma linear das variáveis explicativas, pode-se recorrer tanto a modelos de regressão não espaciais como a modelos de regressão espaciais.

Para a aplicação dos modelos de regressão não espaciais, é essencial verificar um conjunto de pressupostos, como os resíduos apresentarem média zero e variância constante. Para avaliarmos estes pressupostos deve-se analisar a homoscedasticidade (a mesma variância), a normalidade dos resíduos, a independência das observações, e ausência de multicolinearidade. A não conformidade com estes pressupostos pode comprometer a qualidade da modelação [68]. O diagnóstico do modelo envolve a avaliação das estimativas dos parâmetros, dos resíduos e do ajustamento global. A normalidade dos resíduos foi analisada através do teste Jarque-Bera [69], enquanto a homoscedasticidade foi avaliada com o teste Koenker [70]. A ausência de padrões de autocorrelação espacial nos resíduos foi verificada através do teste Local Moran's I, pois se existir autocorrelação nos resíduos, pode indicar que o modelo não considerou corretamente a estrutura espacial dos dados. O ajustamento global do modelo (performance do modelo) foi avaliado com o teste de Wald, o coeficiente de determinação (R^2) ajustado e o Critério de Informação de Akaike corrigido (AICc) [71].

O R^2 indica a proporção da variância na variável dependente (neste caso, prevalência de utentes com diabetes) que é explicada pelo modelo. Um R^2 elevado indica que o modelo explica melhor a variação nos dados. O AICc é utilizado para comparar diferentes modelos, sendo que um valor mais

baixo indica um modelo com melhor ajuste, levando em consideração a penalização por complexidade do modelo.

Afim de gerar o modelo OLS para as variáveis independentes desejadas, utilizou-se a ferramenta "*Ordinary Least Squares (OLS)*". No entanto, apesar da sua eficácia em muitos cenários, este modelo não incorpora a dependência espacial, o que pode ser problemático quando as observações apresentam relações geográficas entre si. Além da modelação tradicional que não considera a dependência espacial, é essencial recorrer a modelos de regressão espacial quando se verifica que as observações apresentam uma relação geograficamente próxima. A omissão da autocorrelação espacial pode resultar em estimativas enviesadas e ineficientes, dado que as observações próximas tendem a partilhar características semelhantes [65].

Para entender a dependência entre freguesias, optou-se pela aplicação do modelo *Spatial Lag*. Este modelo foi desenhado para incluir explicitamente a autocorrelação espacial na variável dependente, o que é feito através da adição do termo de desfasagem espacial, representado pelo coeficiente (ρ) [72]. Este coeficiente mede a intensidade com que os valores da variável dependente nas freguesias vizinhas influenciam o valor da mesma variável na freguesia em análise.

A modelação espacial com o *Spatial Lag* permite, assim, corrigir a distorção nos coeficientes e melhorar o ajuste global do modelo, tendo em conta que as freguesias vizinhas não são independentes entre si [73]. A aplicação do modelo justifica-se devido à autocorrelação espacial presente nos resíduos, indicando que o modelo OLS não captou completamente a estrutura espacial dos dados. Ao incorporar a dependência espacial, o objetivo é obter estimativas mais robustas e fiáveis das relações entre os fatores socioeconómicos e a variável dependente, tendo em consideração o impacto espacial das freguesias adjacentes.

Adicionalmente, foi também aplicada a Regressão Geográfica Ponderada (GWR), que complementa o *Spatial Lag* ao permitir modelar variações locais nos coeficientes das variáveis explicativas. A GWR ajusta os coeficientes de regressão para cada freguesia, o que possibilita a identificação de variações espaciais significativas nas relações entre as variáveis explicativas e a variável dependente.

Na análise multivariada, é comum utilizar métodos de regressão globais que aplicam a mesma equação a todo um conjunto de dados, resultando em estimativas gerais das relações entre os parâmetros. No entanto, em contextos de não estacionariedade espacial, estas técnicas podem ser inadequadas, pois assumem que a relação entre variáveis a nível global se aplica uniformemente em todas as localizações. A não estacionariedade espacial refere-se à variação dessa relação em diferentes áreas, onde o efeito de uma variável pode ser positivo em algumas regiões e negativo noutras [74].

Para lidar com essa variação, a GWR é uma técnica que permite modelar essas diferenças espaciais, ajustando os coeficientes para cada localização com base na proximidade geográfica. Ao contrário dos modelos globais, que tratam os dados como homogéneos, a GWR estima os coeficientes localmente, considerando a heterogeneidade espacial [75], [76].

Para gerar o modelo GWR com as mesmas variáveis independentes, utilizou-se a ferramenta "*Geographically Weighted Regression (GWR)*", com a definição para o número de vizinhos como "*Golden Search*" onde o modelo percorre um intervalo para identificar o número de vizinhos que

produzem os melhores indicadores. Durante a simulação, o modelo de regressão é formulado pela Equação 3.1.

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (3.1)$$

Em que: y_i equivale ao valor da variável dependente, x_{ik} equivale ao valor da variável independente k e ε_i equivale às variáveis aleatórias que correspondem ao erro, na observação, i=1,…, n;

β_k equivale aos parâmetros do modelo e (u_i, v_i) correspondem às coordenadas geográficas.

3.3.6. Modelação preditiva espácio-temporal

Para projetar a incidência de diabetes em 2026, utilizou-se um conjunto de ferramentas de previsão de séries temporais no ArcGIS. O primeiro passo envolveu a instalação de um pacote de *deep learning* [77], indispensável para a execução das próximas ferramentas.

Na ferramenta "*Train Time Series Forecasting Model*", foi utilizado o cubo tridimensional previamente criado na secção 3.3.3 como a base de dados temporal. Esta ferramenta é responsável por treinar um modelo de previsão, utilizando uma série temporal de dados. O principal parâmetro configurado foi o "*Sequence Length*", que determina o número de passos temporais anteriores usados para treinar o modelo. Neste contexto, o "*Sequence Length*" especifica a quantidade de anos anteriores que serão considerados para prever a incidência futura. O modelo foi treinado para identificar padrões temporais e relacionais nos dados de prevalência de diabetes.

Com o modelo treinado, a etapa seguinte foi utilizar a ferramenta "*Forecast Using Time Series Model*". Esta ferramenta aplica o modelo treinado ao cubo tridimensional original para prever os valores futuros, especificamente a incidência de diabetes para 2026. Foram configurados parâmetros específicos, como o "*Number of Time Steps to Forecast*", que foi definido para prever um ano em avanço. A ferramenta combina o modelo preditivo com os dados temporais existentes para gerar uma previsão espacialmente explícita sobre a incidência futura, que é então visualizada na área geográfica em estudo.

O resultado foi um mapa que se pode visualizar a distribuição prevista da incidência de diabetes em 2026, nas diversas freguesias da região. Este processo permitiu uma análise detalhada e baseada em dados históricos, proporcionando uma visualização clara das áreas que podem vir a exigir mais atenção de futuro, facilitando assim o planeamento de recursos e intervenções na área da saúde.

4. Resultados

Neste capítulo apresentam-se e discutem-se os resultados obtidos nas fases de análise exploratória e de modelação.

4.1. Contextualização

Para complementar a caracterização da área em estudo, foi realizado um mapa da densidade das instituições na área continental de Portugal, recorrendo à ferramenta de "*Kernel Density"*, resultando num mapa de calor que destaca as áreas com maior concentração de unidades de saúde. Esta abordagem é útil para identificar padrões de alocação de recursos.

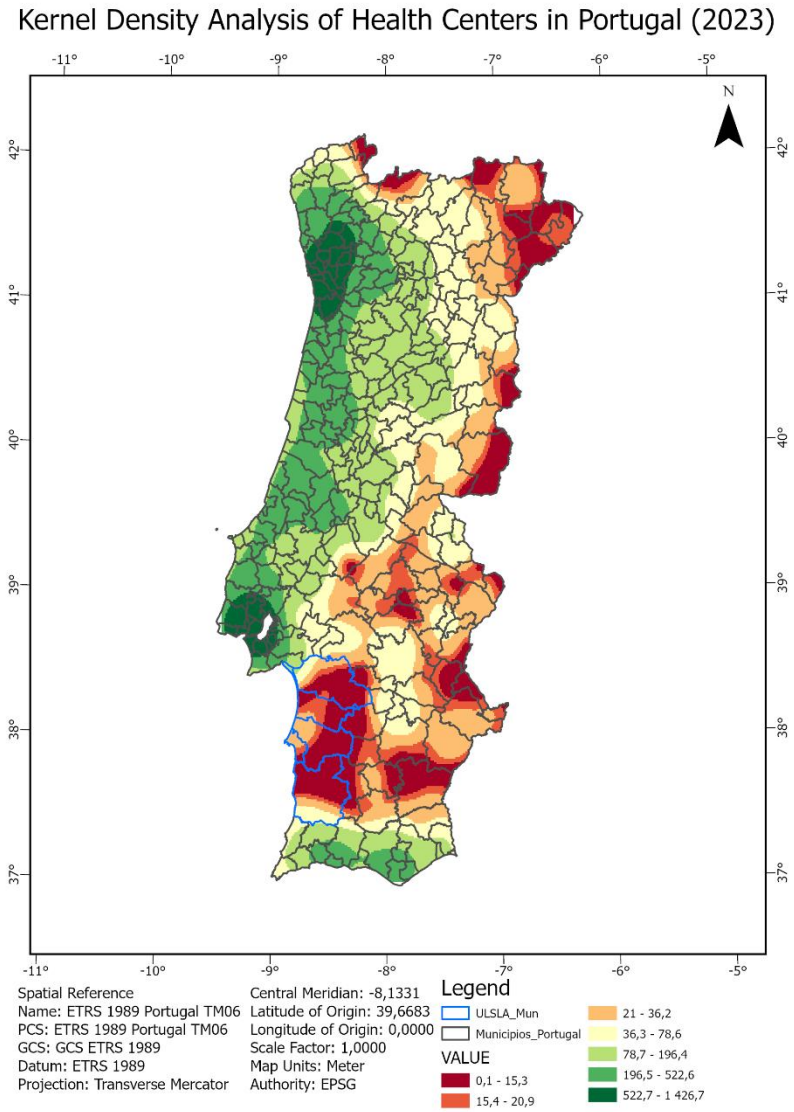


Figura 4.1 *Análise da Densidade de Kernel dos Centros de Saúde em Portugal (2023). A cor vermelha corresponde a zonas com menor densidade de centros de saúde.*

A análise de densidade por *kernel*, presente na Figura 4.1, confirma e reforça a observação anterior. No Litoral Alentejano, a densidade de centros de saúde é consideravelmente menor em comparação com outras regiões do país, assinaladas a vermelho. As áreas do Litoral Norte e Centro, e algumas zonas do Algarve, apresentam densidades mais altas, indicando uma melhor cobertura de serviços de saúde primária, assinaladas a verde.

Esta disparidade no Litoral Alentejano tem impacto direto na acessibilidade aos cuidados de saúde, pois implica que os utentes necessitem de se deslocar frequentemente para outras áreas a fim de obter cuidados médicos, particularmente quando procuram especialidades.

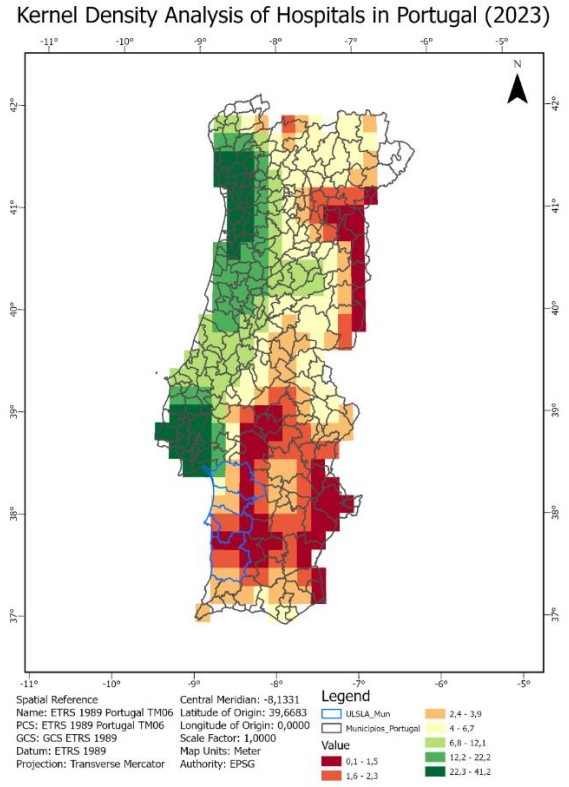


Figura 4.2 *Análise da Densidade de Kernel dos Hospitais em Portugal (2023). A cor vermelha corresponde a zonas com menor densidade de hospitais.*

A análise de densidade por *kernel* para hospitais confirma esta discrepância. Constata-se que no Litoral Alentejano, a densidade hospitalar, assinalado a vermelho na Figura 4.2, é mais baixa contrastando com as regiões do Litoral Norte e Centro onde a densidade de hospitais é notavelmente mais alta.

A comparação entre o Litoral Alentejano, com apenas 1 hospital para os 96 442 habitantes, e o restante país revela que a região sul está em clara desvantagem no que diz respeito à distribuição e acessibilidade de infraestruturas de saúde. Esta situação pode ter importantes repercussões para a saúde pública da região, sublinhando a necessidade de construir um perfil de saúde da comunidade para melhorar a cobertura dos serviços de saúde e garantir que a população do Litoral Alentejano tenha acesso adequado aos cuidados necessários.

4.2. Pré-processamento dos MCDT’s

4.2.1. Dados Brutos

Nesta secção, apresentamos as tabelas e representações gráficas dos dados brutos de diferentes anos, antes de qualquer processo de limpeza ou normalização. O objetivo desta análise é fornecer uma visão preliminar da qualidade dos dados e identificar potenciais problemas que possam comprometer a fiabilidade das análises subsequentes.

As tabelas descritivas, apresentadas no Anexo I, fornecem estatísticas dos MCDT’s em estudo para cada uma das Unidades Funcionais de Saúde (UFS). As variáveis analisadas incluem: Idade, Tensão

Arterial diastólica (TAd), Tensão Arterial sistólica (TAs), Glicemia Capilar, Perímetro Abdominal (Perímetro AB), IMC, Lipoproteínas de Alta Densidade (HDL), Lipoproteínas de Baixa Densidade (LDL), Hba1c, Creatinina e Triglicerídeos.

As tabelas descritivas revelam uma presença de outliers, que indicam potenciais erros de medição ou entrada incorreta de dados. Foram identificados valores irreais, por exemplo na variável Glicemia Capilar, foi encontrado um valor de 994 mg/dL, nos Triglicerídeos um valor de 2164 mg/dL, no Perímetro AB um valor de 603cm ou um valor de IMC de 475.555,6 kg/m². Estes registos têm o potencial de distorcer as médias e outros indicadores estatísticos, afetando a interpretação dos resultados, destacando a necessidade de uma limpeza de dados e revisão cuidadosas, antes de prosseguir com análises detalhadas, para garantir que os resultados sejam fiáveis e representativos da realidade.

Estes problemas tornam-se evidentes, visualmente, nos boxplots apresentados na Figura 4.3 e 4.4. Nalguns gráficos, certos boxplots não são visíveis, devido à distorção da escala provocada pelos valores irrealistas, que acabam por ocultar os dados legítimos. Este fenómeno sublinha a importância de lidar com os outliers e valores anómalos, pois, caso contrário, a análise dos dados torna-se inviável e pode conduzir a interpretações incorretas.

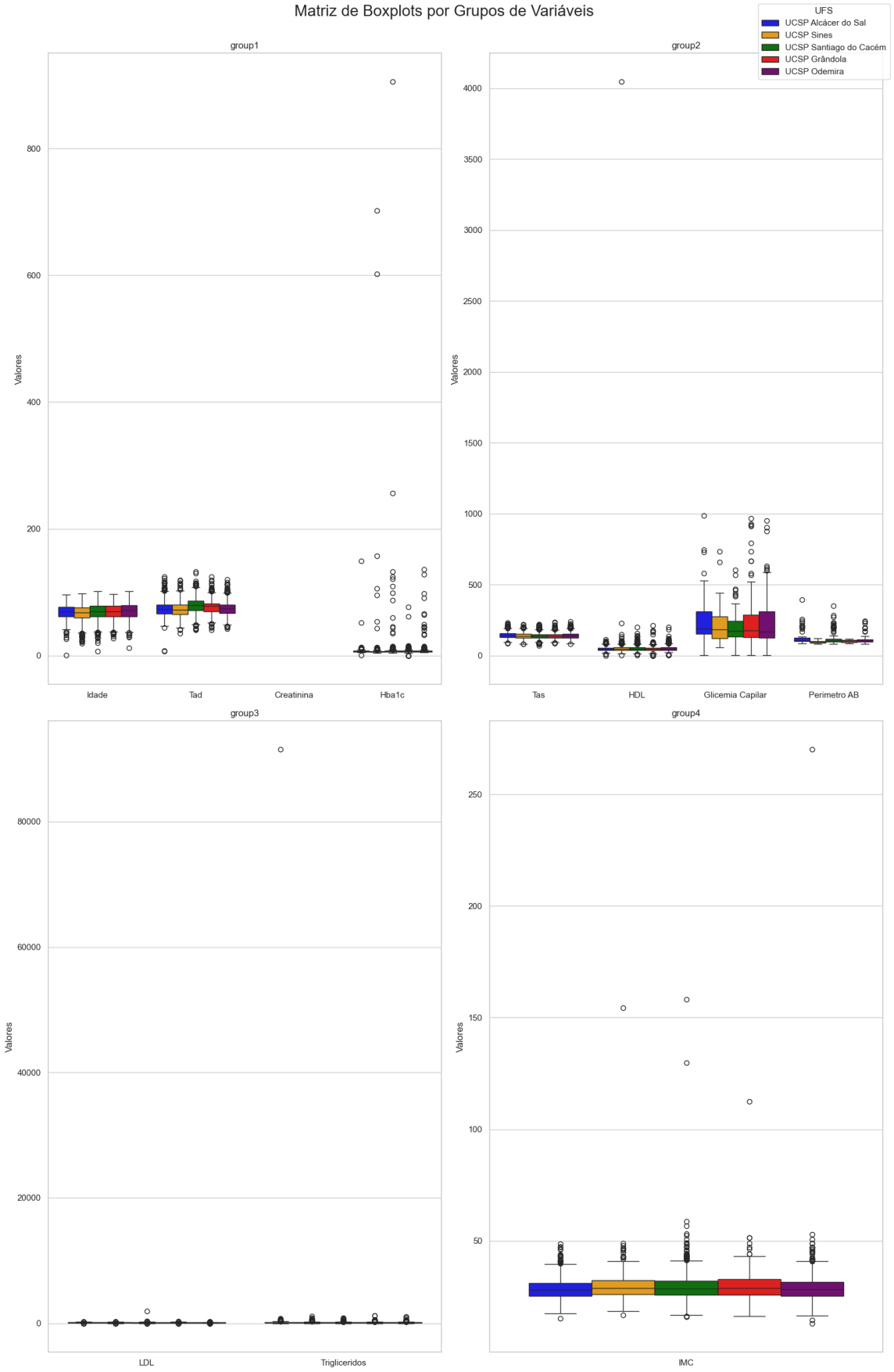


Figura 4.3 *Distribuição das Variáveis Clínicas, antes do processamento, por USF em 2017. A distribuição dos dados é ilustrada através de boxplots, onde cada cor representa uma USF específica, detalhada na legenda à direita.*

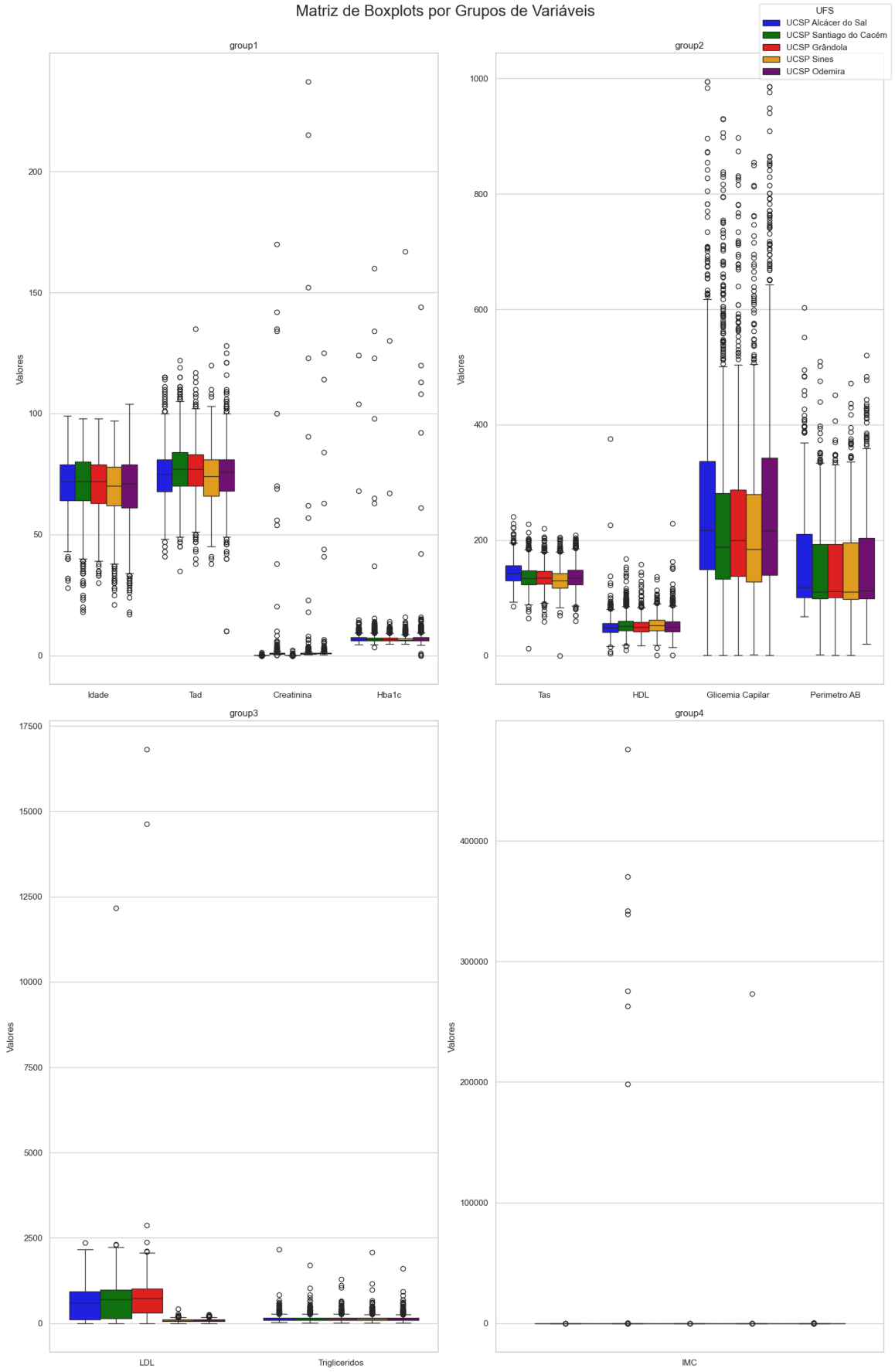


Figura 4.4 *Distribuição das Variáveis Clínicas, antes do processamento, por USF em 2023. A distribuição dos dados é ilustrada através de boxplots, onde cada cor representa uma USF específica, detalhada na legenda à direita.*

4.2.2. Dados Processados

Após a limpeza dos dados, observam-se melhorias na qualidade dos dados, evidenciadas tanto nas tabelas descritivas quanto nos boxplots. As tabelas descritivas pós-processamento apresentam valores estatísticos mais consistentes e plausíveis em comparação com as tabelas de dados brutos. As métricas centrais, como a média e a mediana refletem, agora, de forma mais precisa a realidade das variáveis analisadas.

Observamos que a glicemia capilar e o perímetro abdominal têm vindo a aumentar gradualmente desde 2017 a 2023, onde as médias destes parâmetros apresentaram um aumento percentual de 12.02% e 21.76%, respetivamente. Esta tendência sugere um agravamento no controlo da diabetes e no aumento da adiposidade abdominal, que são fatores de risco para a saúde.

Por outro lado, a média do LDL apresentou uma diminuição significativa de -12,22%. Esta redução sugere uma melhoria no perfil lipídico dos utentes, possivelmente de uma maior adesão a terapêuticas específicas para o controlo do colesterol, e/ou mudanças na alimentação.

As restantes variáveis apresentaram ligeiros aumentos e descidas que, ao longo do tempo, resultaram em pequenas mudanças percentuais, sempre abaixo dos 5%.

Tabela 4.1 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2017. Contêm o número de observações (count), média (mean), desvio padrão (std), valor mínimo (min), percentis 25%, 50% (mediana), 75% e valor máximo (max) para cada variável clínica.*

	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8397	8394	8394	8161	7957	8394	8394	8394	8394	0	8394
mean	69.278	76.014	139.821	203.080	114.200	29.072	48.738	106.685	6.833	-	135.421
std	11.908	10.576	19.054	48.197	16.261	3.891	12.230	33.820	1.230	-	69.849
min	1	50	71	59	80	13.13	10	10.6	4	-	11
25%	62	70	128.14	170	102	27.13	41	84.8	6	-	93
50%	70	76	139	198	115	29.04	47	105	6.6	-	126
75%	78	81.54	150	230	121	30.47	55	122.8	7.2	-	155
max	101	132	240	600	283	49.22	100	287.2	14.7	-	1000

Tabela 4.2 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2023. Contêm o número de observações (count), média (mean), desvio padrão (std), valor mínimo (min), percentis 25%, 50% (mediana), 75% e valor máximo (max) para cada variável clínica.*

	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	9344	9341	9341	9339	9339	9318	9341	9313	9341	8933	9341
mean	70.427	75.941	136.364	227.490	139.045	28.879	51.089	93.645	6.946	0.989	134.823
std	12.108	10.149	17.708	86.841	41.419	2.940	12.426	26.654	1.145	0.474	68.782
min	17	50	60	56	31	15.63	13	10	3.4	0.5	10
25%	63	70	126	177	107	27.38	43	84.8	6.2	0.72	93
50%	71	76	136	222.78	135.61	28.69	50	94.561	6.7934	0.9	125
75%	79	82	145	245.94	144.93	29.86	57	101.704	7.3	1.13	154
max	104	135	229	599	300	49.35	100	298	14.9	10	985

Nos histogramas, presentes na Figura 4.5, podemos observar que algumas variáveis demonstram uma distribuição mais próxima da normal, enquanto outras continuam a apresentar assimetrias.

Ao examinar os histogramas, variáveis como o IMC, TAd, e TAs apresentam distribuições que se aproximam de uma curva normal, embora ainda com algumas assimetrias. Estas variáveis possuem um pico central bem definido e uma diminuição gradual nas extremidades, características típicas de uma distribuição normal.

Por outro lado, variáveis como a glicemia capilar, triglicerídeos e creatinina continuam a apresentar distribuições assimétricas, com caudas longas à direita, indicando a presença de valores mais altos que distorcem a distribuição. Estas caudas longas podem estar associadas a *outliers*.

O HbA1c, embora demonstre uma certa estabilidade nos valores médios ao longo do tempo, também não segue uma distribuição normal. Observa-se uma concentração de valores em torno de um pico central, mas com uma cauda à direita que indica a presença de valores mais altos, característicos de alguns utentes que não estão a conseguir controlar adequadamente a glicemia.

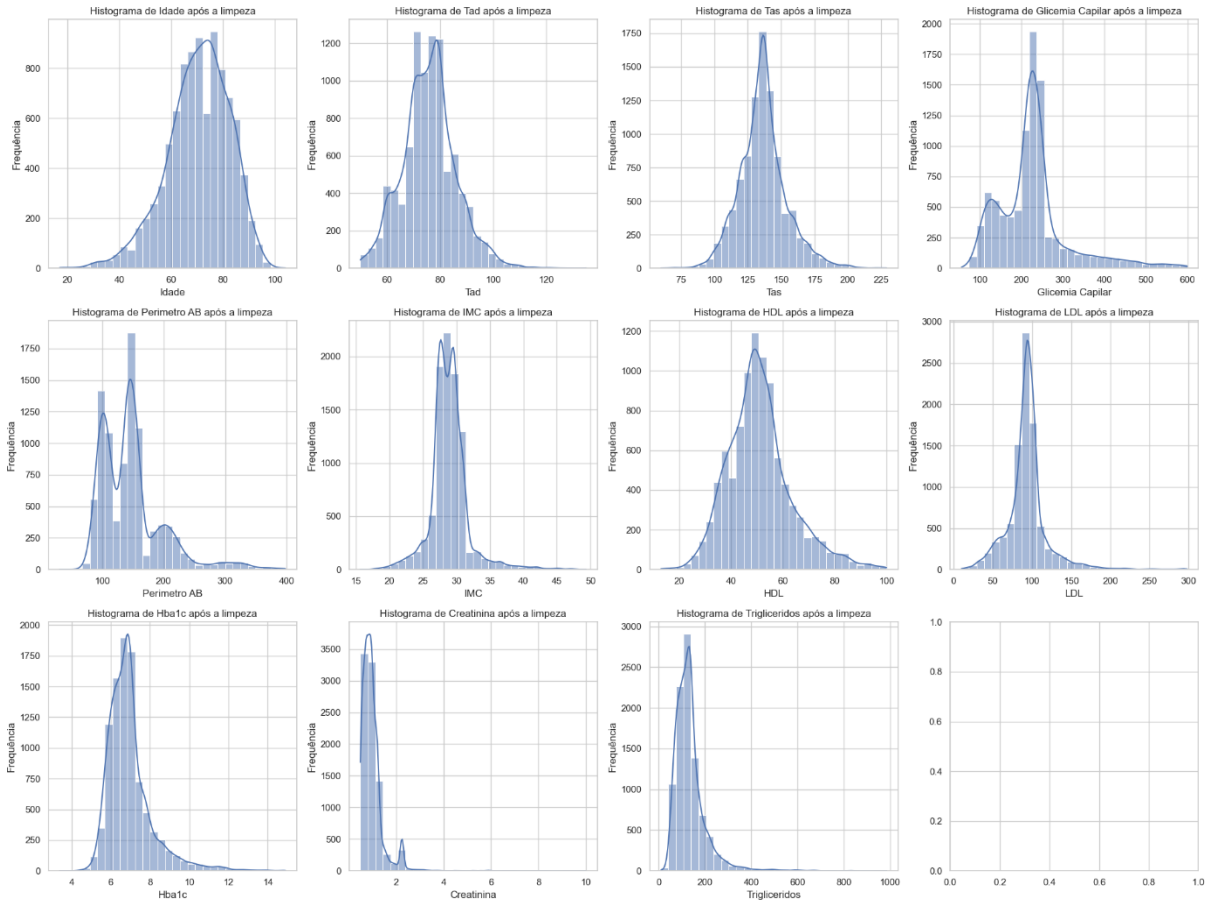


Figura 4.5 *Distribuição das Variáveis Clínicas Após Limpeza de Dados em 2023. Esta matriz de histogramas ilustra a distribuição de várias variáveis clínicas coletadas após o processo de limpeza de dados.*

Analisando mais detalhadamente o histograma do perímetro abdominal, é possível identificar uma distribuição que pode ser considerada trimodal ou multimodal. Os dois picos mais altos sugerem a existência de duas subpopulações distintas com diferentes características de adiposidade abdominal. Além destes dois picos, observa-se uma "onda" mais suave, seguida por uma cauda que se prolonga para a direita, indicando a presença de indivíduos com perímetros abdominais elevados.

Esta complexidade na distribuição aponta para uma evolução ao longo do tempo, onde as subpopulações que antes estavam mais agregadas em 2017 começaram a dispersar nos anos seguintes. Este fenómeno pode ser resultado de variações nos perfis populacionais atendidos pelas UFS ou até mesmo a evolução natural da doença entre os utentes diabéticos.

A dispersão ao longo do tempo indica que, enquanto alguns indivíduos conseguiram manter ou reduzir o perímetro abdominal, outros continuaram a aumentar, contribuindo para a formação destes múltiplos picos na distribuição. Esta tendência de dispersão ressalta a importância de estratégias

personalizadas de intervenção, pois diferentes subgrupos podem necessitar de abordagens distintas para o controlo da obesidade abdominal e das complicações associadas.

4.2.3. Segmentação do Perfil

A partir das análises anteriores, surgiu o interesse de segmentar a população diabética e tentar perceber quais os fatores diferenciais que iriam afetar os subgrupos. Efetuou-se a construção das Tabelas 4.3 e 4.4 dos fatores para tentar perceber a distribuição dos diabéticos por cada fator. A segmentação resultou em quatro *clusters* distintos, apresentados na Figura 4.6.

Tabela 4.3 *Características demográficas e de IMC da população diabética em 2023.*

Característica	N (%)	
Sexo		
Masculino	4632 (49.57%)	
Feminino	4712 (50.43%)	
Total		9344
Nacionalidade		
Portuguesa	9026 (96.60%)	
Outras	318 (3.40%)	
Total		9344
Idade		
<40	135 (1.44%)	
40-55	906 (9.70%)	
56-70	3336 (35.70%)	
≥71	4967 (53.16%)	
Total		9344
IMC (kg/m^2)		
<25	595	
25-28	2592	
29-34	5728	
≥35	403	
Total		9318

Tabela 4.4 *Características clínicas da população diabética em 2023.*

Característica	N	
HbA1c		
<7	5812	
≥7	3529	
Total		9341
HDL		
<50	4638	
≥50	4703	
Total		9341
LDL		
<100	6268	
≥100	3045	
Total		9313
TRIG		
<150	6817	
≥150	2524	
Total		9341
Perímetro Abdominal		
<100	1366	
100-200	6678	
201-300	1056	
≥301	230	
Total		9330

Na população total do Litoral Alentejano, os diabéticos do sexo masculino representam 4,70% enquanto as mulheres são 4,78%, evidenciando uma ligeira predominância feminina. Quanto à nacionalidade, 9,15% da população total corresponde a portugueses diabéticos, e apenas 0,32% são estrangeiros.

A distribuição por faixas etárias revela que a maioria dos diabéticos pertence aos grupos etários mais avançados, com 53,16% dos casos em indivíduos com 71 anos ou mais. Apenas 1,44% dos diabéticos têm menos de 40 anos, o que indica uma maior prevalência da diabetes entre os mais idosos.

No que respeita ao IMC, a maioria dos diabéticos (61,5%) apresenta um IMC entre 29 e 34, indicando uma alta prevalência de sobrepeso ou obesidade nesta população. Apenas 6,4% têm um IMC inferior a 25 kg/m². Relativamente ao perímetro abdominal, a maioria dos diabéticos (71,5%) apresenta medidas entre 100 e 200 cm, sugerindo a presença de fatores de risco metabólicos.

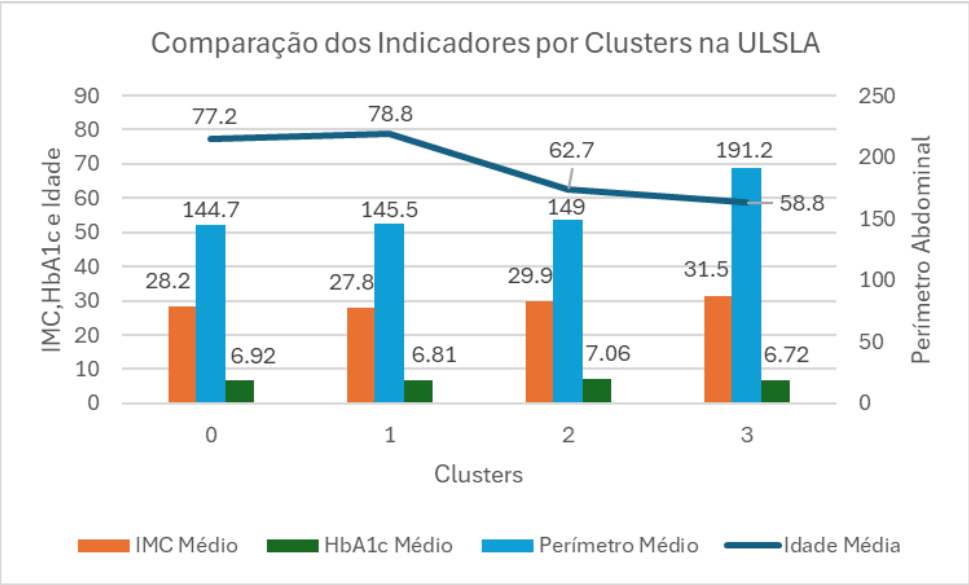


Figura 4.6 *Comparação dos Indicadores Médios por Clusters na ULSLA em 2023. As barras representam o IMC, HbA1c, e o perímetro abdominal, enquanto a linha azul mostra a idade média dos pacientes em cada cluster.*

O *cluster* 0 é composto principalmente por idosos de 77,2 anos é o subgrupo onde as comorbidades são mais prevalentes, destacando-se a Glomerulonefrite (100%), Retinopatia (11,59%), Doença Vascular Cardíaca (9,44%), e Aterosclerose (3,86%). Além disso, uma elevada percentagem destes utentes enfrenta insuficiência económica (39,1%), o que pode dificultar o acesso a cuidados de saúde necessários. Este *cluster* representa um grupo de risco elevado devido à prevalência extremamente alta de Glomerulonefrite (U88), em 100% dos 233 utentes, pode indicar a necessidade de especial atenção para a função renal. A alta incidência de outras comorbilidades, aliada com uma percentagem considerável de insuficiência económica, torna este grupo particularmente vulnerável, exigindo uma abordagem integrada de cuidados de saúde que aborde tanto as necessidades clínicas quanto as barreiras socioeconómicas para impedir a possibilidade de desenvolver complicações adicionais.

O *cluster* 1 é composto por 4280 indivíduos, com uma idade média de 78,8 anos e um IMC médio de 27,8 kg/m², indicando sobrepeso. Embora a prevalência de comorbidades seja menor do que no *Cluster* 0, ainda assim existem riscos significativos, como a Retinopatia (8,88%), Doença Vascular Cardíaca (11,15%) e com Aterosclerose (5,86%). A insuficiência económica é maior neste grupo (46,7%). Apesar da idade similar ao *Cluster* 0, o *Cluster* 1 distingue-se pela ausência de U88 entre os seus utentes, o que pode implicar diferenças no perfil de risco e necessidades de cuidados. A prevalência de outras comorbilidades e a elevada taxa de insuficiência económica sugerem que, embora não enfrentem os desafios da U88, estes utentes ainda requerem uma gestão cuidadosa de outras condições crónicas. As intervenções devem focar-se em melhorar o controlo dessas comorbilidades e em mitigar os impactos das barreiras económicas.

O *cluster* 2 é composto de 4825 indivíduos, com idade média de 62,7 anos e um IMC médio de 29,9 kg/m², indicando obesidade. Este grupo apresenta uma baixa prevalência de comorbidades, incluindo a Aterosclerose (1,76%), Doença Vascular Cardíaca (1,82%) e Retinopatia (1,66%). A insuficiência económica afeta apenas 18,2% dos utentes deste *cluster*. Este grupo apresenta um risco metabólico elevado devido à obesidade e ao controlo glicémico insatisfatório. A implementação de programas de controlo de peso e de educação em diabetes seria benéfica para reduzir o risco de complicações futuras, uma vez que a prevalência das comorbidades graves ainda é baixa. Além disso, a educação destes utentes pode ter um impacto positivo na sua qualidade de vida, dado que ainda estão

numa idade em que a modificação dos hábitos pode trazer benefícios substanciais para a sua vida futura. De acordo Adachi et al. [78], intervenções focadas na educação e na mudança de estilo de vida em indivíduos de meia-idade podem melhorar o controlo glicémico, reduzir o peso corporal e, consequentemente, diminuir o risco de complicações a longo prazo. Devido a uma condição económica mais satisfatória, relativamente aos *clusters* anteriores, este grupo pode ter mais acesso a cuidados de saúde, facilitando a implementação de intervenções preventivas.

O último *cluster* com 6 utentes, é composto pelos indivíduos mais jovens, com uma idade média de 58,8 anos, mas que apresentam obesidade severa (IMC médio de 31,5 kg/m²) e o perímetro abdominal médio elevado (191,2 cm). Todos os indivíduos neste *cluster* estão desempregados. Apesar da menor incidência, estes utentes com obesidade severa estão em risco de desenvolver comorbidades graves no futuro. Intervenções focadas na perda de peso e na prevenção de comorbidades são determinantes, tanto para este grupo como para o subgrupo anterior, para evitar complicações que possam resultar em incapacidades ou mortalidade precoce.

Em síntese, adotar uma abordagem diferenciada e adaptada às necessidades específicas de cada *cluster* pode não só melhorar a qualidade de vida dos utentes, mas também reduzir o impacto da diabetes na saúde pública a longo prazo. Integrar cuidados de saúde com suporte socioeconómico para os *clusters* mais vulneráveis, juntamente com a promoção da educação e prevenção entre os grupos mais jovens, constitui uma estratégia abrangente e eficaz para a gestão da diabetes na população.

4.3. Análise Espácio-temporal

Esta análise espácio-temporal tem como objetivo principal avaliar a distribuição geográfica e a densidade de pacientes com T90 ao longo dos anos em análise. Através de métodos de análise espacial, como a estimativa de densidade *Kernel*, distribuição direcional, a ANN e K-Rippley foi possível identificar padrões e tendências na ocorrência de casos. Este estudo abrange os anos de 2017-2023, permitindo uma comparação temporal que contribui para a compreensão da evolução dos casos no espaço e no tempo.

4.3.1. Distribuição Geográfica dos Utentes

Nesta secção, analisamos a evolução temporal e espacial da distribuição dos pacientes diagnosticados com T90 no Litoral Alentejano, durante os anos de 2017 a 2023. A distribuição espacial dos pacientes foi mapeada anualmente, representada na Figura 4.7 permitindo identificar padrões de concentração geográfica e possíveis mudanças ao longo do tempo.

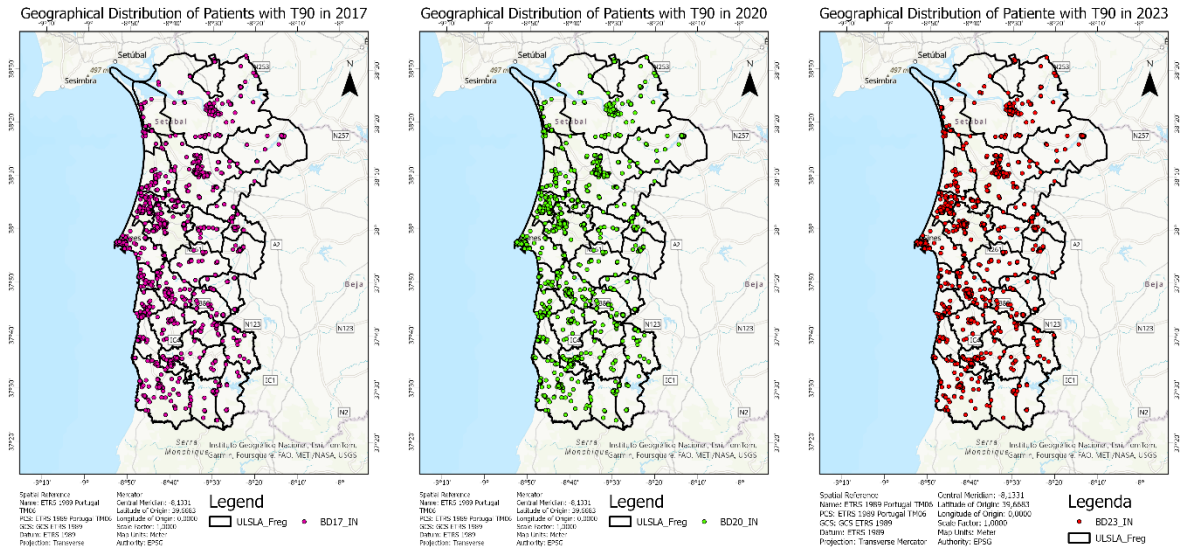


Figura 4.7 *Distribuição geográfica de utentes com T90 nos anos 2017, 2020 e 2023. A distribuição para os anos intermédios está disponível no Anexo II.*

Em 2017, observa-se uma distribuição ampla dos pacientes, com uma concentração significativa nos municípios de Alcácer do Sal, Sines, Grândola e Santiago do Cacém. As áreas de maior densidade localizam-se principalmente ao longo da faixa costeira. No município de Odemira, embora haja um número considerável de casos, os pacientes estão mais dispersos.

Nos anos de 2018 a 2022, a distribuição dos casos mantém-se relativamente constante em termos de padrão espacial. Sines, Grândola e Santiago do Cacém continuam a ser os municípios com maior concentração de casos, refletindo uma tendência persistente de aglomeração nas áreas costeiras e urbanas destes municípios. Em Odemira e Alcácer do Sal, a distribuição permanece dispersa, mas com sinais de aumento gradual na densidade de casos, especialmente em zonas costeiras. As variações de ano para ano são mínimas, com ligeiras intensificações nos *clusters* existentes, sem mudanças significativas na localização dos principais hotspots.

Em 2023, observa-se uma consolidação dos padrões de distribuição identificados nos anos anteriores. Sines, Grândola e Santiago do Cacém continuam a ser os municípios com maior concentração de casos, com hotspots bem estabelecidos e intensificados ao longo da faixa costeira. Odemira, embora mantenha uma distribuição menos densa, agora revela *clusters* mais definidos, especialmente nas áreas costeiras. Alcácer do Sal, por sua vez, mostra uma tendência de aumento na densidade de casos em algumas áreas específicas, sugerindo uma mudança gradual no padrão de distribuição dos pacientes.

Em resumo, pode verificar-se que ao longo dos anos existe uma tendência de aumento na concentração de casos nos municípios de Alcácer do Sal, Sines, Grândola e Santiago do Cacém, juntamente com o desenvolvimento progressivo de *clusters* em Odemira.

4.3.2. Análise do padrão de distribuição espacial de utentes com T90

A utilização da Estimativa de Densidade *Kernel* é importante para uma análise mais detalhada da distribuição espacial dos utentes, permitindo identificar áreas de concentração de casos que poderiam passar despercebidas com a análise de dados pontuais. Esta técnica gera um mapa de calor, destacando os focos, e facilitando a identificação de padrões evolutivos ao longo dos anos. Assim, ao analisar as Figura 4.8 para os anos de 2017, 2020 e 2023 nos municípios de Sines, Grândola, Odemira, Alcácer do

Sal e Santiago do Cacém, é possível compreender melhor a dinâmica da distribuição dos pacientes, em que a cor vermelha está associada às situações mais críticas e que requerem mais atenção.

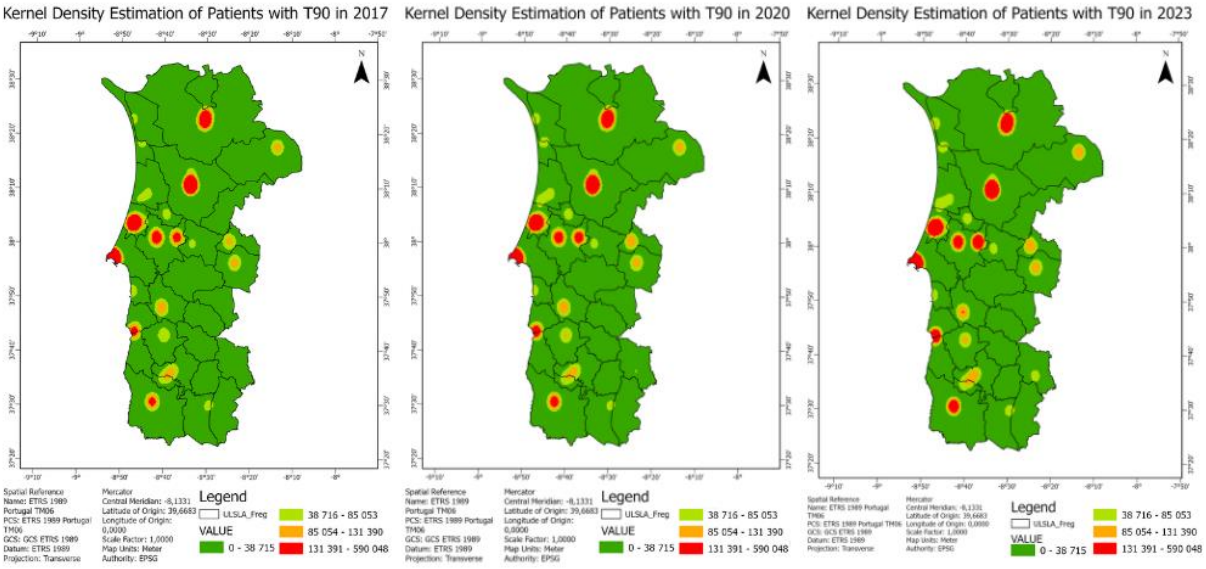


Figura 4.8 *Densidade de Kernel do número de Pacientes com T90 nos Anos de 2017, 2020 e 2023. A cor vermelha indica maior densidade de utentes com DMT2.*

Em 2017, os padrões de densidade revelam áreas de maior concentração de pacientes em regiões específicas, especialmente em torno de Sines e Santiago do Cacém, com focos adicionais em Grândola e Alcácer do Sal. Embora Odemira também apresente áreas com alguma densidade, esta é mais dispersa. As áreas de alta densidade indicam que, apesar de os pacientes estarem relativamente espalhados, há uma concentração nos centros urbanos.

Entre 2018 e 2022, a distribuição dos pontos de alta densidade mantém-se relativamente estável, com pequenas variações na intensidade dos focos de concentração. O núcleo central em Sines e Santiago do Cacém continua a ser uma das áreas com maior concentração de casos, enquanto Grândola, Alcácer do Sal e Odemira mantêm padrões semelhantes aos de 2017, com algumas áreas a apresentar uma densidade ligeiramente mais intensa. Esta estabilidade sugere uma consistência nos padrões de distribuição ao longo dos anos, sem alterações na localização dos principais focos.

Em 2023, observa-se uma ligeira intensificação da concentração dos pacientes em algumas áreas, particularmente em Sines e Santiago do Cacém. Esta mudança sugere que, embora a distribuição geral se mantenha consistente, há uma tendência para uma maior concentração de casos nas regiões anteriormente mencionadas.

A análise da Estimativa de Densidade *Kernel* ao longo dos anos revela que, apesar da distribuição dos pacientes ter permanecido relativamente estável, há uma tendência gradual de aumento na concentração dos casos em áreas específicas, principalmente em Sines e Santiago do Cacém. Este padrão, consistente com a análise anterior, sugere a necessidade de intervenções nessas regiões para mitigar o aparecimento de novos casos com esta patologia.

A análise da Distribuição Direcional dos pacientes com T90, na Figura 4.9, entre 2017 e 2023 revela uma consistência impressionante na orientação e extensão da elipse de distribuição ao longo dos anos. Em cada um destes anos, a elipse mantém uma orientação estável no sentido noroeste-sudeste, abrangendo todos os municípios da região. A forma e a direção da elipse indicam que os padrões espaciais de concentração de casos de T90 permanecem constantes, sem mudanças significativas na

distribuição dos pacientes ao longo do tempo. Esta estabilidade sugere que as áreas de maior concentração de pacientes, e consequentemente os focos de maior densidade de casos, se mantêm relativamente inalterados ao longo dos sete anos analisados. Estes resultados destacam a persistência dos padrões espaciais na região, o que pode ser relevante para a alocação de recursos.

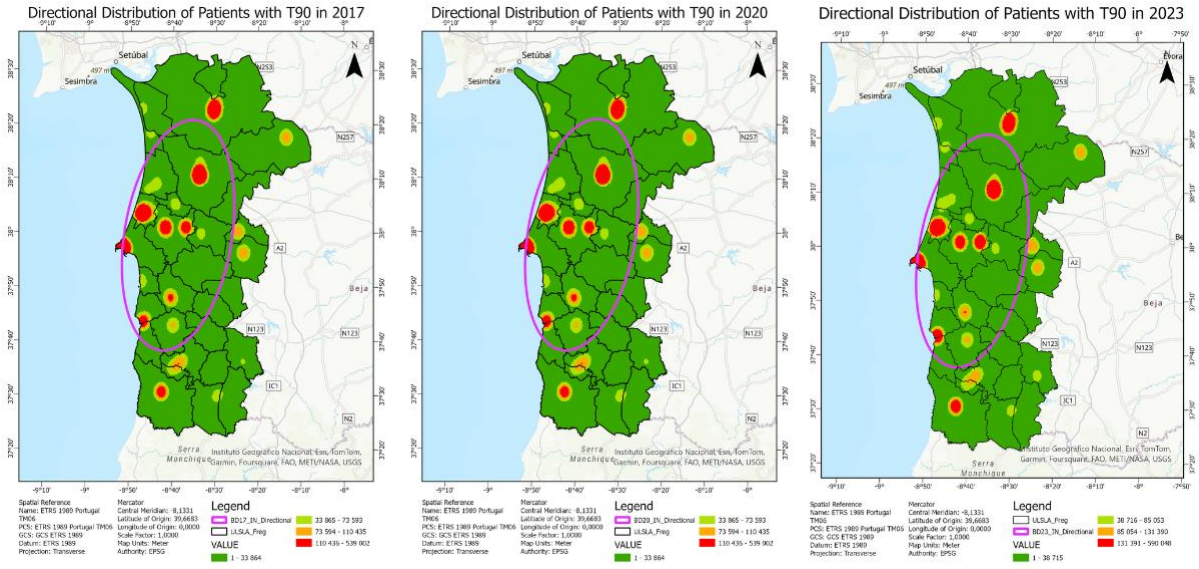


Figura 4.9 *Distribuição Direcional de Pacientes com T90 em LA em 2017, 2020 e 2023. As áreas dentro das elipses indicam regiões de interesse onde há maior concentração de casos.*

Na análise do ANN realizada para os anos de 2017 a 2023, obteve-se um relatório para cada ano, como mostrado na Figura 4.10, indicando o tipo de distribuição dos utentes com T90. Como os valores de *Nearest Neighbor Ratio* (NNR), *p-value* e *z-score* foi criada a Tabela 4.5.

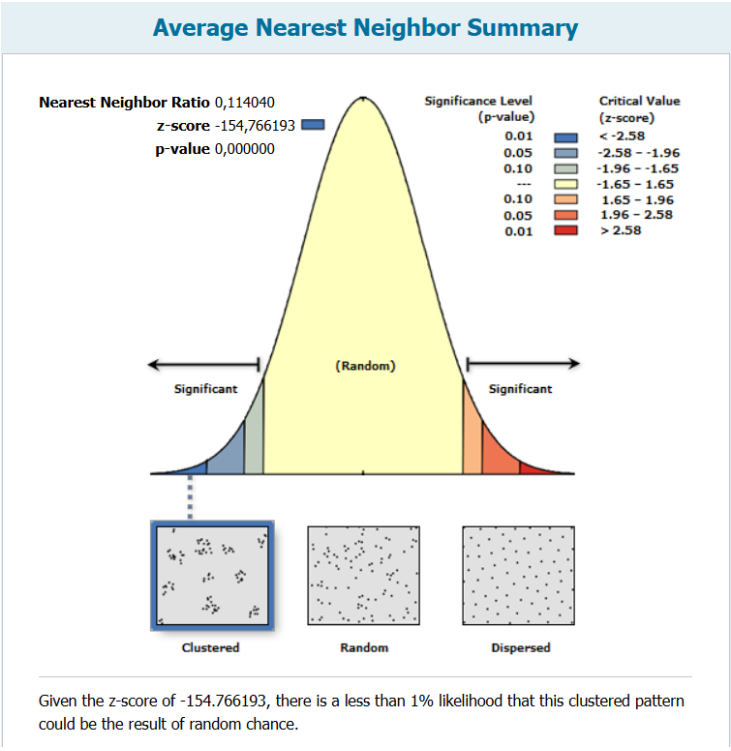


Figura 4.10 *Relatório ANN para Pacientes com T90 de 2017, obtido através do ArcGIS Pro.*

Tabela 4.5 *Análise dos valores de NNR, z-score e p-value para Pacientes com T90 de 2017 a 2023.*

	2017	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**
NNR	0.114	0.111	0.110	0.109	0.101	0.102	0.094
z-score	-154.766	-156.321	-156.916	-157.199	-161.050	-162.965	-166.683
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Esta ferramenta revela um padrão consistente de concentração significativa, anteriormente verificada, dos pacientes com T90 na região em estudo. Os valores de NNR apresentam variações ligeiras ao longo dos anos, começando em 0,114 em 2017 e diminuindo progressivamente para 0,094 em 2023. Esta diminuição indica uma tendência de aumento na densidade dos pacientes, sugerindo que os casos de T90 estão cada vez mais concentrados em áreas específicas.

Os valores de *z-score*, que variam entre -154,766 e -166,683, reforçam a significância estatística deste padrão de aglomeração, com *p-values* de 0,00 em todos os anos, indicando uma probabilidade inferior a 1% de que este padrão seja fruto do acaso. A Figura 4.10 ilustra claramente esta distribuição, demonstrando que o padrão observado se desvia de uma distribuição aleatória, confirmando a forte aglomeração de pacientes ao longo dos anos analisados.

Ao analisar o gráfico obtido para a função *K Rippley* para o ano de 2023, presente na Figura 4.11, podem ser feitas as seguintes observações:

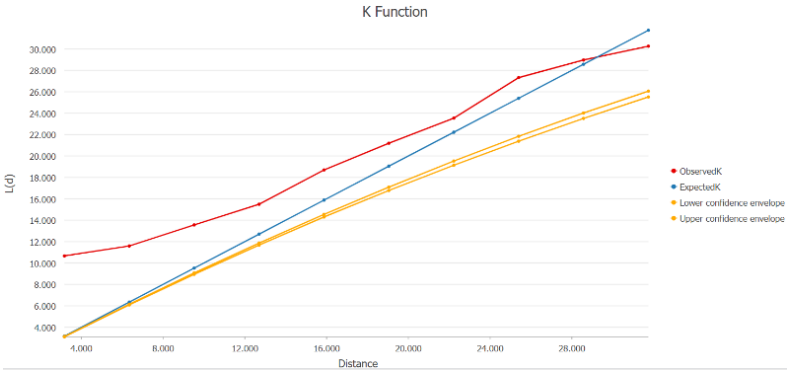


Figura 4.11 *Análise da Função K para Distribuição Espacial. Os gráficos dos anos 2017-2022 encontram-se no Anexo II.*

Em todos os anos analisados, a linha *ObservedK* está consistentemente acima da linha *ExpectedK* e fora dos envelopes de confiança, indicando um *clustering* de casos de diabetes tipo 2 em cada ano, como já havia sido confirmado anteriormente. O *clustering* é mais intenso em distâncias mais curtas (até 22 km) em todos os anos, sugerindo uma alta concentração de casos em áreas específicas que se mantêm constantes ao longo do tempo.

O grau de *clustering* não apresenta variações consideráveis de ano para ano, o que indica uma persistência nas áreas de alto risco. Não há evidências de que o grau de *clustering* esteja a aumentar ou a diminuir de forma substancial ao longo do período analisado. Embora não existam mudanças drásticas, qualquer pequena variação nos padrões pode ser atribuída a fatores locais específicos.

4.3.3. Emerging Hot Spots

A análise de *hot spots* emergentes foi realizada utilizando a base de dados agregada, BD_IN_Merge, que reúne os casos de diabetes tipo 2 ao longo dos sete anos. Para garantir a contextualização espacial, foi aplicada uma máscara por município, com foco na região do Litoral Alentejano. O resultado obtido encontra-se ilustrado na Figura 4.12.

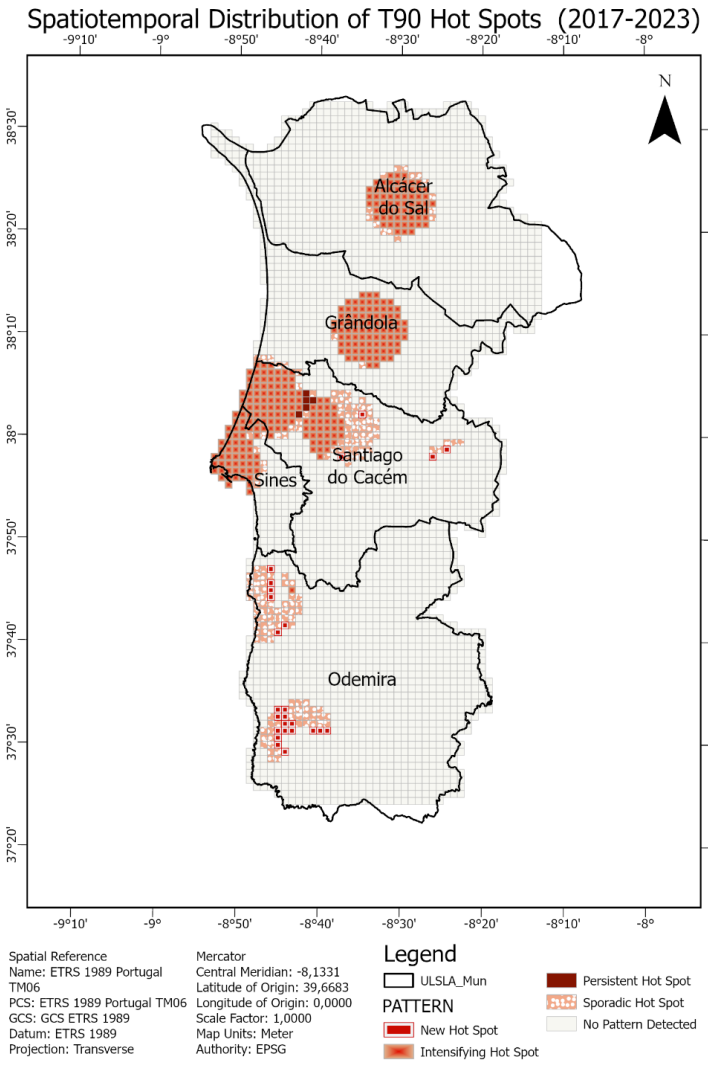


Figura 4.12 *Distribuição Espácio-temporal dos Hot Spots de T90 do LA (2017-2023). O mapa ilustra a localização dos hot spots persistentes, emergentes e esporádicos de T90 ao longo do período de sete anos.*

As áreas com maior intensidade de *hot spots* estão localizadas principalmente na região oeste da área de estudo, como já havia sido observado anteriormente através da ferramenta de Distribuição Direcional. A intensidade dos *hot spots* varia, refletindo diferentes níveis de concentração de casos de DMT2 ao longo dos anos.

Na região estudada foi possível caracterizar as diversas áreas em cinco categorias distintas;

Focos Persistentes: (■) Áreas que foram consistentemente *hot spots* estatisticamente significativos durante 90% do intervalo de tempo, sem nenhuma diminuição na intensidade do agrupamento ao longo do período analisado. As únicas áreas classificadas como persistentes encontram-se na União das freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra, no município de Santiago do Cacém.

Focos Intensos: (■) Áreas que foram *hot spots* estatisticamente significativos durante 90% do intervalo de tempo, incluindo o último ano. Além disso, a intensidade do agrupamento de casos elevados aumentou ao longo do tempo, aumento esse que é estatisticamente significativo. Estas áreas estão localizadas nos municípios de Alcácer do Sal, Grândola, Sines e Santiago do Cacém.

Focos Emergentes: ( ) Áreas que se tornaram *hot spots* estatisticamente significativos apenas no último intervalo de tempo, ou seja, em 2023, e que anteriormente nunca foram classificadas como *hot spots* significativos. Freguesias de Odemira, principalmente, em São Teotónio e Vila Nova de Milfontes, e no município de Santiago do Cacém, mostram a necessidade de monitorização e intervenções rápidas para evitar que estes focos se tornem persistentes, dado que são focos mais recentes.

Focos Esporádicos: ( ) Áreas que foram *hot spots* estatisticamente significativos para o último intervalo de tempo (2023), com um histórico de serem pontos quentes periódicos. Ao longo dos sete anos nunca foram identificados como pontos frios estatisticamente significativos, e em menos de 90% do intervalo de tempo foram *hot spots* significativos. Estes focos encontram-se ao redor dos focos intensos nos municípios de Alcácer do Sal, Grândola e Sines. Em Odemira, encontramos um conjunto de focos esporádicos associados a focos emergentes, e em Santiago do Cacém, surgem focos esporádicos numa região que combina focos emergentes e intensos.

Áreas sem padrão definido: ( ) Por fim, foram identificadas zonas que não apresentam qualquer tipo de padrão consistente, não se enquadrando em nenhuma das categorias mencionadas anteriormente, nem em padrões de pontos frios.

4.4. Análise de Autocorrelação Espacial

Nesta secção são apresentados os resultados da análise exploratória de dados realizada para compreender as correlações entre diversos determinantes sociais de saúde (DSS) e a prevalência de diabetes tipo 2 na região do Litoral Alentejano. Como mencionado na secção 4.2.3, foi utilizado o software GeoDa para avaliar a autocorrelação espacial das variáveis em estudo, e produziram-se mapas e gráficos de dispersão, permitindo identificar como cada variável se comporta na sua vizinhança e de que modo a mesma se correlaciona com o número de casos de diabetes por freguesia.

4.4.1. Variáveis Explicativas

População total e por sexo

Os dados referentes à população total, da área em estudo são apresentados em valores absolutos, para os géneros são exibidos em valores relativos, expressos em percentagem. Quanto mais escura for a cor representada, mais elevada será a percentagem.

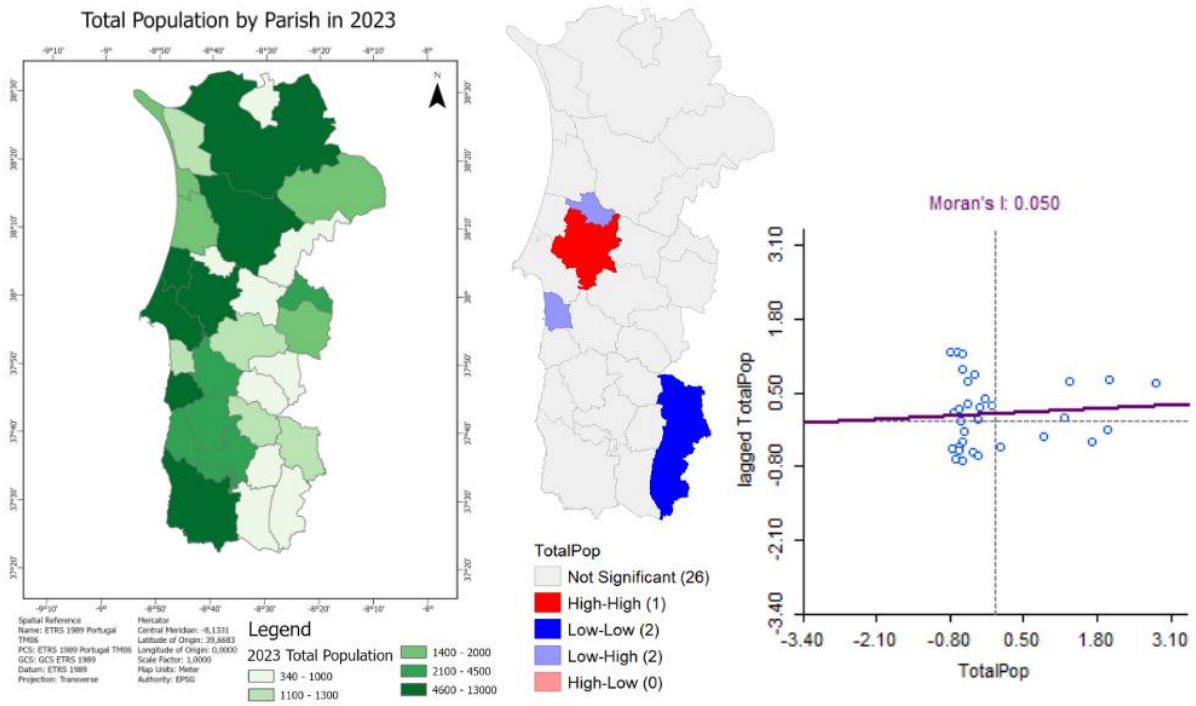


Figura 4.13 *Análise Espacial da População Total por Freguesia e Autocorrelação Espacial em 2023. Da esquerda para a direita, o mapa mostra a distribuição da população por freguesia, indica a significância espacial de clusters e o índice de Moran's I.*

A análise da distribuição da população total por freguesia na região do Litoral Alentejano, na Figura 4.13, revela que a União das Freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra é a única área classificada como *High-High*, indicando uma alta concentração populacional. Esta freguesia sobressai em termos de residentes quando comparada com outras freguesias da região.

Em contraste, as freguesias Santa Clara-a-Velha e São Martinho das Amoreiras, ambas no concelho de Odemira, estão classificadas como *Low-Low*, o que indica áreas com menor população. Adicionalmente, as freguesias de Porto Covo, no concelho de Sines, e São Francisco da Serra, no concelho de Santiago do Cacém, exibem um padrão *Low-High*, sugerindo baixo número de habitantes, mas rodeadas por áreas com mais população.

O índice de Moran's I para a população total é de 0,050, refletindo uma ligeira tendência de autocorrelação espacial. Este valor sugere que, embora existam freguesias com concentrações populacionais distintas, como a União das Freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra, a distribuição geral da população no Litoral Alentejano não apresenta um padrão de agrupamento espacial acentuado, sendo bastante diversificada entre as diferentes freguesias.

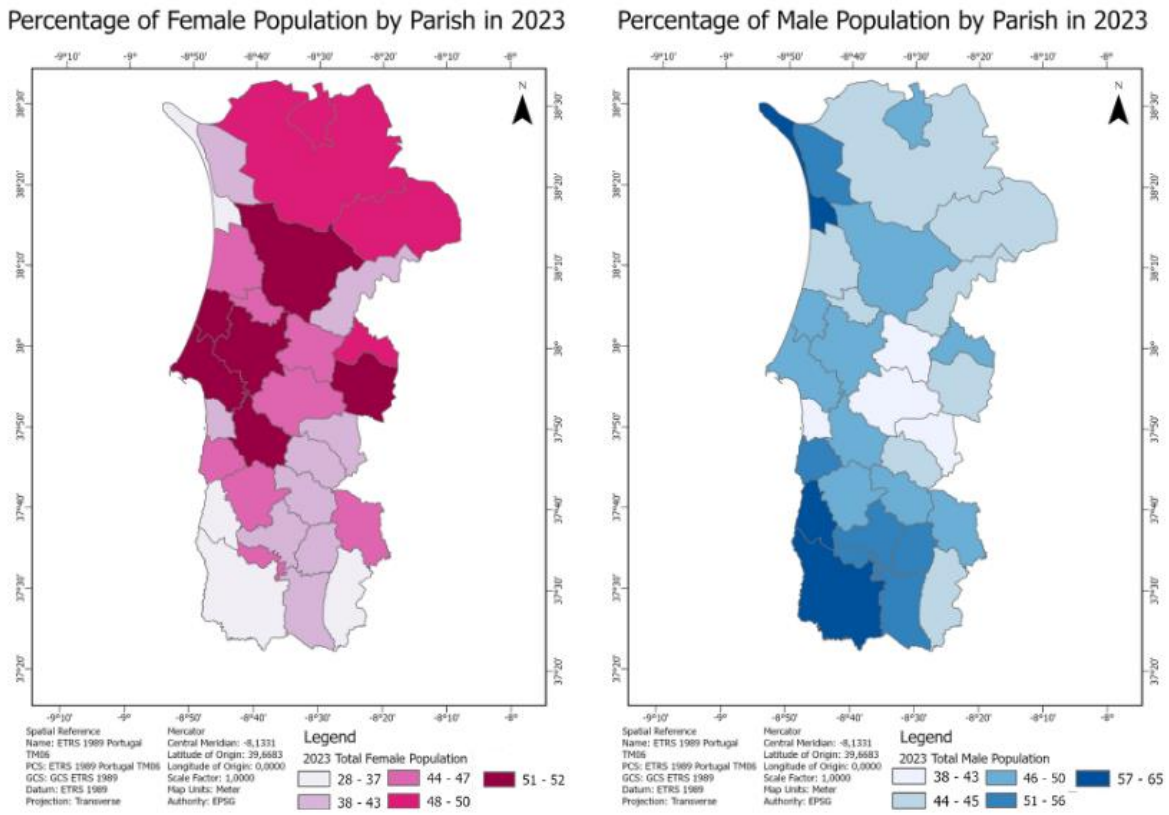


Figura 4.14 *Distribuição Percentual da População por Sexo por Freguesia em 2023. Os mapas ilustram a população feminina (à esquerda) e masculina (à direita).*

Os mapas de distribuição por género, na Figura 4.14, revelam padrões complementares na população da região do Litoral Alentejano. Observa-se que os dois géneros apresentam uma distribuição espacial distinta. Há uma maior percentagem de indivíduos do género masculino nas freguesias do litoral a sul da região, com exceção da freguesia da Comporta. Por outro lado, a população feminina tende a concentrar-se mais na zona central da região, destacando-se freguesias como a União das Freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra, Santo André e Sines.

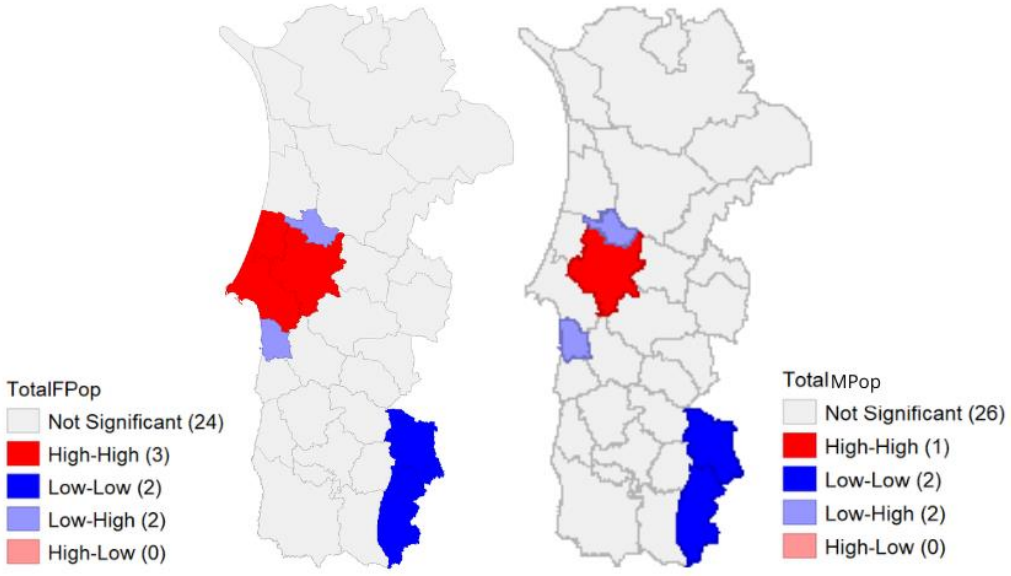


Figura 4.15 *Análise Espacial de Significância para População Feminina e Masculina em 2023. O mapa à esquerda detalha a população feminina e o mapa à direita a masculina.*

A União das Freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra, assim como as freguesias de Santo André e Sines, destacam-se como áreas High-High para ambos os géneros na Figura 4.15. Estas freguesias apresentam uma elevada densidade populacional, com grande concentração de tanto de homens como de mulheres. Este padrão de alta concentração sugere que estas freguesias funcionam como polos de atração populacional, possivelmente devido à presença de atividades económicas que proporcionam maior taxa de empregabilidade ou a outras infraestruturas, como infantários e escolas, que atraem ambos os géneros. Em contraste, as freguesias de Santa Clara-a-Velha e São Martinho das Amoreiras aparecem como *Low-Low*, caracterizadas por baixa densidade populacional e menor atratividade para ambos os géneros, possivelmente refletindo um menor dinamismo económico ou outras condições que influenciam a distribuição demográfica.

Porto Covo e São Francisco da Serra são identificadas como *Low-High*, tanto na população total quanto na distribuição de géneros, sendo vizinhas de freguesias com maior concentração populacional. Isso pode indicar condicionantes locais nas condições sociais e económicas.

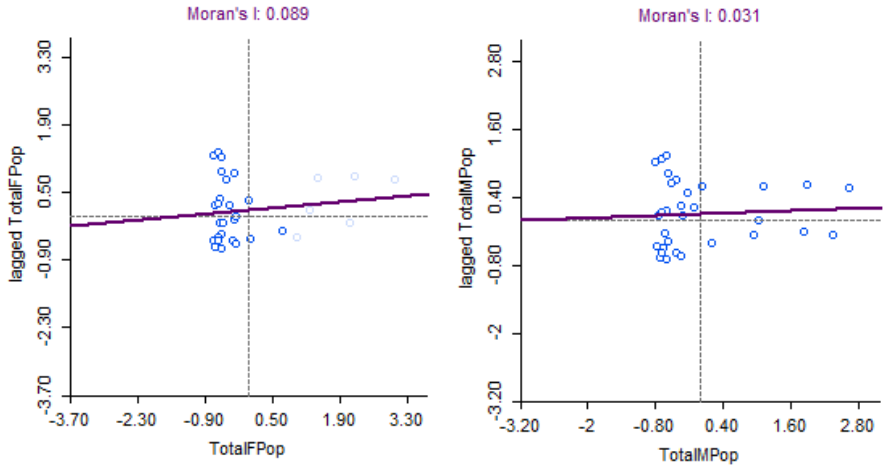


Figura 4.16 *Análise de Moran's I para Autocorrelação Espacial da População Feminina e Masculina. Os gráficos mostram o índice de Moran's I para a autocorrelação espacial das populações feminina (à esquerda) e masculina (à direita).*

Na Figura 4.16, o valor de Moran's I é de 0,089 para a população feminina e 0,031 para a masculina, indicando uma autocorrelação espacial fraca em ambos os géneros, embora ligeiramente mais acentuada na população feminina. Estes resultados sugerem que, apesar de haver áreas com maior ou menor concentração populacional, tanto de homens quanto de mulheres, a distribuição geral não apresenta um padrão de agrupamento espacial claramente definido.

Faixa Etária

A análise demográfica do Litoral Alentejano também permite compreender a distribuição da população por diferentes faixas etárias na Figura 4.17, identificando as áreas com maior concentração de jovens, adultos e idosos.

Figura 4.17 *Distribuição Percentual da População por Faixas Etárias de 2023, da esquerda para a direita e de cima para baixo:0-14, 15-29, 30-44, 45-59 e 60+ anos.*

A distribuição da população jovem concentra-se em áreas específicas, com destaque para as freguesias costeiras como Sines e Vila Nova de Santo André, que apresentam uma maior percentagem de crianças e adolescentes. Na faixa etária dos 15-29 anos, as freguesias do litoral continuam a mostrar uma elevada concentração populacional, especialmente em Sines, São Teotónio e Vila Nova de Milfontes, provavelmente devido às maiores oportunidades de emprego e educação que atraem e retêm esta faixa etária.

A população entre os 30 e os 44 anos encontra-se mais concentrada nas freguesias de Carvalhal e Odemira. O grupo etário dos 45-59 anos apresenta uma distribuição um pouco mais dispersa, com

presença na União das freguesias de Grândola e Santa Margarida da Serra, Santa Clara-a-Velha e Luzianes-Gare, e Sines, sugerindo possíveis sinais de estabilização económica.

Por fim, a análise da população com 60 anos ou mais revela uma forte presença nas freguesias interiores, como Torrão, Alvalade, Vale de Santiago e Cercal. Este padrão sugere que as áreas mais afastadas do litoral, apesar de menos atrativas para as faixas etárias mais jovens, são onde a população idosa é mais prevalente, possivelmente devido a desempenho de atividades agrícolas ou de acomodação deste grupo etário ao longo do tempo.

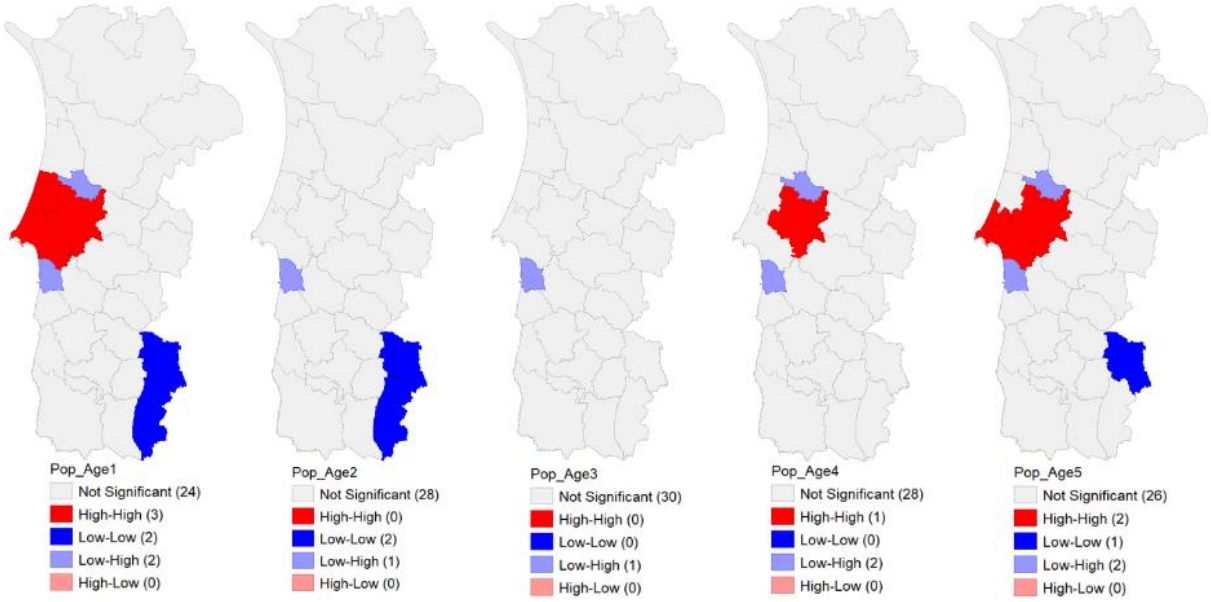


Figura 4.18 *Análise de Autocorrelação Espacial por Faixas Etárias em 2023, da esquerda para a direita, os mapas representam as faixas etárias de 0-14, 15-29, 30-44, 45-59 e 60+ anos.*

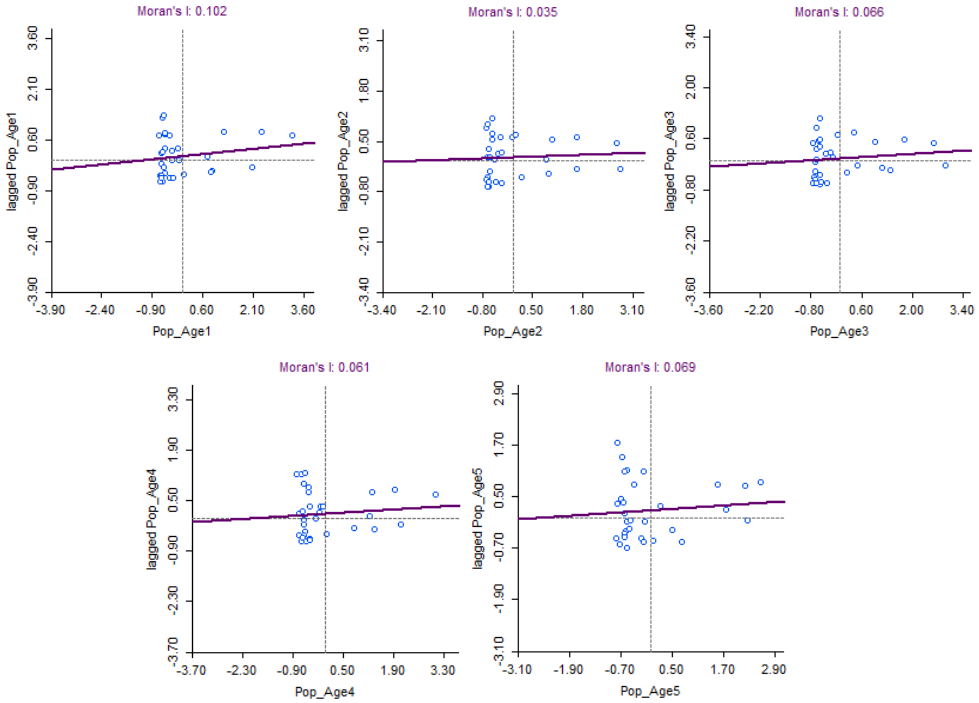


Figura 4.19 *Gráficos de Moran's I para Autocorrelação Espacial por Faixas Etárias em 2023, da esquerda para a direita e de cima para baixo: 0-14, 15-29, 30-44, 45-59 e 60+ anos.*

Relativamente aos índices de Moran, apresentados na Figura 4.19, os valores obtidos para cada faixa etária indicam uma interdependência espacial variável, mas geralmente fraca, sugerindo que a distribuição das faixas etárias no Litoral Alentejano não segue um padrão de agrupamento espacial marcante. Os índices variam entre 0,035 para a faixa etária dos 15-29 anos e 0,102 para a faixa etária dos 0-14 anos, refletindo que, embora existam algumas áreas com concentrações e dispersões, a distribuição geral da população por faixa etária é relativamente dispersa e pouco agrupada.

Habilitações Literárias

Relativamente ao nível de qualificação escolar, os dados foram analisados em dois grupos de indivíduos com 1º, 2º e 3º ciclo completo (Figuras 4.20, 4.21 e 4.22) e indivíduos com o ensino secundário, pós-secundário e superior completo (Figura 4.23, 4.24 e 4.25). A distribuição da escolaridade no Litoral Alentejano revela padrões distintos, refletindo as variações do nível educacional entre as diferentes freguesias.

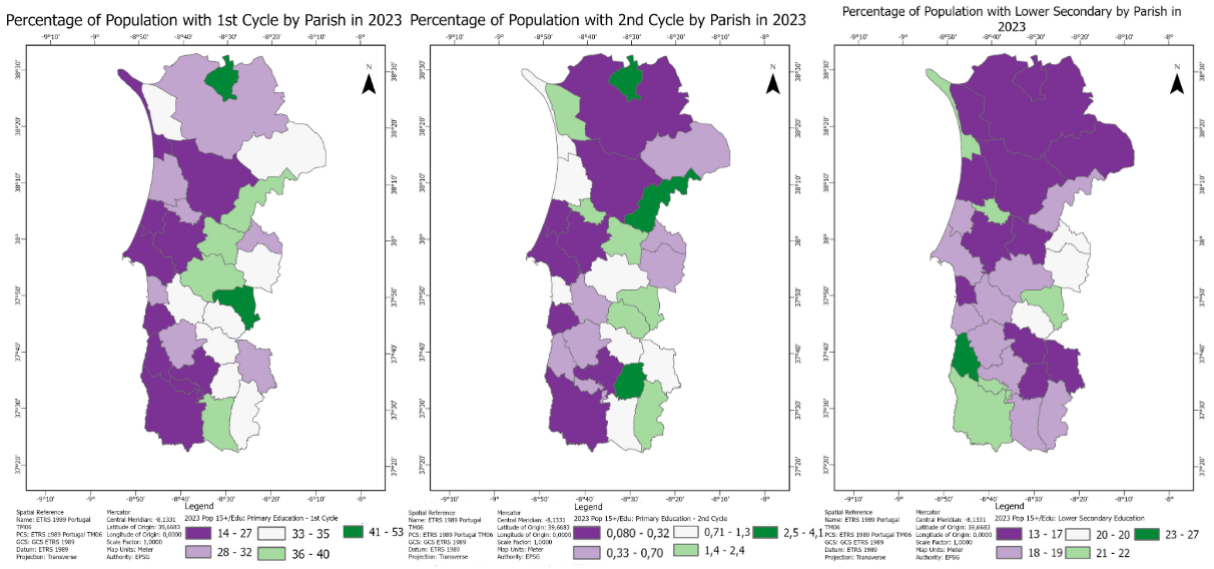


Figura 4.20 *Distribuição das Habilitações Literárias por Freguesia em 2023. Os mapas, da esquerda para a direita, apresentam a percentagem da população que completou o 1º ciclo, 2º ciclo e 3° ciclo.*

No diz respeito ao ensino básico (1º Ciclo) 24.09% da população possui apenas este nível de educação, concentrando-se principalmente nas freguesias de São Martinho e Vale de Santiago. Estas áreas apresentam uma predominância de indivíduos com o perfil educacional mais baixo, tendo concluído apenas o ensino básico.

Além disso, 11.34% da população tem o 2º Ciclo do ensino básico, com maior concentração nas freguesias de Luzianes-Gare, São Martinho e Azinheira dos Barros e São Mamede do Sádão.

Por fim, 18,05% da população possui o 3.º Ciclo do ensino básico, estando concentrada na freguesia de Longueira/Almograve.

Assim, no Litoral Alentejano, existem variações a nível educacional entre freguesias, mais de 50% da população possui apenas a escolaridade básica, o que sugere que estas áreas detêm um perfil educacional muito baixo, evidenciando lacunas a este nível.

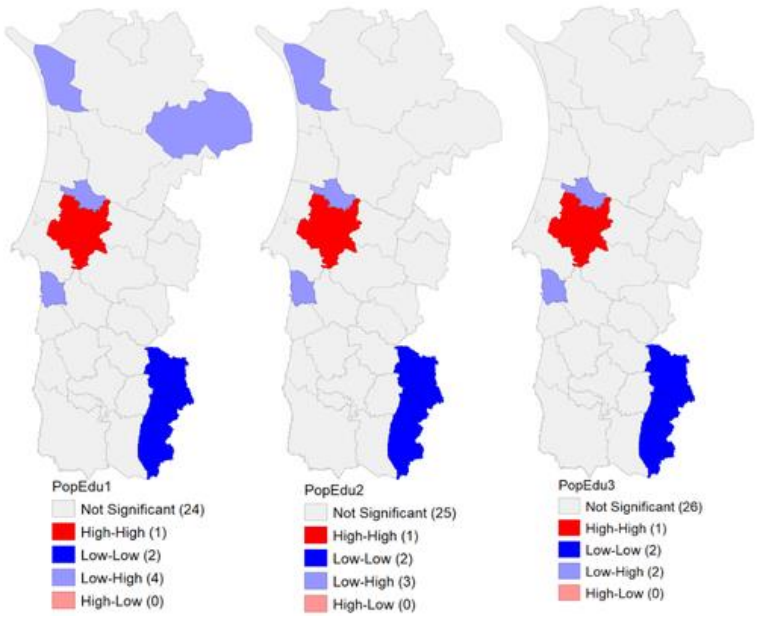


Figura 4.21 *Autocorrelação Espacial das Habilitações Literárias por Freguesia em 2023. Da esquerda para a direita os mapas representam o 1º ciclo (PopEdu1), 2º ciclo (PopEdu2), e 3° ciclo (PopEdu3).*

Na Figura 4.20, são apresentados os mapas de distribuição espacial da escolaridade básica até ao 3.º ciclo no Litoral Alentejano, nos quais se observam padrões distintos em diversas freguesias. Comporta, Torrão, Porto Covo e São Francisco da Serra aparecem consistentemente como áreas com baixas taxas de escolaridade (*Low-High*), sugerindo que estas freguesias possuem uma proporção elevada de população com níveis educacionais mais baixos, estando adjacentes a freguesias onde se verifica o oposto.

Por outro lado, a União das Freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra destaca-se como uma área de alta concentração de população com níveis de escolaridade mais elevados (*High-High*). Esta freguesia revela uma tendência clara de agregação de indivíduos com o nível de educação básica mais alto (3.º ciclo), em contraste com as freguesias vizinhas que apresentam taxas de escolaridade inferiores.

Além disso, Santa Clara-a-Velha e São Martinho das Amoreiras emergem como freguesias com padrões distintos, sendo identificadas como áreas de baixa escolaridade (*Low-Low*) no contexto do ensino básico, o que pode indicar desafios educacionais nestas regiões.

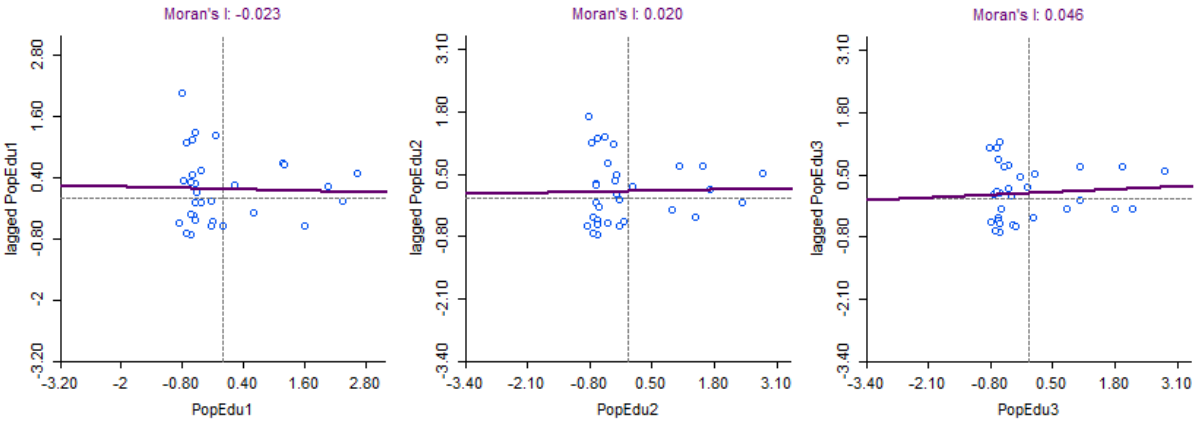


Figura 4.22 *Gráficos de Moran's I para Autocorrelação Espacial das Habilitações Literárias em 2023. Os gráficos representam o índice de Moran's I, da esquerda para a direita, o 1º ciclo (PopEdu1), 2º ciclo (PopEdu2), e 3° ciclo (PopEdu3).*

Os valores do índice de Moran's I, apresentados na Figura 4.22, que indicam a autocorrelação espacial, reforçam a análise espacial anterior. Com valores de -0,023 para o 1º Ciclo, 0,020 para o 2º Ciclo, e 0,046 para o 3º Ciclo do Ensino Básico, estes índices apontam para uma fraca correlação espacial, sugerindo que não existe um padrão forte de agrupamento ou dispersão no que diz respeito ao Ensino Básico nas freguesias do Litoral Alentejano. No entanto, as freguesias mencionadas revelam tendências locais significativas.

Os mapas, apresentados na Figura 4.23, fornecem uma análise das habilitações escolares no Litoral Alentejano, destacando a distribuição da população com ensino secundário, pós-secundário e ensino superior (licenciatura, mestrado e doutoramento) em 2023.

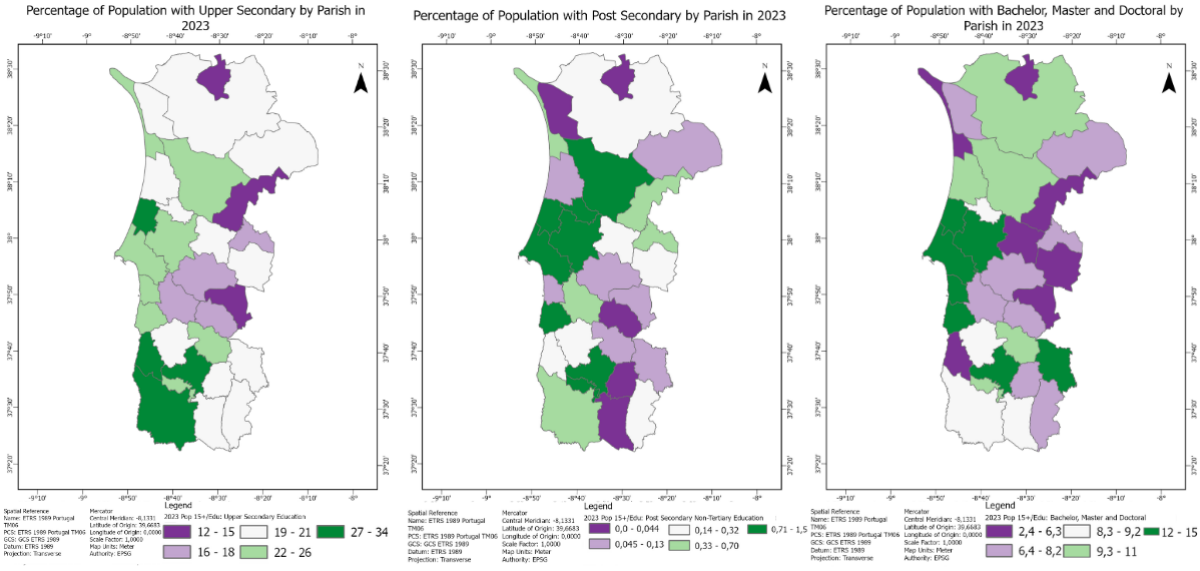


Figura 4.23 *Distribuição das Habilitações Literárias por Freguesia em 2023. Os mapas apresentam a percentagem da população que concluiu o ensino secundário, pós-secundário e graus académicos avançados (licenciatura, mestrado e doutorado), da esquerda para a direita.*

Nesta análise, observa-se que cerca de um quarto da população (23,46%) possui o ensino secundário, com as freguesias de maior percentagem de população com este nível de escolaridade concentradas nas zonas costeiras e urbanas, como São Teotónio e Santo André. Este padrão sugere uma

maior presença de instituições de ensino e, possivelmente, de oportunidades de emprego que incentivam a conclusão do ensino secundário.

Por outro lado, a distribuição da população com habilitações pós-secundárias e superiores varia entre as freguesias. As freguesias de Santiago do Cacém, Sines e a União das Freguesias de Grândola e Santa Margarida da Serra destacam-se com percentagens mais elevadas de população com qualificações superiores, o que pode ser atribuído à presença de atividades económicas que exigem maior qualificação, como a indústria petrolífera e outras indústrias de ponta. Em contraste, as freguesias mais rurais e afastadas dos centros económicos, como Santa Clara-a-Velha, tendem a apresentar percentagens significativamente mais baixas de população com qualificação superior. Este fenómeno pode ser explicado por uma população mais envelhecida, limitações no acesso à educação superior ou pela possível migração de indivíduos mais qualificados para outras regiões.

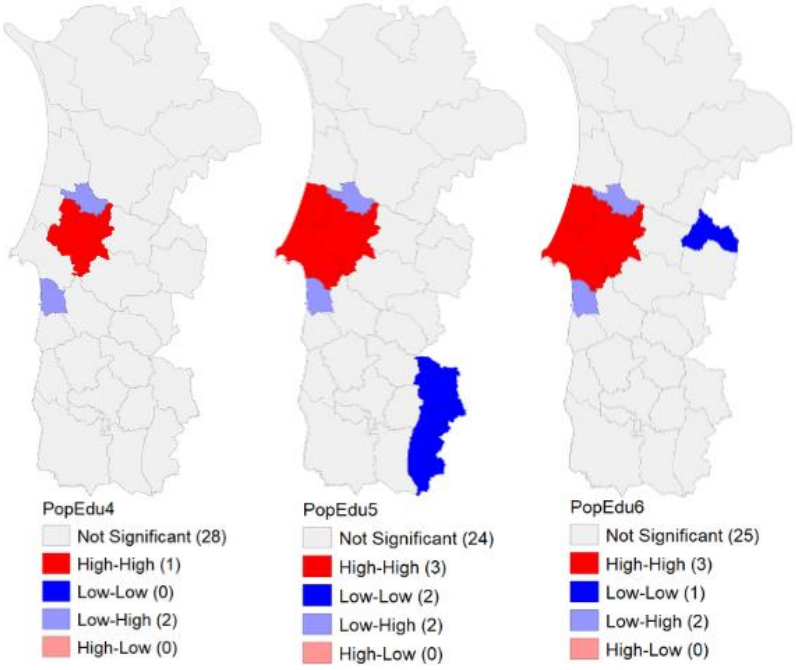


Figura 4.24 *Autocorrelação Espacial das Habilitações Literárias em 2023. Os mapas ilustram a análise de autocorrelação espacial para o ensino secundário (PopEdu4), educação pós-secundária (PopEdu5) e educação superior (PopEdu6), da esquerda para a direita.*

Na análise aos padrões de Moran's I, verificamos que Porto Covo e São Francisco da Serra, classificadas como "*Low-High*", destacam-se por terem uma baixa percentagem de população com níveis educacionais superiores, mas estão rodeadas por freguesias com maiores concentrações de indivíduos com ensino secundário ou superior. Este contraste pode refletir a influência de centros urbanos próximos, onde o acesso a oportunidades educacionais e de emprego qualificado.

Por outro lado, freguesias como Sines, Santo André e a União das Freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra são identificadas como "*High-High*", indicando que são áreas com elevadas percentagens de população com ensino superior e estão rodeadas por freguesias com padrões semelhantes, refletindo a concentração de atividades económicas que exigem alta qualificação. Em contraste, Santa Clara-a-Velha, São Martinho das Amoreiras e Ermidas-Sado são classificadas como "*Low-Low*", caracterizando-se por baixas percentagens de população com educação superior, tanto internamente quanto nas freguesias vizinhas.

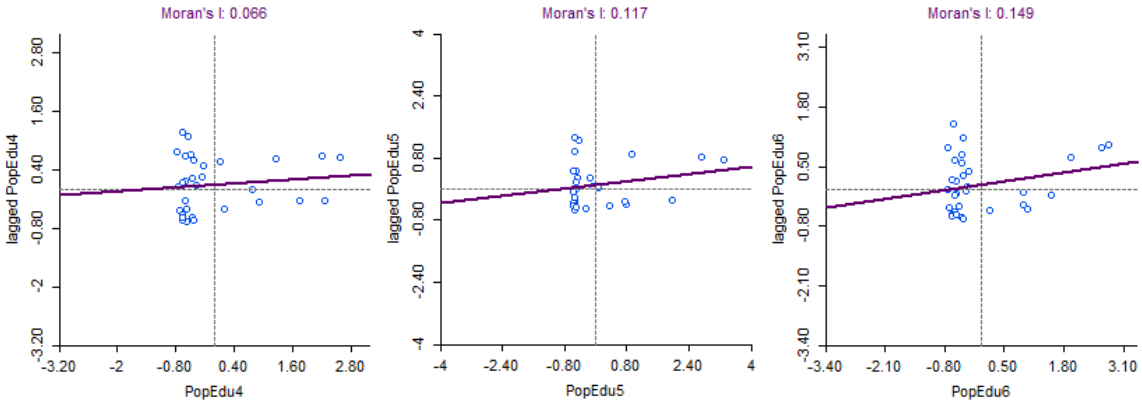


Figura 4.25 *Análise de Moran's I das Habilitações Literárias por Freguesia em 2023. Os gráficos representam o índice de Moran's I, da esquerda para a direita, o ensino secundário (PopEdu4), educação pós-secundária (PopEdu5) e educação superior (PopEdu6).*

Relativamente aos valores de Moran's I obtidos na Figura 4.25, não se observa uma grande variação, o que indica que o grau de autocorrelação espacial entre as freguesias em termos de níveis de escolaridade é relativamente consistente. No caso de ensino secundário, o índice de Moran's I é de 0,066, enquanto para níveis de ensino pós-secundário e superior, os índices são de 0,117 e 0,149, respetivamente. Estes valores sugerem uma ligeira tendência de agrupamento espacial, mas não uma correlação espacial forte. Isso significa que, embora existam algumas freguesias com percentagens elevadas ou baixas de população com educação superior que concentradas em determinadas áreas, a distribuição educacional no Litoral Alentejano não apresenta um padrão de concentração muito acentuado, refletindo uma mistura de áreas com diferentes níveis educacionais sem uma segregação espacial muito marcada.

Rendimentos

Apresentamos, a seguir, uma análise detalhada da distribuição dos rendimentos familiares, evidenciando as disparidades económicas entre as diferentes freguesias. Através da segmentação por faixas de rendimento anuais, identificam-se padrões de rendimentos mais elevados e de vulnerabilidade económica na região. Esta análise proporciona uma compreensão mais profunda das dinâmicas socioeconómicas que caracterizam tanto as áreas urbanas como as rurais desta zona.

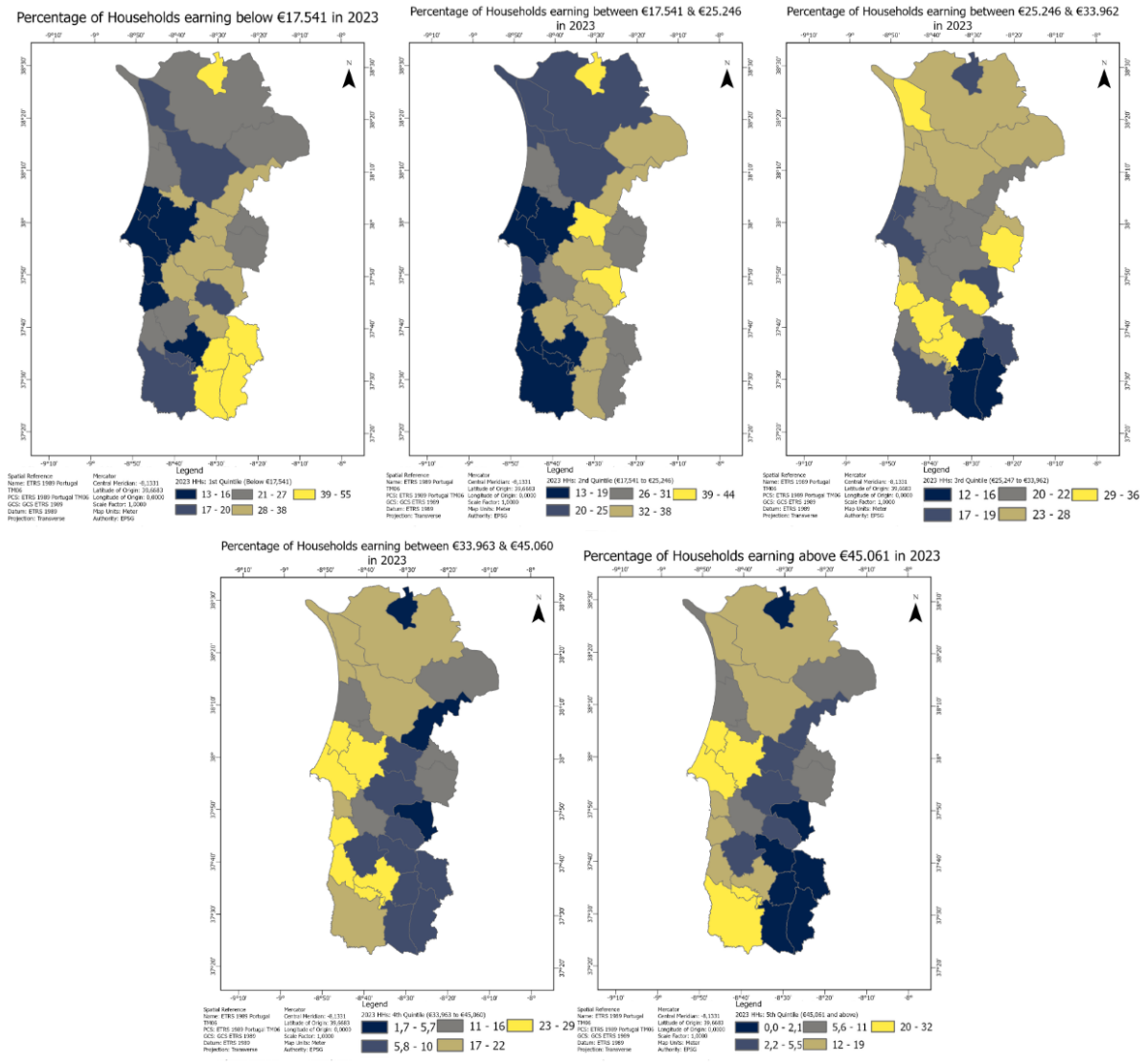


Figura 4.26 *Distribuição de Rendimentos em 2023. Os mapas apresentam a percentagem de agregados familiares em diversas faixas de rendimentos.*

Nas freguesias com rendimentos mais baixos, onde uma grande percentagem de agregados familiares ganha menos de €17.541 anuais, destacam-se zonas como Relíquias, São Martinho, Sabóia, Vale de Santiago e várias outras freguesias do interior. Estas áreas revelam uma acentuada vulnerabilidade económica, com uma parte da população a viver com rendimentos bastante limitados. Este padrão pode estar associado a uma economia predominantemente rural e menos diversificada, com menos oportunidades de emprego bem remunerado.

As freguesias com rendimentos médios, como Colos, Alvalade e São Luís, apresentam uma percentagem significativa de agregados familiares a ganhar entre €17.541 e €45.060 anuais. Estas áreas

parecem ter uma distribuição de rendimentos mais equilibrada, provavelmente devido a uma combinação de atividades económicas que proporcionam rendimentos estáveis, embora não tão elevados quanto nas freguesias urbanas mais industrializadas.

Por outro lado, ao analisar as faixas de rendimento mais altas, acima de €45.061 anuais, as freguesias de Santo André, São Teotónio, Sines, e a União das Freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra emergem como polos de maiores rendimentos. Estas áreas, mais urbanizadas e economicamente dinâmicas, beneficiam da presença de indústrias como a petroquímica, de atividades agrícolas intensivas e algum turismo. Este contraste com freguesias de rendimentos mais baixos traduz a desigualdade económica presente na região, onde a proximidade a centros industriais e urbanos tende a concentrar maiores rendimentos.

Este padrão, visível na Figura 4.26, de rendimentos evidencia uma clara divisão entre as freguesias mais urbanas e economicamente desenvolvidas e as áreas rurais menos prósperas, refletindo os desafios em termos de desenvolvimento regional e de oportunidades económicas.

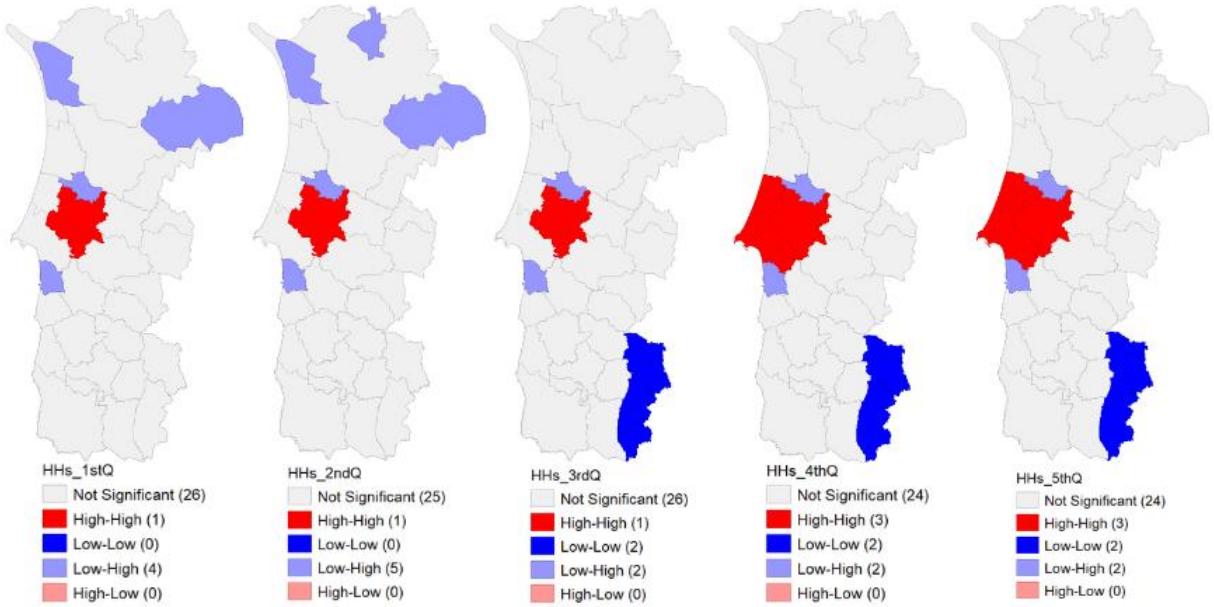


Figura 4.27 *Autocorrelação Espacial das Faixas de Rendimentos em 2023.*

Os mapas de Moran, apresentados na Figura 4.27, revelam padrões interessantes na distribuição espacial dos diferentes agrupamentos de rendimento no Litoral Alentejano. As freguesias de Santo André, Sines, e a União das Freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra, que se destacam consistentemente como "*High-High*", evidenciam uma concentração de rendimentos superiores em comparação com as áreas vizinhas. Em contraste, freguesias como Santa Clara-a-Velha e São Martinho das Amoreiras aparecem como "*Low-Low*" em vários quantis, indicando uma concentração de rendimentos baixos. Adicionalmente, as freguesias de Porto Covo e São Francisco da Serra são identificadas como "*Low-High*", sugerindo que, apesar de apresentarem baixos rendimentos, estão cercadas por freguesias com rendimentos mais elevados.

A zona do Litoral Alentejano revela uma desigualdade de rendimentos entre as freguesias. Os rendimentos mais altos concentram-se em apenas três freguesias, abrangendo uma pequena percentagem da população (18,37%). Em contraste, a restante população, representando 81,63%, enfrenta condições económicas mais desfavoráveis, caracterizando-se por um risco elevado de pobreza ou exclusão social.

Esta situação combina fatores como a pobreza relativa, a privação material e social severa, e uma intensidade laboral per capita reduzida.

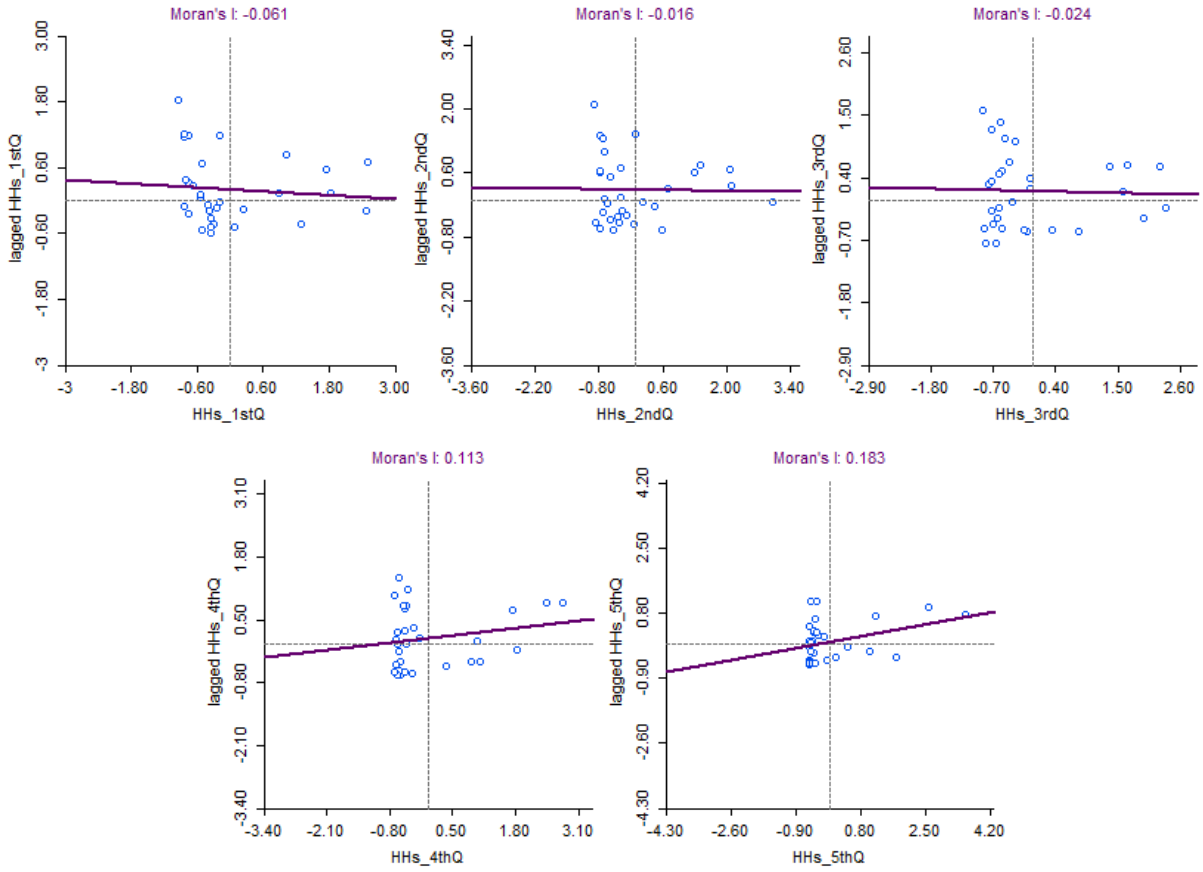


Figura 4.28 *Gráficos de Moran's I para Autocorrelação Espacial das Faixas de Rendimentos em 2023.*

Os valores do índice de Moran's I, na Figura 4.28, obtidos para os diferentes quantis de rendimento refletem uma autocorrelação espacial variada, geralmente fraca. Para o primeiro quintil (HHs_1stQ), o índice de Moran's I é de -0,061, indicando uma leve tendência de dispersão espacial dos rendimentos mais baixos. No segundo (HHs_2ndQ), o valor de Moran's I é de -0,016, ainda sugerindo uma ausência de *clustering* significativo. O valor de Moran's I para o terceiro quintil (HHs_3rdQ) é de -0,024, mantendo a tendência de baixa autocorrelação espacial. A partir do quarto quintil (HHs_4thQ), observamos um ligeiro aumento da autocorrelação positiva, com Moran's I a 0,113, indicando um pequeno grau de *clustering* em torno de rendimentos medianos-altos. Finalmente, no último quintil (HHs_5thQ), o índice de Moran's I atinge 0,183, sugerindo uma maior concentração espacial de rendimentos mais elevados, embora o valor ainda indique uma correlação espacial relativamente moderada.

Despesa em Álcool e tabaco

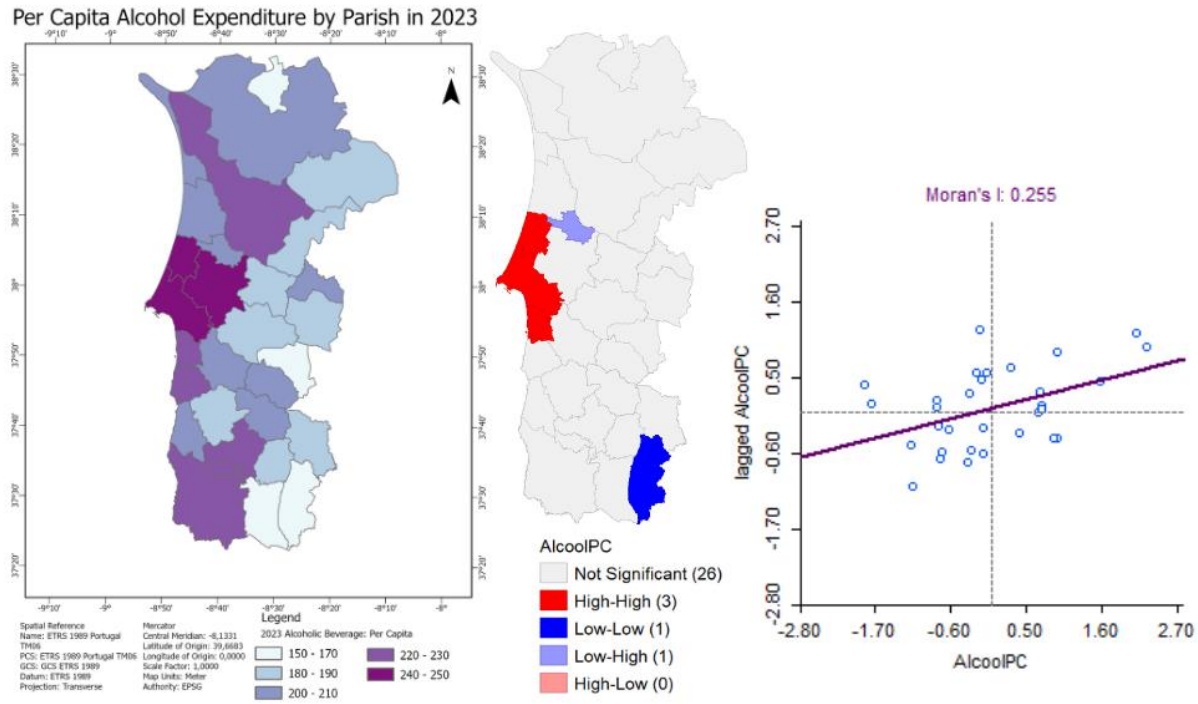


Figura 4.29 *Distribuição Per Capita de Despesas com Álcool e Autocorrelação Espacial por Freguesia em 2023. O mapa detalha a despesa per capita com álcool por freguesia, mostra os padrões espaciais de autocorrelação e apresenta o índice de Moran's I, da esquerda para a direita.*

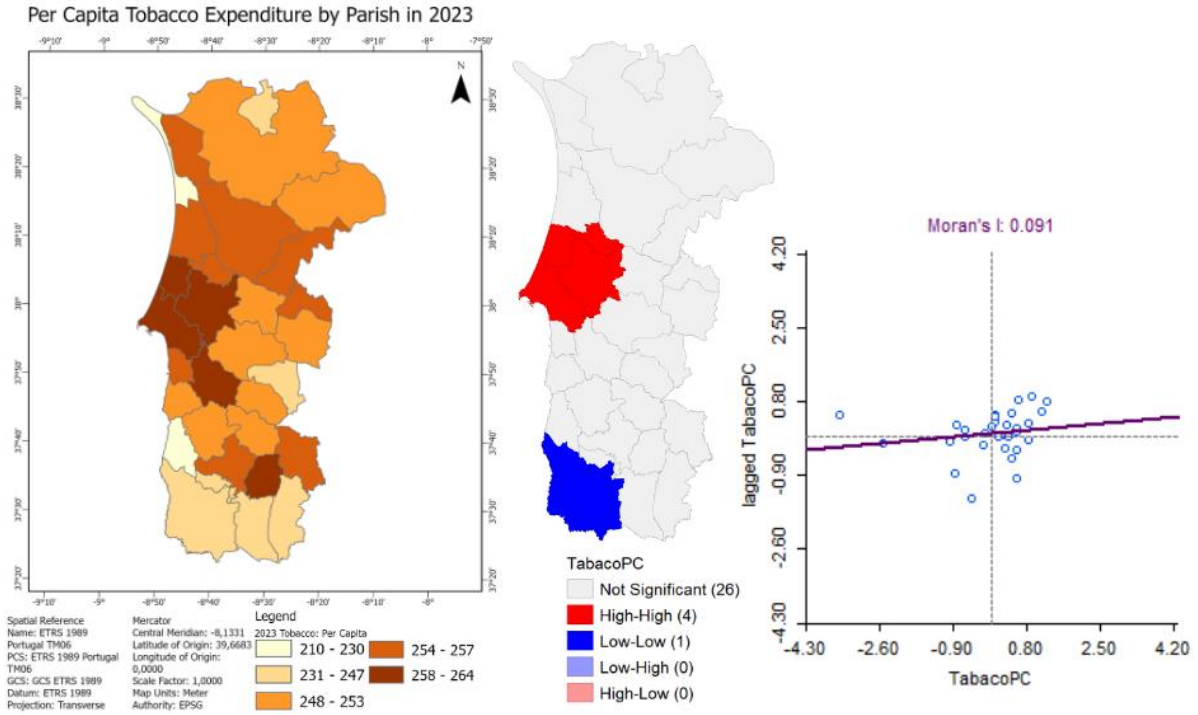


Figura 4.30 *Distribuição Per Capita de Despesas com Tabaco e Autocorrelação Espacial por Freguesia em 2023.*

Nos mapas apresentados na Figura 4.29, que ilustram os valores, em euros, gastos em álcool, observa-se que as freguesias situadas mais a oeste, como Sines, Santo André e a União das freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra, registam os maiores gastos per capita,

oscilando entre 240€ e 250€. Estas áreas são destacadas como "*High-High*" nos mapas de Moran, sugerindo uma elevada concentração de consumo em freguesias adjacentes. Em contraste, freguesias mais interiores como Vale de Santiago e Santa Clara-a-Velha registam gastos de menor valor, situando-se entre 150€ e 170€ per capita, com Santa Clara-a-Velha destacada como "*Low-Low*". O índice de Moran, com o valor de 0,255, confirma uma autocorrelação espacial moderada, sugerindo que as freguesias com altos ou baixos níveis de consumo de álcool tendem a agrupar-se geograficamente.

A análise dos gastos per capita em tabaco, presente na Figura 4.30, revela que as freguesias centrais do Litoral Alentejano, como Cercal, Sines e Santo André e União das freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra, apresentam os valores mais elevados que vão de 258€ e 264€, refletindo um padrão de consumo acentuado nestas áreas. Os mapas de Moran reforçam estes dados, com as já referidas freguesias identificadas como "*High-High*". Em contraste, as regiões mais a sul, como Longueira/Almograve e Boavista dos Pinheiros, registam gastos um pouco mais baixos, variando entre 210€ e 230€. Na análise do mapa de Moran São Teotónio surge como "*Low-Low*", mostrando um padrão de consumo menor e concentrado em áreas próximas com comportamentos semelhantes. O índice de Moran para o tabaco, surge com o valor de 0,091, o que indica uma correlação espacial fraca, mas ainda assim, aponta para a existência de padrões não aleatórios de consumo.

A comparação entre os padrões de despesas com tabaco e álcool no Litoral Alentejano revela algumas semelhanças consideráveis. As freguesias de Santo André, Sines e a União das freguesias de Santiago do Cacém, Santa Cruz e São Bartolomeu da Serra são consistentemente identificadas como áreas de maiores gastos em ambos os produtos, sugerindo uma correlação entre os hábitos de consumo de tabaco e álcool nessas regiões. Por outro lado, freguesias como Santa Clara-a-Velha e São Teotónio aparecem com gastos menores em ambos os itens. Estes resultados indicam que, apesar das diferenças na intensidade da autocorrelação espacial para cada produto, existem zonas no Litoral Alentejano onde a despesa com tabaco e álcool é simultaneamente elevada ou reduzida, reforçando a ideia de que existem fatores socioeconómicos e culturais comuns que podem estar a influenciar os padrões de consumo.

Empregabilidade

Para analisar a empregabilidade no Litoral Alentejano, foi representado na Figura 4.31 o desemprego por freguesia, proporcionando uma visão clara das áreas com maior e menor vulnerabilidade ao desemprego.

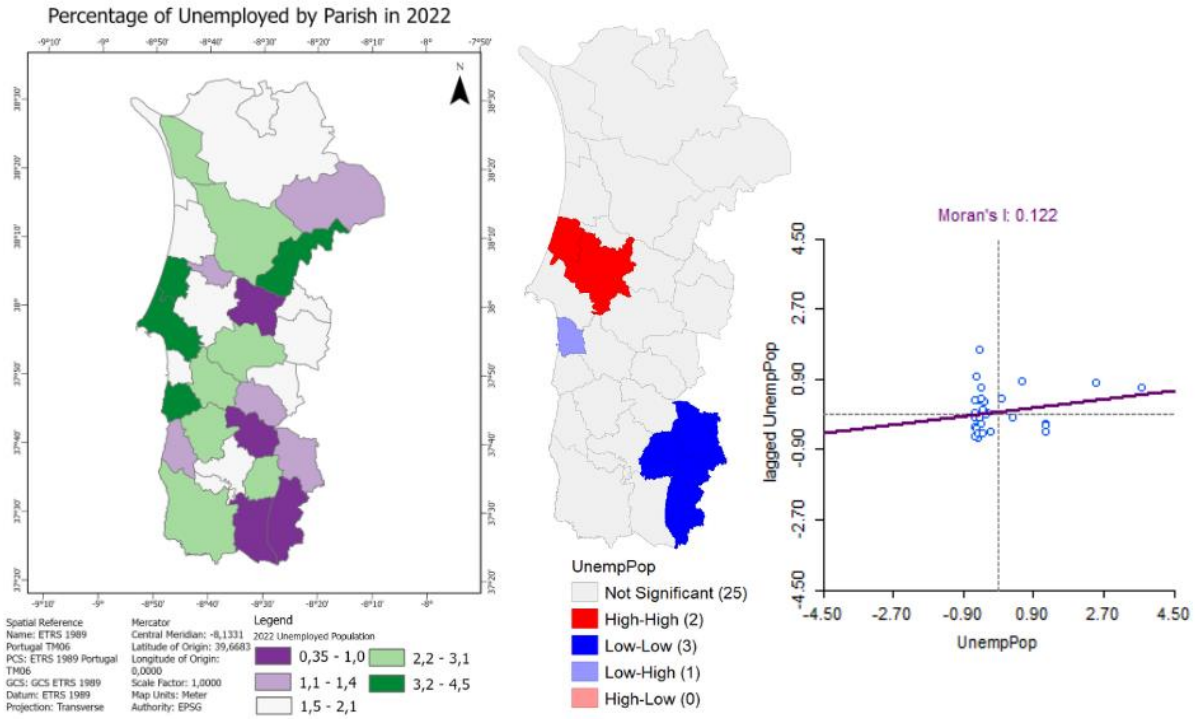


Figura 4.31 *Distribuição Percentual do Desemprego e Autocorrelação Espacial por Freguesia em 2022. Da esquerda para a direita, o mapa detalha a percentagem de desempregados por freguesia, ilustra padrões espaciais e o índice de Moran's I.*

Os mapas apresentados oferecem uma análise detalhada do desemprego por freguesia no Litoral Alentejano em 2022, destacando as variações na taxa de desemprego entre diferentes regiões. No mapa da percentagem de desempregados por freguesia, observa-se que as áreas com maior vulnerabilidade ao desemprego estão localizadas predominantemente nas zonas do litoral centro, como Sines, Santo André e Vila Nova de Milfontes, com exceção da freguesia mais interior de Azinheira dos Barros e São Mamede do Sádão. Estas áreas, onde as percentagens de desemprego variam entre 3,2% e 4,5%, indicam uma maior vulnerabilidade económica, possivelmente devido à dependência de setores mais voláteis ou à natureza específica das indústrias locais, que podem não absorver a totalidade da força de trabalho.

Em contraste, outras freguesias mais interiores, apresentam as menores taxas de desemprego, situando-se entre 0,35% e 1%. Este valor de desemprego mais baixo no interior pode estar relacionado com uma população envelhecida e estabilizada, muitas vezes reformada ou com empregos duradouros, o que reduz a pressão sobre o mercado de trabalho e resulta em taxas de desemprego mais baixas em comparação com as freguesias do litoral, que enfrentam maior pressão populacional.

O mapa de Moran revela a existência de *clusters* de desemprego. As freguesias classificadas como "*High-High*" estão localizadas no litoral centro, onde o desemprego é elevado tanto localmente quanto nas freguesias vizinhas. Este padrão pode refletir desafios estruturais no mercado de trabalho, como a dependência de setores específicos. Por outro lado, as freguesias "*Low-Low*" no interior

mostram um padrão de estabilidade, onde o baixo desemprego numa freguesia correlacionado com baixos níveis de desemprego nas freguesias vizinhas.

O gráfico de dispersão de Moran, reforça esta análise com um valor de Moran's I de 0,122, sugerindo uma autocorrelação espacial positiva, embora fraca. Este valor indica que, no Litoral Alentejano, o desemprego tende a convergir espacialmente, mas de forma pouco acentuada. Em síntese, a análise aponta para uma realidade socioeconómica desigual na região, onde algumas freguesias enfrentam desafios em termos de desemprego, enquanto outras beneficiam de condições mais favoráveis.

Estado Civil

Para analisar a distribuição do estado civil no Litoral Alentejano, foram representadas as percentagens da população solteira, casada e divorciada por freguesia. Esta análise pretende compreender as variações espaciais dos diferentes estados civis e identificar possíveis padrões na região.

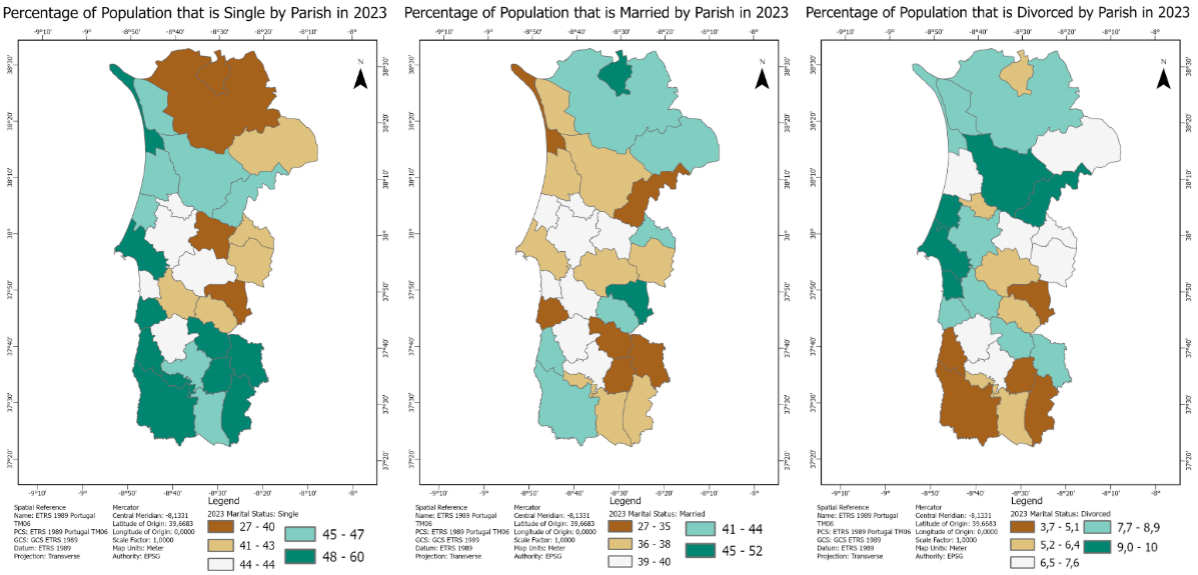


Figura 4.32 *Distribuição de Estado Civil por Freguesia em 2023. Da esquerda para a direita, os mapas mostram a percentagem de solteiros, de casados e de divorciados.*

Os mapas apresentados oferecem uma visão abrangente sobre a distribuição por estado civil no Litoral Alentejano, em 2023, revelando complementares entre si. No mapa relativo à percentagem de população solteira, observa-se que as freguesias localizadas no litoral norte e no sul da região apresentam as percentagens mais elevadas de solteiros, variando entre 44% e 60%, em contraste com as freguesias do interior. Este padrão pode ser influenciado por fatores como a migração de jovens para centros urbanos ou por características demográficas específicas dessas áreas. Por outro lado, nas freguesias a norte e no litoral sul, como São Teotónio e Longueira/Almograve, predomina a população casada. No mapa da percentagem de população divorciada, a maior concentração de indivíduos divorciados ocorre nas freguesias costeiras, como Sines, Porto Covo e Santo André, e também em Grândola e na União das Freguesias de Santa Margarida da Serra.

Estas diferenças refletem a influência de fatores económicos, sociais e culturais nas escolhas e na estabilidade conjugal das populações das diferentes freguesias da região.

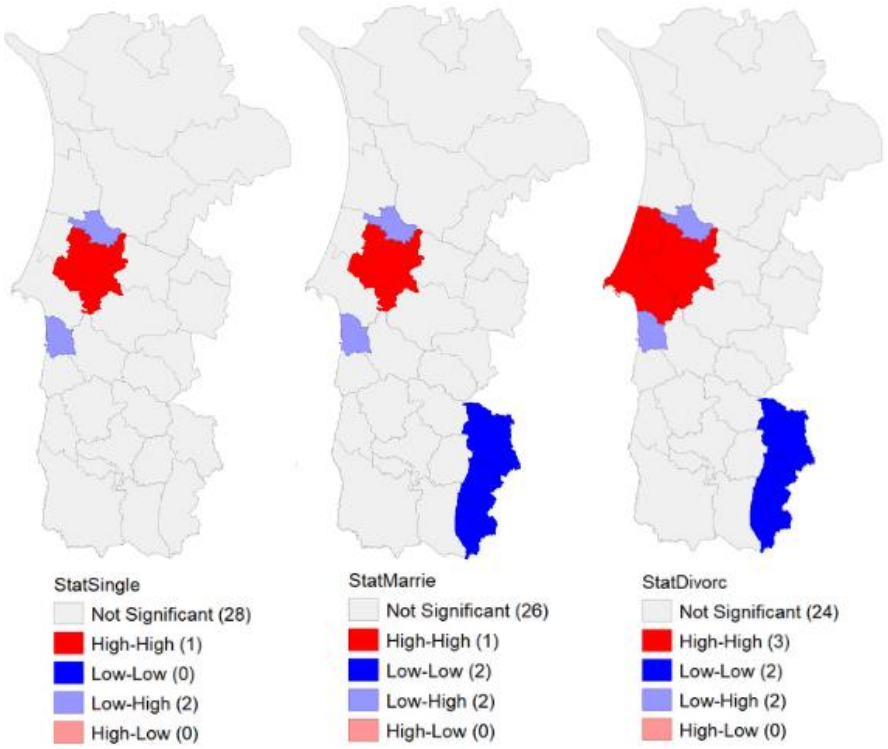


Figura 4.33 *Autocorrelação Espacial dos Estados Civis em Freguesias Selecionadas em 2023. Os mapas analisam a autocorrelação espacial de estados civis: solteiros, casados e divorciados, da esquerda para a direita.*

Os mapas de Moran, na Figura 4.33, indicam a presença de *clusters* espaciais relacionados ao estado civil. No caso da população solteira e casada, observam-se *clusters* "*High-High*" nas freguesias do litoral, como Santiago do Cacém, sugerindo que essas áreas são caracterizadas por altas percentagens de indivíduos solteiros e casados. Em contraste, nas freguesias do interior, como São Martinho das Amoreiras, os *clusters* "*Low-Low*" indicam uma menor concentração de população casada e divorciada.

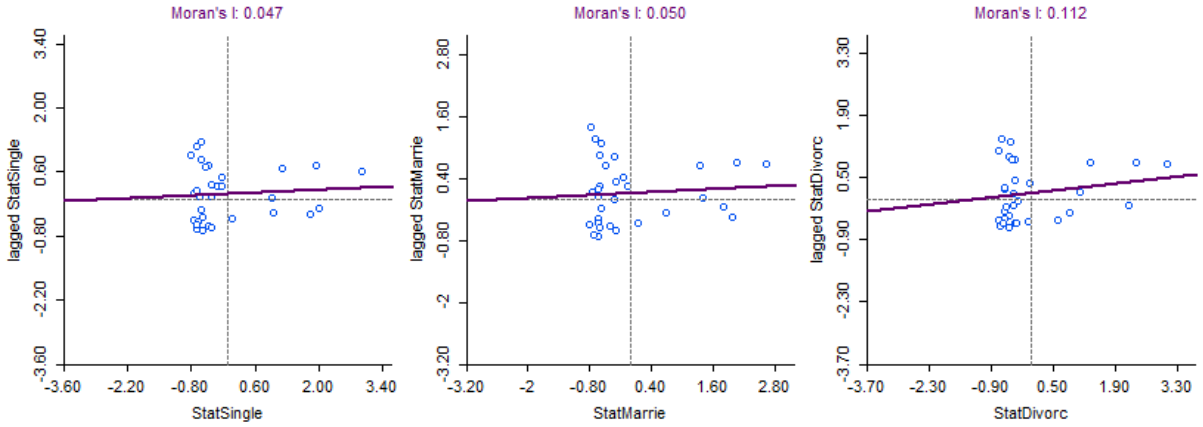


Figura 4.34 *Análise de Moran's I para Autocorrelação Espacial de Estados Civis em 2023. Os três gráficos mostram o índice de Moran's I com base no estado civil: solteiro, casado e divorciado, da esquerda para a direita.*

Os índices de Moran, na Figura 4.34, calculados para os estados civis reforçam as observações feitas anteriormente. O índice de Moran para a população solteira foi de 0,047, enquanto para os divorciados foi de 0,112 e para os casados, 0,050. Esses valores positivos, embora baixos, indicam a existência de uma pequena autocorrelação espacial, especialmente no caso dos divorciados, onde o índice mais elevado sugere uma maior concentração espacial deste estado civil em determinadas freguesias do litoral, como Sines.

Poder de Compra

A análise do poder de compra per capita nas freguesias da região em 2023 revela tendências importantes, evidenciando as disparidades socioeconómicas regionais. Os mapas coropléticos e de Moran, permitem identificar padrões espaciais que diferenciam as freguesias de maior e menor poder de compra.

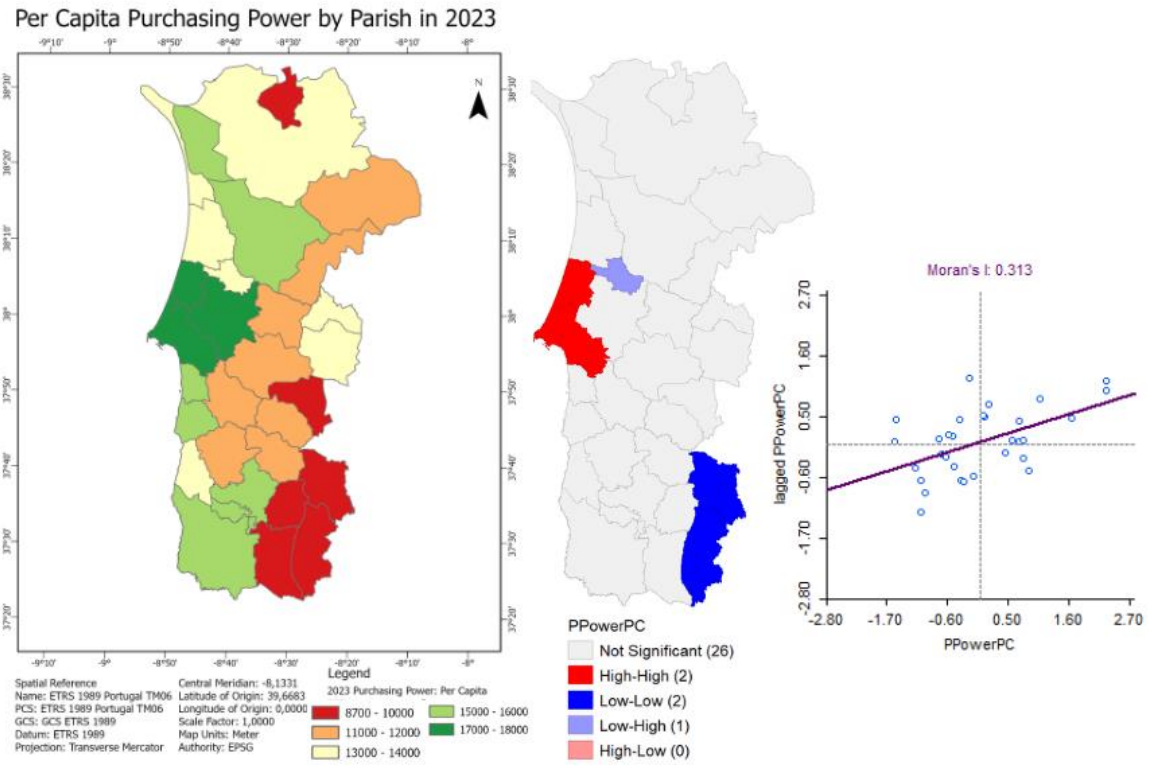


Figura 4.35 *Distribuição do Poder de Compra Per Capita por Freguesia e Autocorrelação Espacial em 2023. Da esquerda para a direita, o poder de compra per capita por freguesia, os padrões de autocorrelação e o gráfico do índice de Moran's I.*

O mapa coroplético, na Figura 4.35, evidencia uma clara divisão entre as freguesias quanto ao poder de compra per capita. Nas freguesias do interior a laranja e vermelho, especialmente a sul, destacam-se pelos níveis muito baixos de poder de compra, com valores entre 8.700€ e 12.000€ per capita. Em contraste as freguesias do litoral, como Sines, Santo André e Santiago do Cacém, apresentam um poder de compra mais elevado, rondando os 18.000€. Esta discrepância pode ser atribuída à maior atividade económica e à oferta de serviços nas regiões costeiras, com a maior atividade económica e oferta de serviços nas regiões litorais, em oposição às áreas interiores, onde a economia é menos dinâmica.

O mapa de Moran revela uma concentração de áreas com elevado poder de compra per capita (*High-High*), principalmente concentradas na zona costeira de Sines, uma área economicamente mais desenvolvida, com infraestruturas robustas e maior concentração de indústria e serviços. Em contrapartida, as áreas *Low-Low*, localizadas no interior, a sul da região, refletem as zonas economicamente mais desfavorecidas.

O índice de Moran, com um valor de 0,313, indica uma autocorrelação espacial positiva moderada, sugerindo que freguesias com altos valores de poder de compra tendem a estar próximas umas das outras, assim como as freguesias com baixos valores. Este resultado comprova a análise dos mapas coropléticos e de Moran, onde se observa uma clara separação entre as regiões mais prósperas,

no litoral e as menos favorecidas no interior. A existência desta autocorrelação espacial positiva reflete as dinâmicas regionais do Litoral Alentejano, onde o desenvolvimento económico tende a concentrar - se em áreas geograficamente contíguas, influenciado por fatores como oportunidades de emprego, infraestruturas de saúde e acesso a serviços.

4.4.2. Variável de Interesse

Diabetes

A análise da distribuição de pacientes com diabetes por freguesia, acompanhada da prevalência desta condição, proporciona uma visão detalhada da distribuição espacial da diabetes, na região em estudo. O primeiro mapa apresenta a contagem absoluta de utentes por freguesia, a prevalência, apresentada no segundo mapa, ajusta esta informação em função da população total, permitindo uma análise mais precisa e comparativa entre freguesias com diferentes densidades populacionais.

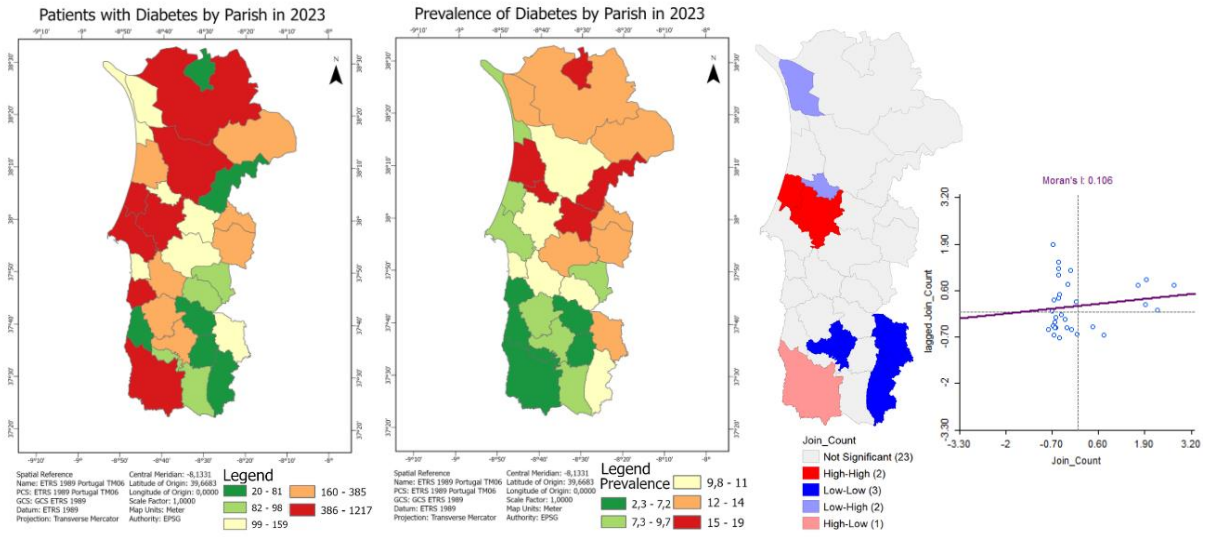


Figura 4.36 *Distribuição de Pacientes com Diabetes (primeiro mapa à esquerda) e Prevalência (segundo mapa) por Freguesia em 2023. O mapa à direita e o gráfico de Moran's I analisam a autocorrelação espacial.*

Os mapas coropléticos, na Figura 4.36, evidenciam uma distribuição desigual tanto na contagem de utentes com diabetes quanto na prevalência da patologia. Notam-se áreas com alta concentração de utentes, principalmente na zona centro e norte da região, que se destacam tanto em número absoluto de pacientes quanto em elevada prevalência. O litoral, em geral, apresenta uma prevalência relativamente mais baixa, com exceção de Melides e São Francisco da Serra.

Os mapas de Moran indicam uma correlação espacial na distribuição dos utentes com diabetes. As freguesias classificadas como "*High-High*" mostram um elevado número absoluto de casos de diabetes, mas com uma menor prevalência relativa da doença. o que pode ser explicado pelo facto de estas freguesias terem uma população total maior, resultando numa diluição da prevalência, apesar do elevado número de casos. Em contrapartida, as áreas classificadas como "*Low-Low*" revelam freguesias com menos utentes diabéticos e uma maior prevalência, o que pode estar associado a populações menores.

Um Índice de Moran de 0.106 sugere uma presença moderada de autocorrelação espacial. Isto indica que as áreas próximas tendem a apresentar valores semelhantes, embora essa semelhança não seja particularmente forte. Noutras palavras, existe uma tendência para o agrupamento espacial, mas não é

muito acentuada, o que justifica o uso da GWR para explorar como as relações entre variáveis independentes e a variável dependente variam no espaço.

Tabela 4.6 *Resumo da análise de autocorrelação espacial para diferentes variáveis no LA e respetivos índices de Moran's I.*

Variável	Índice de Moran's I	Resumo das Conclusções
População Total e por Gênero	0,050 (População Total) 0,089 (Feminina) 0,031 (Masculina)	A União das Freguesias de Santiago do Cacém é classificada como High-High, indicando alta concentração populacional. A distribuição por gênero é distinta, com uma maior percentagem de homens nas freguesias do litoral sul.
Faixa Etária	0,102 (0-14 anos) 0,035 (15-29 anos) 0,077 (30-44 anos) 0,045 (45-59 anos) 0,090 (60+ anos)	As freguesias costeiras como Sines e Vila Nova de Santo André apresentam maior percentagem de jovens.
Habilitações Literárias	0,023(1º Ciclo) 0,020 (2º Ciclo) 0,046 (3º Ciclo) 0,066 (Ensino Secundário) 0,117 (Pós-Secundário) 0,149 (Superior)	Mais de 50% da população possui apenas a escolaridade básica.
Rendimentos	0,061 (1º Quantil) 0,016 (2º Quantil) 0,024 (3º Quantil) 0,113 (4º Quantil) 0,183 (5º Quantil)	As freguesias urbanas como Santo André e Sines têm rendimentos mais elevados.
Despesas com Ãlcool e Tabaco	0,255 (Álcool) 0,091 (Tabaco)	Despesas per capita mais altas em álcool e tabaco foram encontradas nas freguesias como Sines e Santo André.
Empregabilidade	0,122	Zonas como Sines e Santo André têm taxas de desemprego mais altas (3,2% a 4,5%).
Estado Civil	0,047 (Solteiros) 0,050 (Casados) 0,112 (Divorciados)	Freguesias do litoral como Sines têm altas percentagens de indivíduos solteiros e divorciados.
Poder de Compra	0,313	Freguesias costeiras como Sines apresentam um poder de compra mais elevado.
Diabetes	0,106	Zonas com alta concentração de diabetes estão no centro e norte, com excessão de São Teotónio.

4.5. Modelação

Para aprofundar a análise e compreender melhor as inter-relações entre os diferentes fatores e a quantidade de utentes com diabetes, foi elaborado um *scatter plot matrix*, presente na Figura 4.37. Esta representação gráfica permite visualizar as correlações entre as variáveis analisadas, como poder de compra, estado civil, desemprego, entre outras, e o número de utentes com diabetes em cada freguesia. Este processo é essencial para identificar quais as variáveis que exercem uma influência significativa, que serão posteriormente incorporadas num modelo de GWR, permitindo uma modelação precisa e adaptada às especificidades espaciais da região do litoral alentejano.

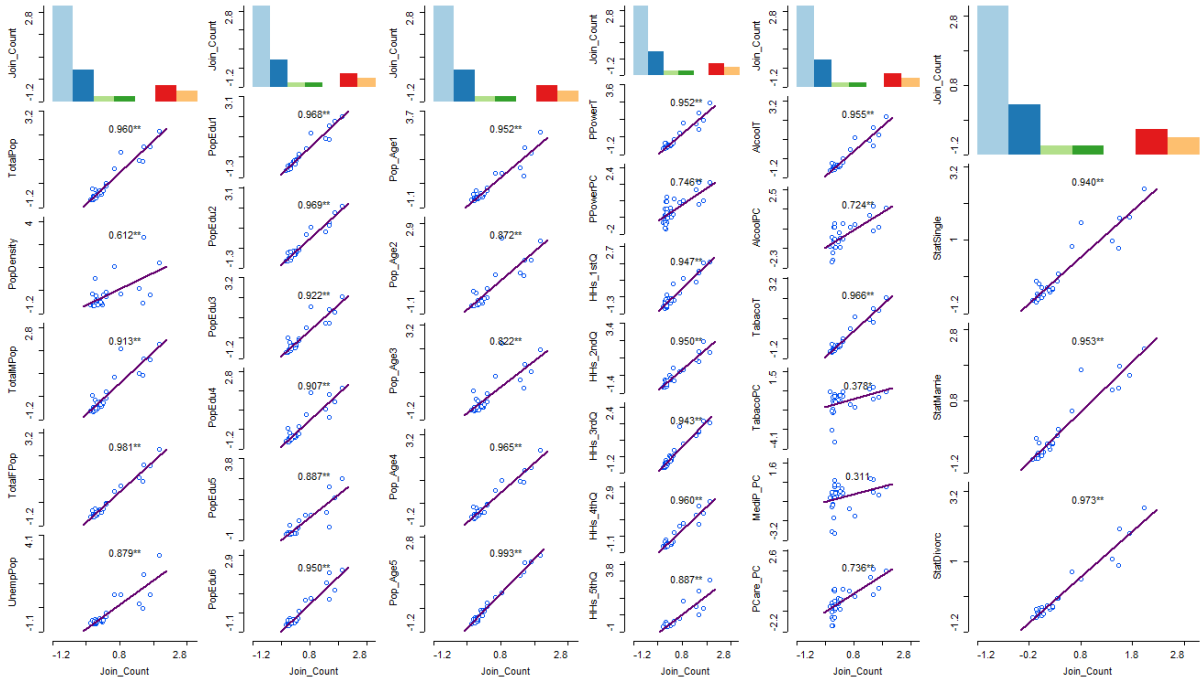


Figura 4.37 *Matriz de Gráficos de Dispersão, que analisa as relações entre as variáveis independentes e a variável dependente.*

Nesta análise dos gráficos de dispersão, verificou-se que a maior prevalência da doença está consistentemente associada a uma população mais idosa, do sexo feminino e a níveis de escolaridade baixos. Adicionalmente, indicadores como o consumo de álcool e tabaco mostraram uma correlação elevada com a prevalência da diabetes, o que reforça a necessidade de políticas de saúde pública direcionadas para a diminuição do consumo e incentivo de estilos de vida mais saudáveis. Por outro lado, a correlação entre o estado civil e a prevalência da diabetes sugere que fatores sociais e de suporte familiar podem desempenhar um papel significativo na saúde.

Os resultados desta análise exploratória são importantes para a formulação de intervenções de saúde que sejam direcionadas e eficazes, concentrando-se nos DSS mais relevantes associados à prevalência de diabetes tipo 2 nesta região específica. Esta compreensão aprofundada contribuirá para a melhoria das condições de saúde da população local, permitindo uma aplicação mais precisa dos recursos de saúde pública.

4.5.1. Modelo de Regressão

No âmbito deste estudo, foi desenvolvido um Modelo de Regressão (Ordinary Least Square OLS) com o objetivo de, posteriormente, se proceder à análise comparativa dos resultados entre as duas ferramentas (OLS e GWR). O principal propósito deste modelo OLS foi identificar as relações globais entre a prevalência de diabetes e as variáveis explicativas selecionadas, servindo como referência para avaliar a pertinência e as vantagens de aplicar técnicas de análise espacial mais avançadas.

O modelo OLS foi desenvolvido com duas variáveis independentes: "Percentagem de População com o Primeiro Ciclo" (PER1EDU) e "Percentagem de População cujo Estado Civil é Divorciado" (PERDIVORC). A seleção destas variáveis baseou-se nos valores de correlação obtidos na matriz do gráfico de dispersão, escolhendo-se aquelas que melhor explicam a patologia na região estudada e que não apresentassem elevados valores de fatores de inflação da variância (VIF) no modelo.

Este modelo apresentou um coeficiente de determinação ajustado (R^2 ajustado) de 0,5515, indicando que cerca de 55,2% da variabilidade da prevalência de diabetes é explicada pelas duas variáveis independentes. O AICc é de 145,958.

A análise dos VIF não revelou problemas de multicolinearidade, com todos os valores de VIF abaixo de 2, sugerindo que as variáveis independentes não estão fortemente correlacionadas entre si. Como demonstrado na Tabela 4.5, tanto a variável PER1EDU como a variável PERDIVORC apresentam coeficientes estatisticamente significativos (*p-value* < 0,01), indicando um efeito positivo na prevalência de diabetes.

Tabela 4.7 *Resultados do modelo OLS. Apresenta os coeficientes estimados, com seus erros-padrão, estatísticas-t, probabilidades associadas para cada variável e o VIF.*

Variable	Coefficient	StdError	t-Statistic	Probability	Robust_SE	Robust_t	Robust_Pr	VIF
Intercept	-4.709	2.648	-1.778	0.086	2.094	-2.249	0.033	-------
PER1EDU	0.319	0.053	6.027	0.000	0.036	8.926	0.000	1.083
PERDIVORC	0.787	0.245	3.208	0.003	0.224	3.529	0.001	1.083

O mapa de resíduos apresentado na Figura 4.38 indica que o modelo OLS não capta todas as variações espaciais, o que é comum neste tipo de modelação. Idealmente, num modelo perfeito, os resíduos deveriam estar distribuídos de forma aleatória, sem qualquer padrão espacial evidente. No entanto, ao analisar o mapa de resíduos gerado, verifica-se que existem áreas onde os resíduos não estão aleatoriamente distribuídos, mas sim agrupados espacialmente (Moran's I = 0,359; z = 2,97; P < .003). Este agrupamento sugere que o modelo OLS pode não estar a recolher totalmente todas as variações espaciais na prevalência de diabetes.

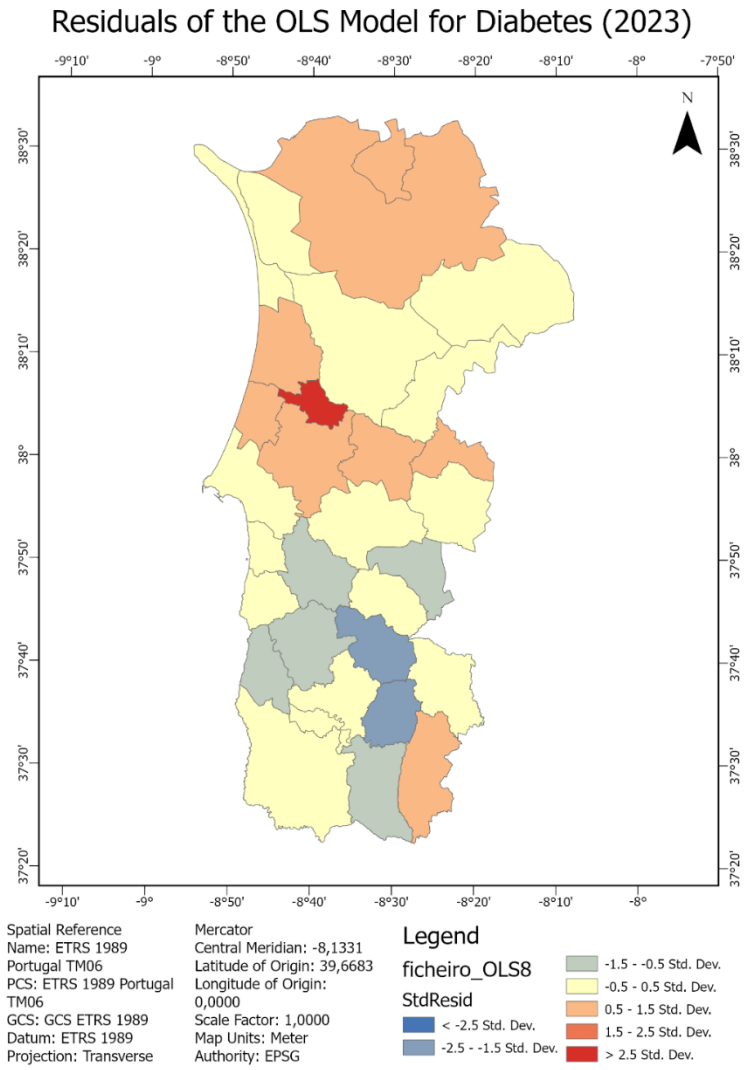


Figura 4.38 *Distribuição espacial dos Resíduos do Modelo OLS para a prevalência de diabetes em 2023.*

Esta análise sublinha a importância de explorar técnicas de análise espacial, como o Spatial Lag e a GWR, que levam em consideração a variação espacial e podem aumentar o poder explicativo do modelo, assim como a compreensão da distribuição espacial da prevalência de diabetes da região.

4.5.2. Modelo *Spatial Lag*

Os resultados do modelo *Spatial Lag* indicam que o coeficiente de desfasagem espacial é de $\rho=0.383$, com um valor de probabilidade de $p=0.021$, o que indica que a autocorrelação espacial é estatisticamente significativa e tem um impacto positivo nos valores da variável dependente. Este coeficiente sugere que freguesias vizinhas influenciam-se mutuamente, reforçando a importância de captar a dependência espacial no modelo.

O R^2 do modelo é de 0.655, o que indica que o modelo explica cerca de 65,5% da variação na prevalência de utentes com diabetes nas diferentes freguesias. Este valor, superior ao valor observado no modelo OLS, sugere que o modelo se ajusta adequadamente aos dados e que as variáveis incluídas têm um poder explicativo considerável.

Ao analisar a Figura 4.39, podemos observar a distribuição espacial dos resíduos obtidos através do modelo de *Spatial Lag*. A maioria das áreas apresenta resíduos próximos de zero, sugerindo que o

modelo tem um bom ajustamento a maioria das freguesias. No entanto, algumas freguesias destacam-se a azul escuro e vermelho, indicando, respetivamente, superestimação e subestimação dos valores ajustados em relação aos observados.

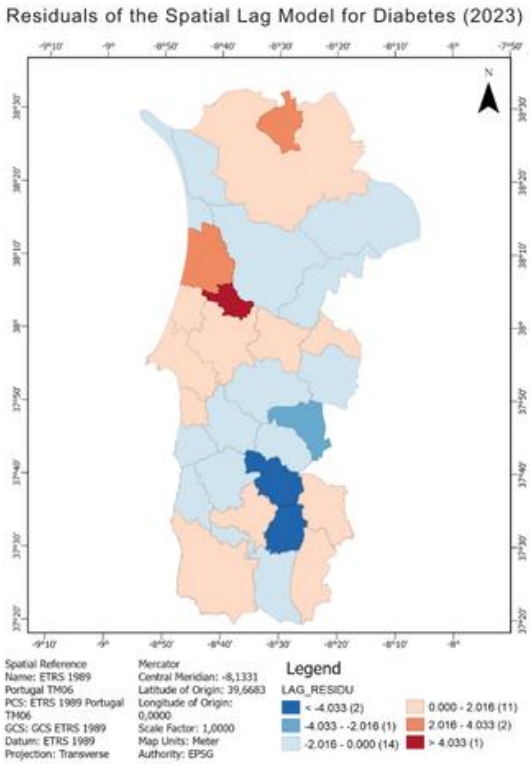


Figura 4.39 *Distribuição espacial dos Resíduos do Modelo Spatial Lag para a prevalência de diabetes em 2023.*

O AICc do modelo é de 139.605, o que confirma que o modelo é mais adequado à estrutura dos dados do que o modelo OLS ($\Delta AICc = 6.353$). Ao incorporar a vertente espacial com o modelo *Spatial Lag*, foi possível obter de forma mais precisa as interações espaciais entre as freguesias, permitindo compreender melhor a realidade da distribuição da prevalência de utentes com diabetes e as dinâmicas entre os fatores explicativos, reforçando a importância de considerar a dependência espacial nas análises.

4.5.3. Modelo GWR

Assim, foi construído um modelo GWR utilizando as mesmas variáveis independentes: "Percentagem de População com o Primeiro Ciclo" (PER1EDU) e "Percentagem de População cujo Estado Civil é Divorciado" (PERDIVORC). O objetivo principal do modelo GWR foi captar a variação espacial das relações entre as variáveis independentes e a prevalência de diabetes, permitindo uma análise mais detalhada das variações locais.

Para o modelo, foram obtidos valores para os coeficientes de Determinação, R^2 e R^2 ajustado, o valor para o AICc, um histograma de distribuição dos Resíduos, um gráfico de resíduos estandardizados versus valores preditos e um mapa coroplético da distribuição dos resíduos. Estes resultados permitem avaliar a adequação das variáveis ao modelo, identificar possíveis vieses sistemáticos, analisar a distribuição dos erros e comparar a eficácia de diferentes modelos.

Para o modelo GWR, o R^2 ajustado foi de 0,7383, com uma variação local entre 0,5805 e 0,8173. Isso indica que aproximadamente 73,8% da variabilidade na prevalência de diabetes é explicada pelas variáveis independentes selecionadas, considerando a variação espacial. Em comparação, o R^2 ajustado

do modelo OLS foi de 0,5515, ou seja, 55,2%. Assim, o modelo GWR demonstra uma melhoria de 18% na qualidade do ajuste dos dados.

Adicionalmente, o valor de AICc do modelo GWR é 60,164, significativamente inferior ao obtido no modelo de Spatial Lag ($\Delta AICc = 79.441$). Um AICc mais baixo indica que o modelo GWR proporciona um melhor equilíbrio entre ajuste e complexidade do modelo, mostrando que este modelo se ajusta melhor aos dados sem sobreajuste.

Para a variável PER1EDU, os coeficientes variam entre 0,32 e 0,65. Esta variação, constatada na Figura 4.40, sugere que o impacto do nível de educação (1° Ciclo) sobre a prevalência de diabetes é positivo em todas as freguesias, embora a intensidade deste impacto varie. As freguesias a verde escuro representam coeficientes GWR positivos mais elevados. Isto significa que, nas freguesias da área central, há uma associação positiva forte entre a baixa escolaridade e a prevalência de diabetes. Noutras palavras, em regiões onde a cor verde predomina, uma maior percentagem de pessoas com baixa escolaridade está fortemente associada a uma maior prevalência de diabetes. Por outro lado, nas freguesias a roxo escuro estão representados coeficientes GWR positivos, mas de magnitude menor, indicando que há uma associação positiva entre a baixa escolaridade e a prevalência de diabetes, mas essa associação é fraca.

Em contrapartida, os coeficientes para PERDIVORC mostram uma variação mais acentuada, oscilando entre -0,35 e 0,42. Esta amplitude de valores, que inclui tanto coeficientes negativos quanto positivos, indica que o impacto de ser divorciado na prevalência de diabetes é mais heterogéneo. Assim nas freguesias situadas no norte do Litoral Alentejano, assinaladas a roxo, indicam coeficientes GWR negativos. Isto é, uma maior percentagem de pessoas divorciadas está associada a uma menor prevalência de diabetes. Portanto o divórcio tem uma relação inversa com a prevalência de diabetes, sugerindo que quanto maior o número de pessoas divorciadas, menor a prevalência de diabetes. Para os coeficientes GWR positivos, freguesias a verde, há uma associação positiva entre a percentagem de pessoas divorciadas e a prevalência de diabetes, mostrando que uma maior percentagem de pessoas divorciadas está associada a uma maior prevalência de diabetes. Esta variabilidade espacial não é captada no modelo OLS, onde um único coeficiente global é aplicado a todas as freguesias, ignorando as nuances locais.

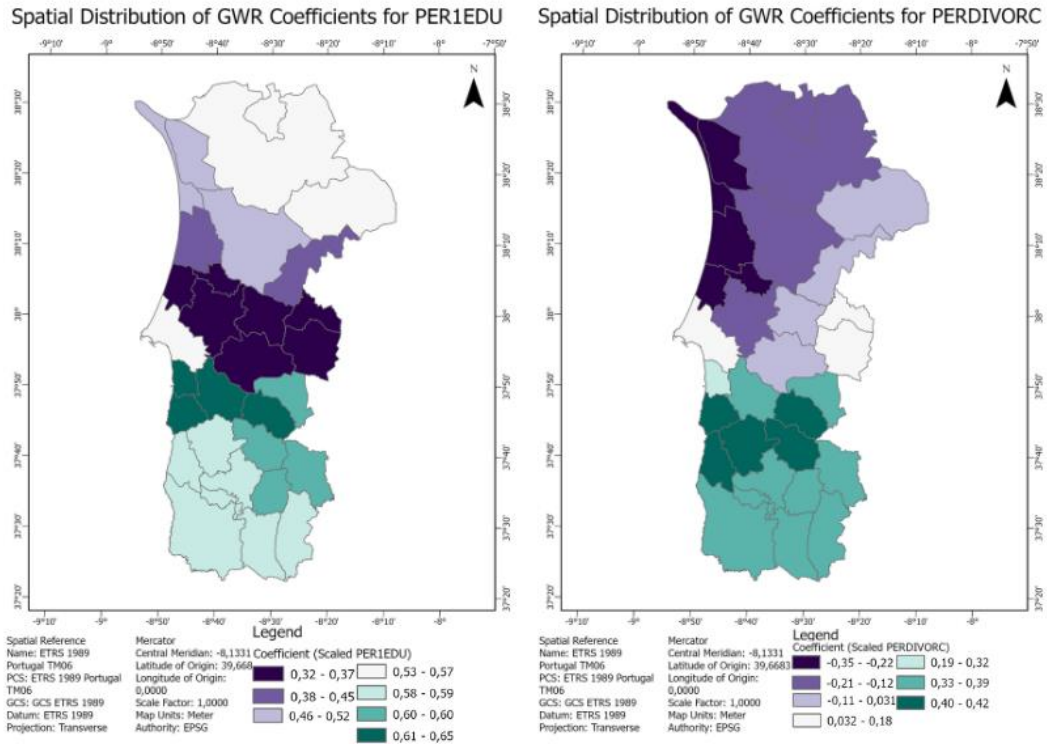


Figura 4.40 *Distribuição Espacial dos Coeficientes GWR para as variáveis PER1EDU (à esquerda) e PERDIVORC (à direita).*

A análise do mapa de resíduos do modelo GWR, apresentado na Figura 4.41, para a prevalência de diabetes revela uma variação espacial no desempenho do modelo. A maioria das regiões apresenta resíduos de ±1,5 desvios padrão, indicando uma previsão razoavelmente precisa. Contudo, destacam-se áreas específicas onde o modelo apresenta discrepâncias mais significativas: na freguesia de Relíquias observam-se resíduos muito negativos (sobrestimação), enquanto em algumas áreas do centro e nordeste, registam-se resíduos positivos mais elevados (subestimação). Esta heterogeneidade espacial nos resíduos (Moran's I = 0,076; z = 0,808; P < 0.5) sublinha a importância de considerar fatores locais na modelação da prevalência de diabetes e demonstra a vantagem do GWR sobre modelos globais, como o OLS e o Spatial Lag, ao detetar estas variações geográficas. Tal padrão sugere que existem fatores locais adicionais, não contemplados no modelo atual, que influenciam a prevalência de diabetes de forma diferenciada ao longo do território.

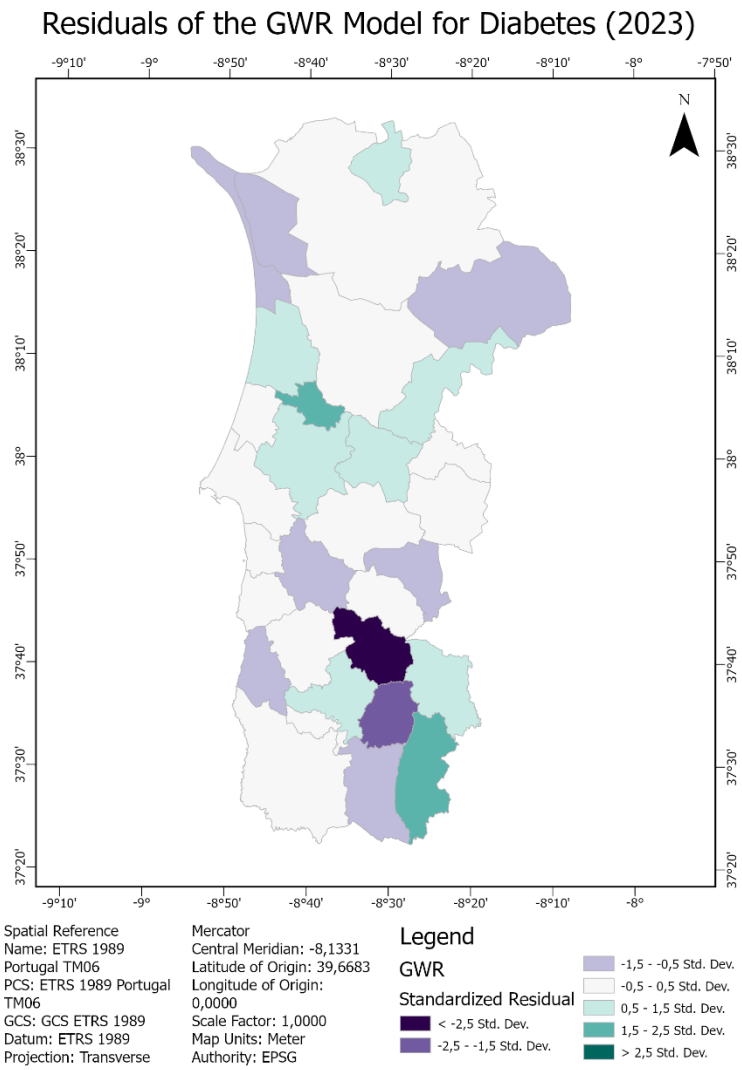


Figura 4.41 *Resíduos do Modelo GWR para Diabetes em 2023 por Freguesia.*

Conforme ilustrado na Figura 4.42, os valores do *Condition Number* variam de 0,55 a 2,2, permanecendo dentro de limites aceitáveis (<30). Estes resultados sugerem que, apesar de existir alguma variação espacial na estabilidade do modelo, as variáveis utilizadas não apresentam multicolinearidade. Assim, pode-se afirmar que o modelo é fiável em todas as freguesias analisadas.

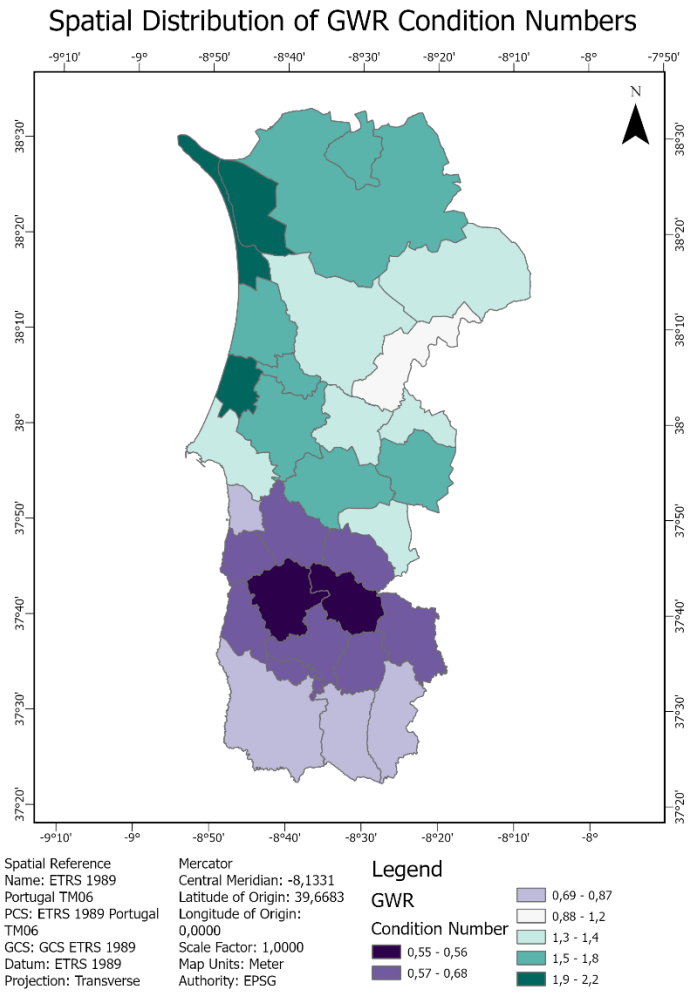


Figura 4.42 *Distribuição Espacial dos Condition Number do Modelo GWR.*

4.5.4. Previsão para 2026

Com o propósito de obter uma projeção da prevalência de diabetes para o ano de 2026, foi utilizado um modelo de previsão baseado em séries temporais. A análise espacial, apresentada na Figura 4.43, permite identificar as áreas onde se espera que a prevalência de diabetes sofra alterações, tanto positivas quanto negativas. Esta previsão é fundamental para permitir um planeamento antecipado, direcionando as intervenções de saúde pública na região de forma a suprir necessidades futuras.

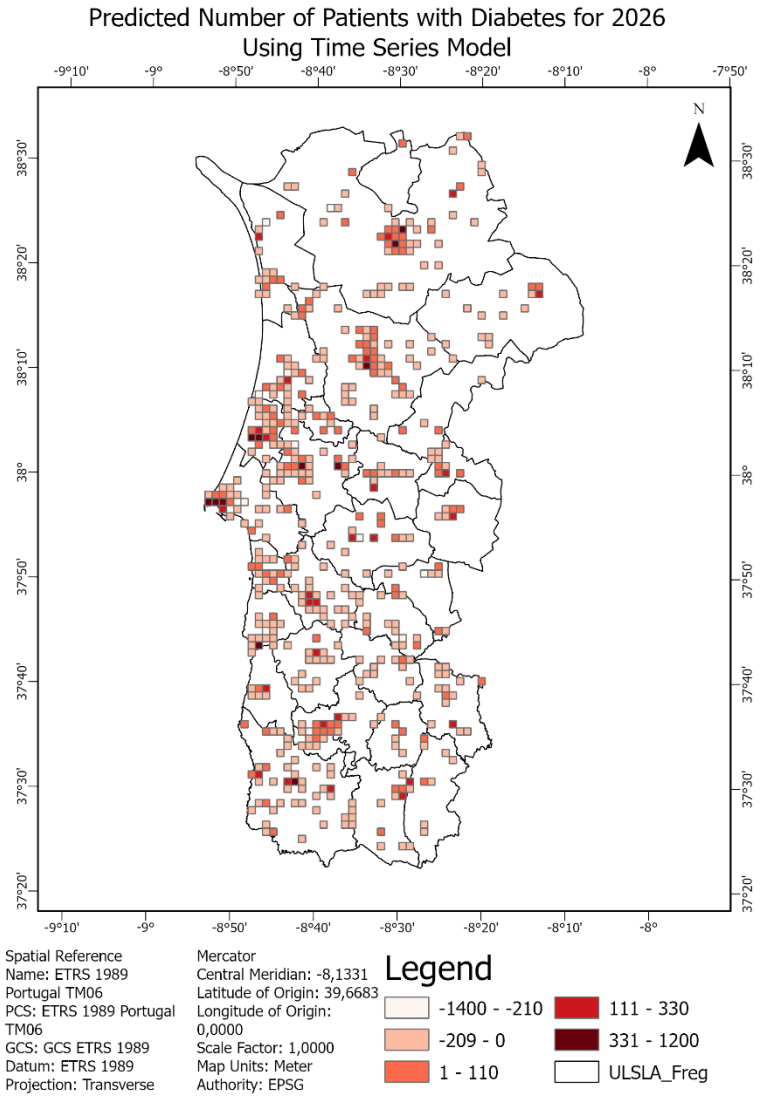


Figura 4.43 *Estimativa de novos casos de Diabetes para 2026, através do Modelo de Séries Temporais.*

O mapa permite observar uma distribuição heterogénea da incidência de diabetes ao longo da região, com algumas áreas a apresentarem valores mais elevados. As áreas mais escuras no mapa, que representam valores entre 331 e 1200 casos, destacam-se como zonas onde se prevê um aumento acentuado de novos casos de diabetes. Estas áreas estão concentradas principalmente nos maiores centros urbanos, onde se encontra maior parte da população, indicando que essas regiões poderão enfrentar maiores desafios no controlo e gestão da diabetes nos próximos anos.

Por outro lado, as áreas de coloração mais clara (-1400 a 0 casos) refletem uma redução de casos. Estas zonas, predominantemente localizadas nas redondezas das áreas urbanas, podem indicar uma deslocação populacional, seja devido ao abandono das zonas rurais ou ao envelhecimento populacional, com o consequente aumento de mortalidade.

Esta análise oferece uma importante ferramenta para a antecipação das necessidades em termos de recursos de saúde, com intervalo de confiança de 90%, proporcionando uma base sólida para o planeamento estratégico e a alocação eficaz de recursos nos próximos anos.

5. Discussão

A presente discussão aborda os resultados obtidos na análise da prevalência de diabetes tipo 2 no Litoral Alentejano, destacando a influência de fatores demográficos e socioeconómicos na distribuição espacial da patologia. Serão também exploradas as implicações destes achados para a formulação de políticas de saúde pública mais eficazes e direcionadas.

Um dos principais desafios identificados, para a região do Litoral Alentejano, foi a disparidade na acessibilidade aos cuidados de saúde no Litoral Alentejano. A menor densidade de centros de saúde e a presença de apenas um hospital na região, em comparação com outras áreas de Portugal, podem estar a influenciar as variações na prevalência de DMT2. Esta situação suscita preocupações sobre a equidade no acesso aos cuidados de saúde e o impacto dessas desigualdades na gestão de doenças crónicas como a diabetes. Para reduzir essas disparidades, as autoridades de saúde devem considerar a implementação de estratégias como unidades móveis de saúde ou a utilização da telemedicina, especialmente em áreas menos servidas, para garantir que todos os residentes tenham acesso adequado aos cuidados necessários.

Os resultados da análise espacial da prevalência de DMT2 no Litoral Alentejano revelaram uma distribuição geográfica heterogénea da doença, com uma concentração acentuada nas freguesias de Alcácer do Sal, Grândola, Santo André, Sines, Santiago do Cacém e São Teotónio. Esta distribuição sugere que fatores locais, como demográficos e socioeconómicos, desempenham um papel crucial na prevalência da doença. A concentração em áreas urbanas e costeiras pode estar relacionada com estilos de vida específicos dessas regiões, como maior sedentarismo e dietas menos saudáveis, influenciando assim a incidência de DMT2.

A análise temporal mostrou um aumento preocupante da prevalência da DMT2 entre 2017 e 2023, especialmente nas áreas já mencionadas. Este dado reforça a necessidade de políticas de saúde pública adaptadas às especificidades locais, com foco na prevenção e gestão da doença nas áreas mais afetadas. Desenvolver estratégias de intervenção, como campanhas de sensibilização e programas de rastreio em áreas de maior risco, pode ser essencial para travar o aumento da prevalência.

A relação observada entre DSS e a prevalência de DMT2 também é significativa. A correlação observada entre a doença e fatores como baixos níveis de educação e estado civil (divorciado) destaca a importância de considerar os contextos sociais e culturais na formulação de políticas de saúde. Estes determinantes devem ser incorporados em intervenções que visem reduzir as desigualdades na saúde.

No que diz respeito aos modelos de análise, GWR demonstrou ser mais eficaz que o modelo OLS na deteção das variações espaciais na prevalência da DMT2. A aplicação de um modelo espacial revelou-se essencial para este tipo de análises, uma vez que permite capturar de forma mais eficaz as variações geográficas que os modelos tradicionais, como o OLS, não conseguem detetar, oferecendo assim uma compreensão mais detalhada e ajustada da prevalência da DMT2 nas diferentes freguesias.

A aplicação do modelo OLS revelou que variáveis como o nível de educação (PER1EDU) e o estado civil (PERDIVORC) têm uma influência na prevalência de diabetes. No entanto, os resíduos do modelo indicam que o OLS não conseguiu captar todas as variações espaciais, sugerindo a existência de padrões geográficos que requerem uma abordagem mais refinada. A análise espacial dos resíduos, através do índice de Moran's I, confirmou a existência de agrupamentos espaciais não observados pelo modelo global, evidenciando a necessidade de técnicas que considerem a variação espacial.

O modelo GWR, por outro lado, demonstrou uma capacidade superior para captar essas variações, como evidenciado pelos valores ajustados de R^2 e AICc significativamente melhores em comparação com o modelo OLS. Esta melhoria sublinha a importância de considerar a diversidade espacial ao

estudar a distribuição da diabetes, uma vez que as relações entre variáveis podem variar substancialmente entre freguesias. O GWR permitiu identificar variações locais importantes, como a influência positiva e variável do nível de educação e do estado civil sobre a prevalência de diabetes, oferecendo clareza para políticas de saúde pública mais direcionadas.

A formulação de políticas de saúde pública, direcionadas para o Litoral Alentejano, deve ter em consideração os resultados que apontam para uma correlação significativa entre baixos níveis de escolaridade e o estado civil de divorciado com a prevalência de diabetes. O estudo de Kposowa et al. (2021) indica que indivíduos divorciados ou separados, em particular os homens, enfrentam um risco mais elevado de mortalidade associada à diabetes [79]. Estes dados sublinham a importância de implementar programas de apoio psicossocial e de acompanhamento em saúde mental para indivíduos divorciados, com o objetivo de minimizar os efeitos do stress relacionado com o divórcio na sua saúde física. Paralelamente, a literacia em saúde reduzida, comum nas pessoas com níveis educacionais mais baixos, também foi identificada como um fator de risco relevante para o controlo inadequado da diabetes [80]. Programas de educação comunitária focados na prevenção da diabetes e na promoção de hábitos de vida saudáveis são essenciais para a redução destes riscos.

Em resumo, os resultados sugerem que a prevalência de DMT2 no Litoral Alentejano é influenciada por uma combinação de fatores demográficos e socioeconómicos, que variam espacialmente. Esta constatação realça a necessidade de intervenções de saúde pública específicas para cada localidade, visando uma resposta mais eficaz à gestão da DMT2 na região. Além disso, a utilização de técnicas de análise espacial avançadas, como a GWR, mostrou-se essencial para captar as nuances locais e deve ser considerada em estudos futuros.

6. Conclusão

O presente estudo cumpre integralmente os objetivos inicialmente definidos, ao analisar o padrão espacial da prevalência da diabetes nas freguesias sob a gestão da ULSLA, recorrendo a modelos de regressão que integram os principais indicadores socioeconómicos. Foram identificados os fatores mais relevantes que influenciam a variável dependente, e os padrões geográficos correspondentes foram evidenciados através da utilização de modelos de regressão espacial, como o GWR e o Spatial Lag, o que permitiu captar a variação local da influência de cada indicador. Foi realizada uma previsão, recorrendo a um modelo de regressão, com bom nível de ajuste. A utilização de um modelo de regressão espacial local possibilitou uma análise aprofundada da influência dos fatores mais impactantes em cada freguesia. Desta forma, o estudo cumpre os seus objetivos, oferecendo uma visão abrangente sobre a distribuição geográfica e os determinantes da prevalência da diabetes na região.

Este estudo reafirma a importância de uma abordagem localizada e personalizada na gestão e prevenção da DMT2 no Litoral Alentejano. Os padrões geográficos e os determinantes sociais de saúde identificados ao longo da investigação evidenciam que fatores contextuais específicos desempenham um papel determinante na prevalência da DMT2 nesta região. Este facto sublinha a necessidade de estratégias de saúde pública adaptadas às particularidades locais, capazes de responder de forma mais eficaz às necessidades da população.

Uma recomendação que emerge deste estudo é a criação de uma base de dados integrada para o SNS, que facilite uma navegação e manipulação mais eficiente dos dados de saúde. Atualmente, a dispersão de fontes e a falta de interoperabilidade constituem desafios significativos para a integração dos dados de saúde. A centralização dessas informações numa base de dados a nível nacional, seria benéfica a muitos níveis, permitindo criar um historial do utente, acessível aos prestadores de saúde e em qualquer instituição, ao longo da vida do indivíduo.

Sugere-se que investigações futuras incluam dados provenientes das urgências hospitalares, especialmente dado o limitado número de recursos hospitalares na região. A criação de um modelo específico para a analisar o uso dos serviços de urgência pode proporcionar uma visão mais aprofundada sobre o acesso e a gestão de condições crónicas, no Litoral Alentejano. Além disso, a inclusão de dados sobre poluição ambiental por freguesia poderá enriquecer a análise dos determinantes de saúde, oferecendo uma compreensão mais abrangente dos fatores que influenciam a prevalência da DMT2.

A validação das previsões para o ano de 2026, através da comparação com dados reais, será fundamental para aferir a precisão do modelo desenvolvido. Este processo permitirá ajustar e refinar as previsões futuras, garantindo que as intervenções de saúde pública sejam cada vez mais direcionadas e eficazes.

Em suma, este projeto estabelece uma base sólida para uma abordagem mais informada e personalizada na gestão da DMT2 no Litoral Alentejano. A implementação das recomendações propostas, conciliadas com o desenvolvimento contínuo de ferramentas de análise espacial e a integração de novos dados, tem o potencial de transformar a forma como os desafios de saúde pública são enfrentados nesta região. Este modelo, além de beneficiar a população do Litoral Alentejano, pode servir de referência para outras regiões de Portugal, promovendo um sistema de saúde mais eficiente, equitativo e adaptado às necessidades locais.

7. Referências.

[1] N. M. Capeletti, C. A. F. Amado, and S. P. Santos, "Performance assessment of primary health care services using data envelopment analysis and the quality-adjusted malmquist index," *Journal of the Operational Research Society*, 2023, doi: 10.1080/01605682.2023.2197928.

[2] Banco Mundial (n.d), "Life expectancy at birth, total (years) - Portugal," ttps://data.worldbank.org/indicator/SP.DYN.LE00.IN?end=2022&locations=PT&start=2012.

[3] S. M. Shortell, R. R. Gillies, and D. A. Anderson, "The new world of managed care: Creating organized delivery systems," *Health Aff*, vol. 13, no. 5, pp. 46–64, 1994, doi: 10.1377/hlthaff.13.5.46.

[4] O. Grone and M. Garcia-Barbero, "Integrated care: A position paper of the WHO European office for integrated health care services," *Int J Integr Care*, vol. 1, no. 1, 2001, [Online]. Available: http://www.ijic.org/

[5] World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), "Towards universal health coverage and the sustainable development goals A VISION FOR PRIMARY HEALTH CARE IN THE 21ST CENTURY," 2018.

[6] OECD, *Geographic Variations in Health Care*. in OECD Health Policy Studies. OECD, 2014. doi: 10.1787/9789264216594-en.

[7] A. C. Gatrell and S. J. Elliott, *Geographies of Health - An Introduction*, 3rd ed. Wiley, Blackwell, 2015.

[8] Ministério da Saúde, "Retrato da Saúde 2018," 2018.

[9] V. Ribeiro, P. Remoaldo, and J. Gutiérrez, "Acessibilidade e SIG no planeamento em saúde : Uma abordagem baseada em modelos de alocação-localização," 2015.

[10] N. Wang *et al.*, "Lung cancer and particulate pollution: A critical review of spatial and temporal analysis evidence," Jul. 01, 2018, *Academic Press Inc.* doi: 10.1016/j.envres.2018.03.034.

[11] P. Santana, *Introdução à Geografia da Saúde Território, Saúde e Bem-Estar*, vol. 0. Imprensa da Universidade de Coimbra, 2014. [Online]. Available: https://doi.org/http://dx.doi.org/10.14195/978-989-26-0727-6

[12] P. Pednekar and A. Peterson, "Mapping pharmacy deserts and determining accessibility to community pharmacy services for elderly enrolled in a State Pharmaceutical Assistance Program.," *PLoS One*, vol. 13, no. 6, Jun. 2018, doi: 10.1371/journal.pone.0198173.

[13] E. Hatef *et al.*, "The Impact of Social Determinants of Health on Hospitalization in the Veterans Health Administration," *Am J Prev Med*, vol. 56, no. 6, pp. 811–818, Jun. 2019, doi: 10.1016/j.amepre.2018.12.012.

[14] P. Sousa *et al.*, "COMPRIME - COnhecer Mais PaRa Intervir MElhor: Preliminary Mapping of Municipal Level Determinants of COVID-19 Transmission in Portugal at Different Moments of the 1st Epidemic Wave," *Portuguese Journal of Public Health*, vol. 38, no. 1, pp. 18–25, Jun. 2021, doi: 10.1159/000514334.

[15] International Health Conference, “Minutes of the Technical Preparatory Committee for the International Health Conference held in Paris from 18 March to 5 April 1946,” 1947. [Online]. Available: https://apps.who.int/iris/handle/10665/85572

[16] C. Albuquerque and C. Ferreira de Oliveira, “Saúde e doença: significações e perspectivas em mudança,” 2002, [Online]. Available: http://hdl.handle.net/10400.19/635

[17] S. Shaw, R. Rosen, and B. Rumbold, “What is integrated care? Research report,” 2011. [Online]. Available: www.nuffieldtrust.org.uk/integratedcare

[18] G. D. Armitage, E. Suter, N. D. Oelke, and C. E. Adair, “Health systems integration: state of the evidence,” *Int J Integr Care*, vol. 9, 2009, [Online]. Available: http://www.ijic.org/

[19] WORLD HEALTH ORGANIZATION (WHO), “Integrating health services,” 2018.

[20] M. M. Byrne and C. M. Ashton, “Incentives for Vertical Integration in Healthcare: The Effect of Reimbursement Systems,” *Journal of Healthcare Management*, vol. 44, no. 1, 1999, [Online]. Available: https://journals.lww.com/jhmonline/fulltext/1999/01000/incentives_for_vertical_integration_in_healthcare_.7.aspx

[21] D. A. Conrad and S. M. Shortell, “Integrated health systems: promise and performance,” *Front Health Serv Manage*, vol. 13, no. 1, pp. 3—40; discussion 57—8, 1996, [Online]. Available: http://europepmc.org/abstract/MED/10159629

[22] OECD, *Health at a Glance 2023*. in Health at a Glance. OECD, 2023. doi: 10.1787/7a7afb35-en.

[23] Entidade Reguladora da Saúde (ERS), “Estudo sobre a Organização e Desempenho das Unidades Locais de Saúde – Relatório Preliminar I,” 2015.

[24] P. LEATT, M. GUERRIERE, and G. PINK, “Towards a Canadian model of integrated healthcare,” *Healthcare Papers*, 2000.

[25] J. Hänninen, J. Takala, and S. Keinänen-Kiukaanniemi, “Good continuity of care may improve quality of life in Type 2 diabetes,” 2001. [Online]. Available: www.elsevier.com/locate/diabres

[26] G. Arch, Mainous Ill, R. J. Koopman, J. M. Gill, R. Baker, and W. S. Pearson, “Relationship Between Continuity of Care and Diabetes Control: Evidence From the Third National Health and Nutrition Examination Survey,” *Am J Public Health*, vol. 94, no. 1, 2004.

[27] D. A. Christakis, C. Feudtner, C. Pihoker, and F. A. Connell, “Continuity and Quality of Care for Children With Diabetes Who Are Covered by Medicaid,” *AMBULATORY PEDIATRICS*, vol. 1, no. 2, 2001.

[28] J. Lloyd and S. Wait, “Integrated care : a guide for polciymakers,” 2006.

[29] PopulationPyramid.net, “Pirâmides Populacionais de Portugal desde 1950 até 2100,” https://www.populationpyramid.net/pt/portugal/2024/.

[30] P. Braveman and L. Gottlieb, “The Social Determinants of Health: It’s Time to Consider the Causes of the Causes,” *Public Health Reports®*, vol. 129, no. 1_suppl2, pp. 19–31, Jan. 2014, doi: 10.1177/00333549141291S206.

[31] R. J. Lawrence, “Understanding Environmental Quality Through Quality of Life (QOL) Studies,” in *Encyclopedia of Environmental Health*, J. O. Nriagu, Ed., Burlington: Elsevier, 2011, pp. 518–525. doi: https://doi.org/10.1016/B978-0-444-52272-6.00226-9.

[32] William F. Davenhall and Christopher Kinabrew, “GIS in Health and Human Services,” no. Springer Handbook of Geographic Information, pp. 557–578, 2012, [Online]. Available: https://doi.org/10.1007/978-3-540-72680-7_29

[33] WHO, “Webinar Launch of the WHO European Geospatial Coordination Hub,” WHO. [Online]. Available: https://dreambroker.com/channel/usef75hm/uke55kdr

[34] World Health Organization, “Connecting maps, data and health: Introducing the WHO European Geospatial Coordination Hub,” World Health Organization. [Online]. Available: https://www.who.int/europe/news/item/05-01-2024-connecting-maps--data-and-health--introducing-the-who-european-geospatial-coordination-hub

[35] G. M. Foody, “GIS: health applications,” *Progress in Physical Geography: Earth and Environment*, vol. 30, no. 5, pp. 691–695, Oct. 2006, doi: 10.1177/0309133306071152.

[36] International Diabetes Federation, “IDF Diabetes Atlas 10th edition,” 2021. [Online]. Available: www.diabetesatlas.org

[37] M. Barreto *et al.*, “Prevalence, awareness, treatment and control of diabetes in Portugal: Results from the first National Health examination Survey (INSEF 2015),” *Diabetes Res Clin Pract*, vol. 140, pp. 271–278, 2018, doi: https://doi.org/10.1016/j.diabres.2018.03.052.

[38] Programa Nacional para a Diabetes (PND), “PROGRAMA NACIONAL PARA A DIABETES Desafios e Estratégias,” Nov. 2023. [Online]. Available: www.dgs.pt

[39] M. Uusitupa *et al.*, “Prevention of type 2 diabetes by lifestyle changes: A systematic review and meta-analysis,” Nov. 01, 2019, *MDPI AG*. doi: 10.3390/nu11112611.

[40] J. S. Skyler *et al.*, “Differentiation of Diabetes by Pathophysiology, Natural History, and Prognosis,” *Diabetes*, vol. 66, no. 2, pp. 241–255, Dec. 2016, doi: 10.2337/db16-0806.

[41] T. Zhou, X. Liu, Y. Liu, and X. Li, “Meta-analytic evaluation for the spatio-temporal patterns of the associations between common risk factors and type 2 diabetes in mainland China,” *Medicine*, vol. 98, no. 20, 2019, [Online]. Available: https://journals.lww.com/md-journal/fulltext/2019/05170/meta_analytic_evaluation_for_the_spatio_temporal.32.aspx

[42] A. B. Curtis, M. Catherine Kothari, M. Rajib Paul, and E. Connors, “Using GIS and Secondary Data to Target Diabetes-Related Public Health Efforts.” Accessed: Jan. 18, 2024. [Online]. Available: https://journals.sagepub.com/doi/10.1177/003335491312800311

[43] J. Wu, Y. Wang, X. Xiao, X. Shang, M. He, and L. Zhang, “Spatial Analysis of Incidence of Diagnosed Type 2 Diabetes Mellitus and Its Association With Obesity and Physical Inactivity,” *Front Endocrinol (Lausanne)*, vol. 12, Oct. 2021, doi: 10.3389/fendo.2021.755575.

[44] M. Jiwa, O. Gudes, R. Varhol, and N. Mullan, “Impact of geography on the control of type 2 diabetes mellitus: a review of geocoded clinical data from general practice,” *BMJ Open*, vol. 5, p. 9504, 2015, doi: 10.1136/bmjopen-2015.

[45] Y. Krishnamoorthy, S. Rajaa, M. Verma, R. Kakkar, and S. Kalra, “Spatial Patterns and Determinants of Diabetes Mellitus in Indian Adult Population: a Secondary Data Analysis from Nationally Representative Surveys,” *Diabetes Therapy*, vol. 14, no. 1, pp. 63–75, Jan. 2023, doi: 10.1007/s13300-022-01329-6.

[46] L. Wang *et al.*, “Geographical variation in diabetes prevalence and detection in China: Multilevel spatial analysis of 98,058 adults,” *Diabetes Care*, vol. 38, no. 1, pp. 72–81, Jan. 2015, doi: 10.2337/dc14-1100.

[47] N. Nurjannah, K. M. Baker, and D. P. Mashinini, “Spatiotemporal trends in diabetes-related mortality by school district in the state of Michigan, United States,” *Epidemiol Health*, vol. 43, 2021, doi: 10.4178/epih.e2021098.

[48] K. Schwab, *The Fourth Industrial Revolution*. 2016. [Online]. Available: www.weforum.org

[49] L. Anselin, “Local Indicators of Spatial Association—LISA,” *Geogr Anal*, vol. 27, no. 2, pp. 93–115, Apr. 1995, doi: https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.

[50] Instituto da Conservação da Natureza e das Florestas (ICNF), “Parque Natural do Sudoeste Alentejano e Costa Vicentina.” [Online]. Available: https://www.icnf.pt/conservacao/rnapareasprotegidas/parquesnaturais/pnsudoestealentejanoecostavicentina

[51] Instituto Português do Mar e da Atmosfera (IPMA), “Clima de Portugal Continental.” [Online]. Available: https://www.ipma.pt/

[52] Instituto Nacional de Estatística (INE), “Censos 2021- Resultados definitivos.” [Online]. Available: https://www.ine.pt/

[53] Comissão de Coordenação e Desenvolvimento Regional do Alentejo, “Programa Operacional Regional do Alentejo 2014-2020,” Dec. 2018. [Online]. Available: http://www.alentejo.portugal2020.pt/phocadownload/Documentacao/Programa/PORALENTEJO_dez18.pdf

[54] Wikipédia, “Litoral Alentejano.” Accessed: Mar. 17, 2024. [Online]. Available: https://pt.wikipedia.org/wiki/Alentejo_Litoral

[55] Associação de Desenvolvimento do Litoral Alentejano, “Zona de Intervenção.” Accessed: Feb. 10, 2024. [Online]. Available: https://litoralalentejano.pt/associacao/zona-de-intervencao/

[56] ESRI, “Enrich Layer (Analysis).” [Online]. Available: https://pro.arcgis.com/en/pro-app/latest/tool-reference/analysis/enrich-layer.htm

[57] ESRI, “Esri Demographic Data Release Notes: Portugal,” May 2022. [Online]. Available: https://downloads.esri.com/esri_content_doc/dbl/int/PRT_MBR_2021_Release_Notes.pdf

[58] ESRI, “Esri Demographics Regional Data for Portugal.” [Online]. Available: https://doc.arcgis.com/en/esri-demographics/latest/regional-data/portugal.htm

[59] Geoapify, “Geocoding in Python Tutorial.” [Online]. Available: ttps://www.geoapify.com/tutorial/geocoding-python/

[60] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, J. Neyman, Ed., Berkele, CA: Univ.California, 1967, pp. 281–297. [Online]. Available: https://api.semanticscholar.org/CorpusID:6278891

[61] P. Diggle, *Statistical analysis of spatial and spatio-temporal point patterns*, Third edition. in Monographs on statistics and applied probability ; 128. Hoboken: Taylor and Francis, 2014.

[62] ESRI, “Create Space Time Cube By Aggregating Points (Space Time Pattern Mining).” [Online]. Available: https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/create-space-time-cube.htm

[63] ESRI, “Emerging Hot Spot Analysis (Space Time Pattern Mining).” [Online]. Available: https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/emerginghotspots.htm

[64] C. Zhang, W. Lv, P. Zhang, and J. Song, “Multidimensional spatial autocorrelation analysis and it’s application based on improved Moran’s I,” *Earth Sci Inform*, vol. 16, no. 4, pp. 3355–3368, 2023, doi: 10.1007/s12145-023-01090-9.

[65] L. Anselin, I. Syabri, and Y. Kho, “GeoDa: Spatial Data Analysis Lab. GeoDa Center for Geospatial Analysis and Computation.” [Online]. Available: https://geodacenter.github.io/

[66] ESRI, “How Cluster and Outlier Analysis (Anselin Local Moran’s I) works.” [Online]. Available: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-cluster-and-outlier-analysis-anselin-local-m.htm

[67] Samprit Chatterjee and Jeffrey S. Simonoff, “Multiple Linear Regression,” in *Handbook of Regression Analysis*, 2012, pp. 1–21. doi: https://doi.org/10.1002/9781118532843.ch1.

[68] I. Choi, “ECONOMETRICS: by Fumio Hayashi, Princeton University Press, 2000,” *Econ Theory*, vol. 18, no. 4, pp. 1000–1006, 2002, doi: DOI: 10.1017/S0266466602004115.

[69] C. M. Jarque and A. K. Bera, “Efficient tests for normality, homoscedasticity and serial independence of regression residuals,” *Econ Lett*, vol. 6, no. 3, pp. 255–259, 1980, doi: https://doi.org/10.1016/0165-1765(80)90024-5.

[70] R. Koenker and G. Bassett, “Robust Tests for Heteroscedasticity Based on Regression Quantiles,” *Econometrica*, vol. 50, no. 1, pp. 43–61, 1982, [Online]. Available: https://EconPapers.repec.org/RePEc:ecm:emetrp:v:50:y:1982:i:1:p:43-61

[71] C. M. HURVICH and C.-L. TSAI, “Regression and time series model selection in small samples,” *Biometrika*, vol. 76, no. 2, pp. 297–307, Jun. 1989, doi: 10.1093/biomet/76.2.297.

[72] J. Lesage and R. K. Pace, *Introduction to Spatial Econometrics*, 1st ed. Chapman and Hall/CRC, 2009.

[73] J. P. Elhorst, *From Cross-Sectional Data to Spatial Panels*. Springer, 2014. [Online]. Available: http://www.springer.com/series/10096

[74] C. Brunsdon, A. S. Fotheringham, and M. E. Charlton, “Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity,” *Geogr Anal*, vol. 28, no. 4, pp. 281–298, Oct. 1996, doi: https://doi.org/10.1111/j.1538-4632.1996.tb00936.x.

[75] A. Fotheringham, C. Brunsdon, and M. Charlton, “Geographically Weighted Regression: The Analysis of Spatially Varying Relationships,” *John Wiley & Sons*, vol. 13, Jan. 2002.

[76] S. Matthews and T.-C. Yang, “Mapping the results of local statistics: Using geographically weighted regression,” *Demogr Res*, vol. S13, no. 6, pp. 151–166, Mar. 2012, doi: 10.4054/DemRes.2012.26.6.

[77] ESRI, “Deep Learning Frameworks,” Github. [Online]. Available: https://github.com/Esri/deep-learning-frameworks?tab=readme-ov-file

[78] M. Adachi *et al.*, “Effects of lifestyle education program for type 2 diabetes patients in clinics: a cluster randomized controlled trial,” 2013. [Online]. Available: http://www.biomedcentral.com/1471-2458/13/467

[79] A. J. Kposowa, D. A. Ezzat, and K. Breault, “Diabetes mellitus and marital status: Evidence from the national longitudinal mortality study on the effect of marital dissolution and the death of a spouse,” *Int J Gen Med*, vol. 14, pp. 1881–1888, 2021, doi: 10.2147/IJGM.S307436.

[80] C. J. Steele *et al.*, “Education achievement and type 2 diabetes - What mediates the relationship in older adults? Data from the ESTHER study: A population-based cohort study,” *BMJ Open*, vol. 7, no. 4, Apr. 2017, doi: 10.1136/bmjopen-2016-013569.

ANEXOS

Anexo I- Tabelas Descritivas

Tabela Anexo I. 1 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2017.*

2017	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8397	7740	7741	353	270	4909	7177	6969	7220	0	7125
mean	69.278	75.734	139.869	234.379	120.552	29.216	49.708	105.775	7.499	-	148.508
std	11.908	11.300	19.796	175.546	45.503	6.874	49.499	44.776	16.345	-	1084.952
min	1	7	71	1.069	80	13.13	0	0	0	-	0
25%	62	69	127	132	97	25.68	40	79	6	-	87
50%	70	76	139	176	105	28.48	47	102.4	6.5	-	118
75%	78	82	150	276	116	31.89	57	128	7.4	-	162
max	101	132	240	988	394	269.97	4045	1958.6	905	-	91476

Tabela Anexo I. 2 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2018.*

2018	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8493	7835	7836	4115	3782	4709	7242	7128	7296	0	7234
mean	69.721	75.162	138.283	199.308	118.001	127.335	50.896	98.764	7.203	-	137.670
std	11.806	11.210	19.806	113.323	42.110	4768.309	17.168	36.973	6.245	-	99.179
min	8	7	1	1.035	1	3.21	0	0	0	-	0
25%	62	68	126	127	97	25.64	41	74.2	6.1	-	89
50%	70	75	138	166	104	28.52	49	95	6.6	-	119
75%	79	82	149	233	116	31.83	58	120	7.4	-	164.75
max	102	138	247	979	523	241964.94	554	431.8	262	-	5454

Tabela Anexo I. 3 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2019.*

2019	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8555	7863	7696	5699	5617	3343	6627	6720	7337	7203	7038
mean	69.963	75.100	137.779	274.903	164.268	292.493	50.348	372.010	7.170	0.789	139.002
std	11.848	11.160	19.332	167.031	85.061	8872.062	16.027	527.762	4.602	4.523	141.374
min	9	21	67	1.012	1	3.45	0	0	0	0.001	0
25%	62	68	125	147	101	25.77	40	82.4	6.1	0.096	89
50%	71	75	137	233	117	28.62	48	113	6.6	0.71	120
75%	79	82	149	361	210	32.05	58	682	7.4	0.94	164
max	103	145	250	993	900	343934.91	379	14381	150	262	10009

Tabela Anexo I. 4 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2020.*

2020	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8560	3579	7549	4155	3967	2928	7158	7107	7258	7101	7173
mean	70.225	76.263	137.803	228.084	133.031	363.143	50.939	117.569	7.303	1.299	138.047
std	11.831	11.130	19.344	131.809	60.148	10465	15.066	1945.466	7.128	6.262	78.870
min	2	20	57	1.015	1	15.61	0	0	0	0	0
25%	63	70	125	136	98	25.71	41	70.6	6.1	0.71	90
50%	71	77	137	188	107	28.575	49	91	6.6	0.86	120
75%	79	83	149	280	139	32.05	58	115	7.4	1.06	164
max	104	125	245	996	977	364480.6	351	164073.8	295	289	2009

Tabela Anexo I. 5 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2021.*

2021	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8840	7738	7899	3359	3128	2563	7392	7346	7163	7328	7402
mean	70.184	75.560	136.720	227.797	136.778	588.319	50.913	94.040	7.201	1.370	140.129
std	11.918	10.844	18.639	132.538	60.444	14431.624	15.374	36.067	4.735	7.935	82.263
min	3	0	1	1.003	1	3.45	0.56	0.4	0	0	1.3
25%	63	69	125	134	99	25.745	41	70	6.1	0.73	91
50%	71	76	136	189	109	28.67	49	90	6.7	0.88	122
75%	79	82	148	277	182	32.24	58	115.6	7.5	1.09	166
max	102	130	238	996	802	475555.56	466	300.4	164	438	1603

Tabela Anexo I. 6 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2022.*

2022	Idade	Tad	Tas	Glicemia Ca	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	9067	8005	8005	4949	4862	2203	7548	7503	7649	7516	7564
mean	70.270	75.287	136.844	236.342	145.848	802.826	50.329	92.523	7.191	1.544	140.422
std	12.038	10.871	18.764	138.935	67.310	16636.625	17.012	37.615	5.249	7.904	91.625
min	7	5	60	1.008	1.15	5.95	0.56	0.4	0	0	1.8
25%	63	69	125	135	99	25.755	40	67.6	6.1	0.73	91
50%	71	75	136	197	111	28.89	48	89	6.6	0.88	121
75%	79	82	148	297	197	32.37	58	114.4	7.4	1.09	166
max	103	121	235	998	511	475555.56	522	452	209	237	2534

Tabela Anexo I. 7 *Estatísticas Descritivas das Variáveis Clínicas antes do processamento de dados em 2023.*

2023	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	9344	8247	7768	5534	5510	2583	7941	6982	8012	7899	7937
mean	70.427	75.694	136.269	243.183	149.192	1011.361	51.721	424.639	7.178	1.012	136.131
std	12.108	11.057	19.487	148.438	71.234	18162.613	15.420	542.040	5.199	6.278	87.914
min	17	10	0	1.02	1.041	2.45	0.56	0.4	0	0	5.2
25%	63	69	123	137	100	25.595	42	78	6.1	0.113	87
50%	71	76	135	200	112	28.51	50	126.3	6.6	0.74	117
75%	79	82	148	301	199	31.98	59	763.5	7.4	0.98	159
max	104	135	241	994	603	475555.56	376	16806	167	237	2164

Tabela Anexo I. 8 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2017.*

2017	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8397	8394	8394	8161	7957	8394	8394	8394	8394	0	8394
mean	69.278	76.014	139.821	203.080	114.200	29.072	48.738	106.685	6.833	-	135.421
std	11.908	10.576	19.054	48.197	16.261	3.891	12.230	33.820	1.230	-	69.849
min	1	50	71	59	80	13.13	10	10.6	4	-	11
25%	62	70	128.14	170	102	27.13	41	84.8	6	-	93
50%	70	76	139	198	115	29.04	47	105	6.6	-	126
75%	78	81.54	150	230	121	30.47	55	122.8	7.2	-	155
max	101	132	240	600	283	49.22	100	287.2	14.7	-	1000

Tabela Anexo I. 9 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2018.*

2018	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8493	8491	8491	8471	8471	8491	8491	8491	8491	0	8491
mean	69.721	75.505	138.315	192.134	115.989	29.050	50.206	99.909	6.888	-	136.797
std	11.806	10.383	18.859	65.492	25.965	3.742	12.557	32.067	1.189	-	68.999
min	8	50	63	58	50	15.06	10	10.2	1.8	-	10
25%	62	70	127	161	105	27.41	42	79.4	6.1	-	94
50%	70	75	138	183	110	29.04	49	97	6.7	-	128
75%	79	81	148	204	121	30.45	56	115.2	7.2	-	157
max	102	138	236	595	299	49.29	100	294.4	15	-	1000

Tabela Anexo I. 10 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2019.*

2019	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8555	8552	8552	8552	8552	8552	8552	8549	8552	6861	8552
mean	69.963	75.427	137.745	252.374	146.010	29.108	49.748	98.981	6.942	0.993	137.171
std	11.848	10.404	18.273	100.320	46.121	3.213	11.602	25.587	1.214	0.460	66.705
min	9	50	67	62	65	15.82	10	11.4	1	0.5	10
25%	62	69	127	188	107	27.74	43	89.4	6.2	0.706	95
50%	71	76	137.69	249.95	143	28.84	48	98	6.7	0.89	129
75%	79	81	147	275	161	30.35	54.095	106.259	7.3	1.1	155.6
max	103	145	231	599	300	49.25	100	294.4	15	11.6	827

Tabela Anexo I. 11 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2020.*

2020	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8560	8554	8556	8554	8554	8556	8556	8556	8556	8556	8556
mean	70.225	76.524	137.750	216.864	125.315	29.047	50.379	95.508	6.953	0.963	137.726
std	11.831	7.587	18.176	77.179	33.548	3.096	12.059	30.799	1.170	0.411	66.703
min	2	50	65	60	53	16.22	11	11.8	1	0.5	14
25%	63	72.5	128	183	106	27.58	43	75.763	6.2	0.75	95
50%	71	75.9	137.4	216.5	118	28.74	49	92.965	6.8	0.89	129
75%	79	80.7	147	227.2	130.23	30.56	56	109.8	7.3	1.043	158
max	104	125	232	600	297	50	100	294.4	15	10.6	965

Tabela Anexo I. 12 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2021.*

2021	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	8840	8835	8835	8810	8788	8822	8835	8835	8835	8835	8835
mean	70.184	75.864	136.779	216.353	127.636	29.117	50.374	95.315	6.984	0.979	139.602
std	11.918	9.881	17.553	67.978	30.769	2.963	11.998	31.427	1.159	0.436	69.800
min	3	50	62	64	30	11.72	16	10	1	0.5	18
25%	63	70	126	195	110.11	27.73	43	75.2	6.3	0.76	96
50%	71	76	136.9	210.862	124.49	28.93	49	92.8	6.8	0.9	131
75%	79	81	145	232.319	135.34	30.66	56	110.4	7.3	1.06	159
max	102	130	238	596	300	50	100	289.6	14.6	10.2	1000

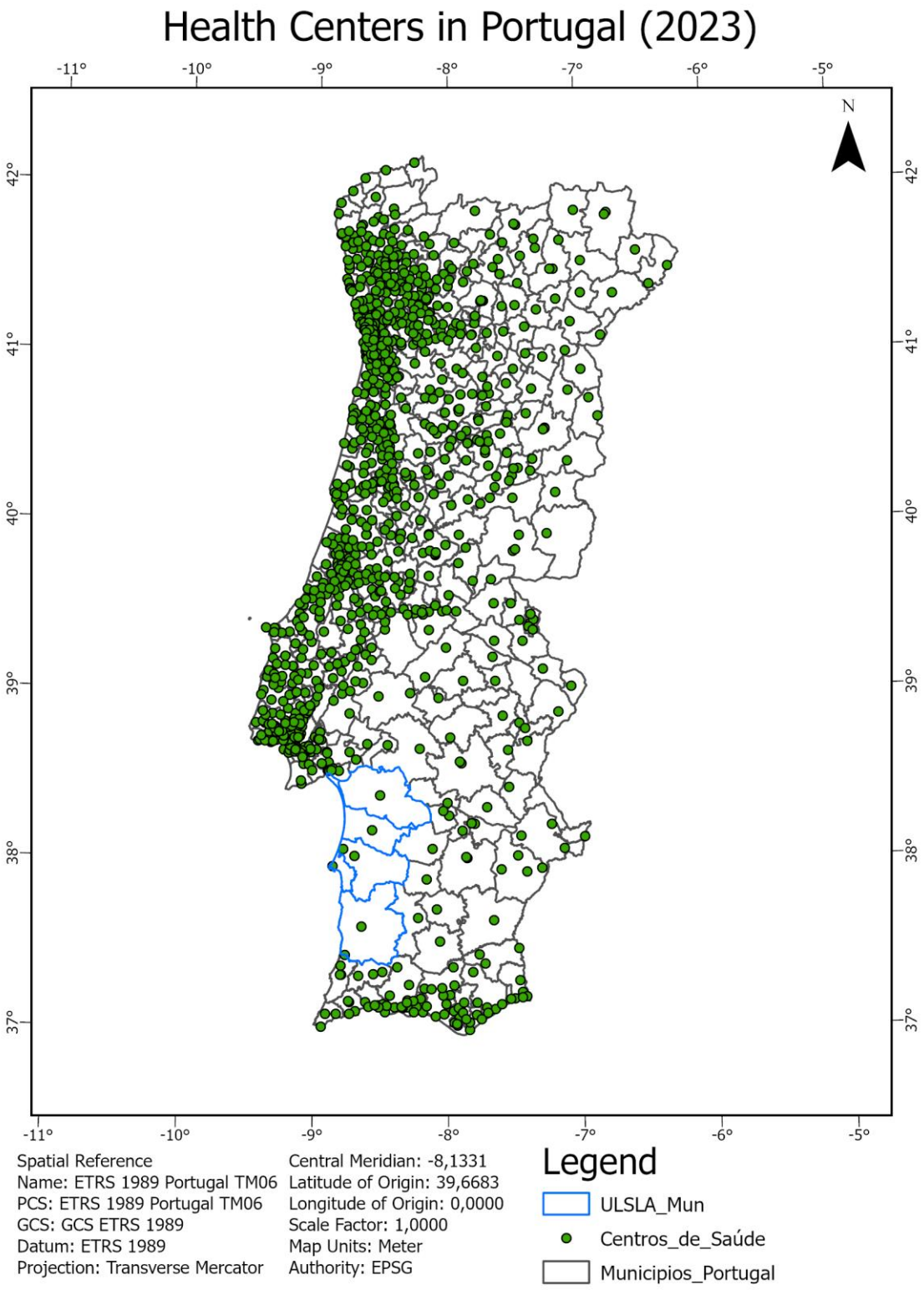
Tabela Anexo I. 13 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2022.*

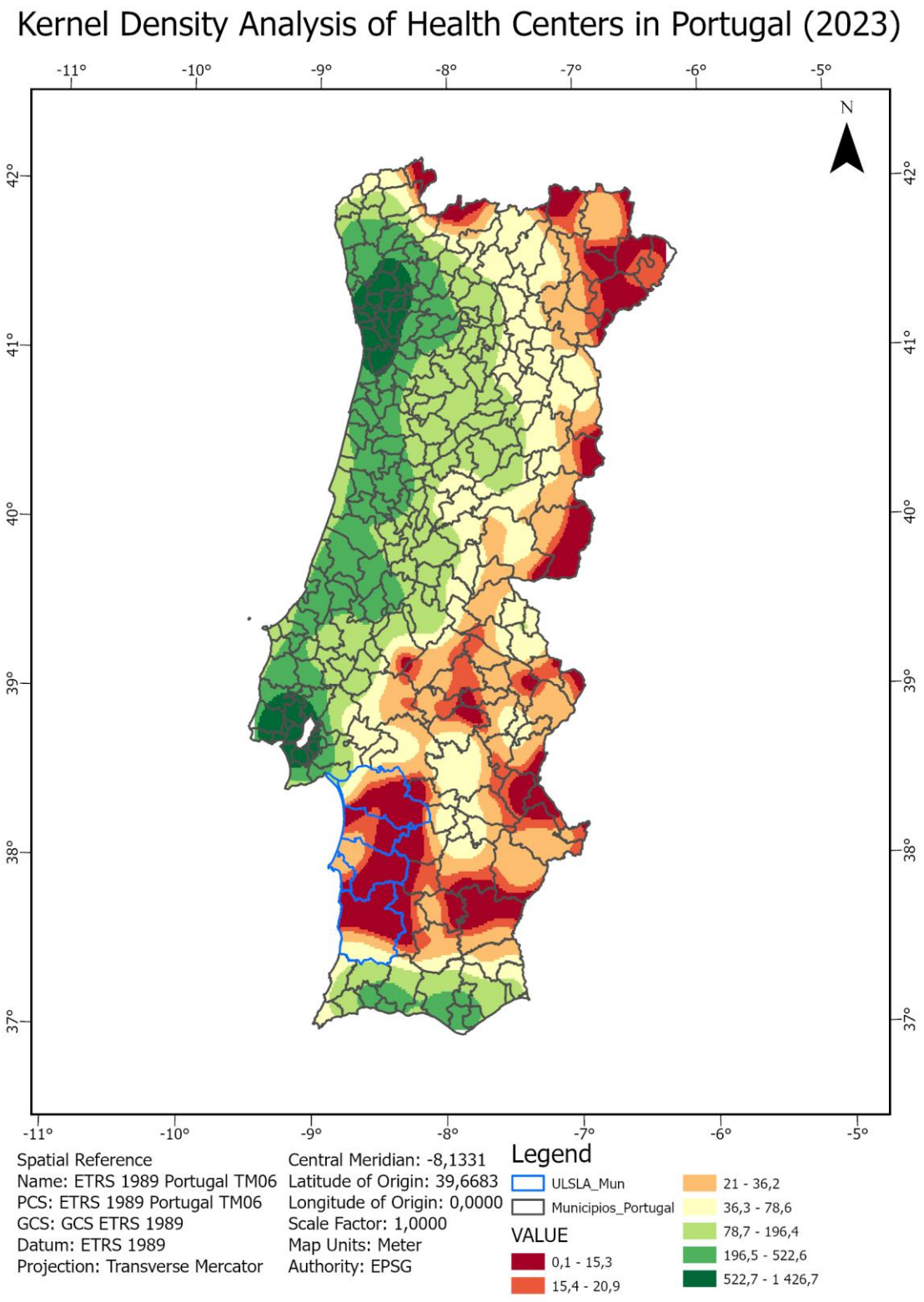
2022	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	9067	9061	9061	9061	9061	9049	9061	9061	9061	9061	9061
mean	70.270	75.566	136.844	225.887	137.364	29.181	49.521	93.779	6.952	0.982	138.918
std	12.038	9.924	17.669	82.653	39.813	2.864	12.110	32.393	1.169	0.445	70.017
min	7	50	60	63	64	15.73	10	10	3.8	0.5	10
25%	63	70	126	180	109	27.78	42	73.4	6.2	0.76	96
50%	71	76	137	224.774	133.34	29.06	48	92.2	6.762	0.9	130
75%	79	81	146	242.285	145.08	30.51	55	108.4	7.3	1.06	157
max	103	121	235	600	300	49.45	100	289.6	14.9	10.6	1000

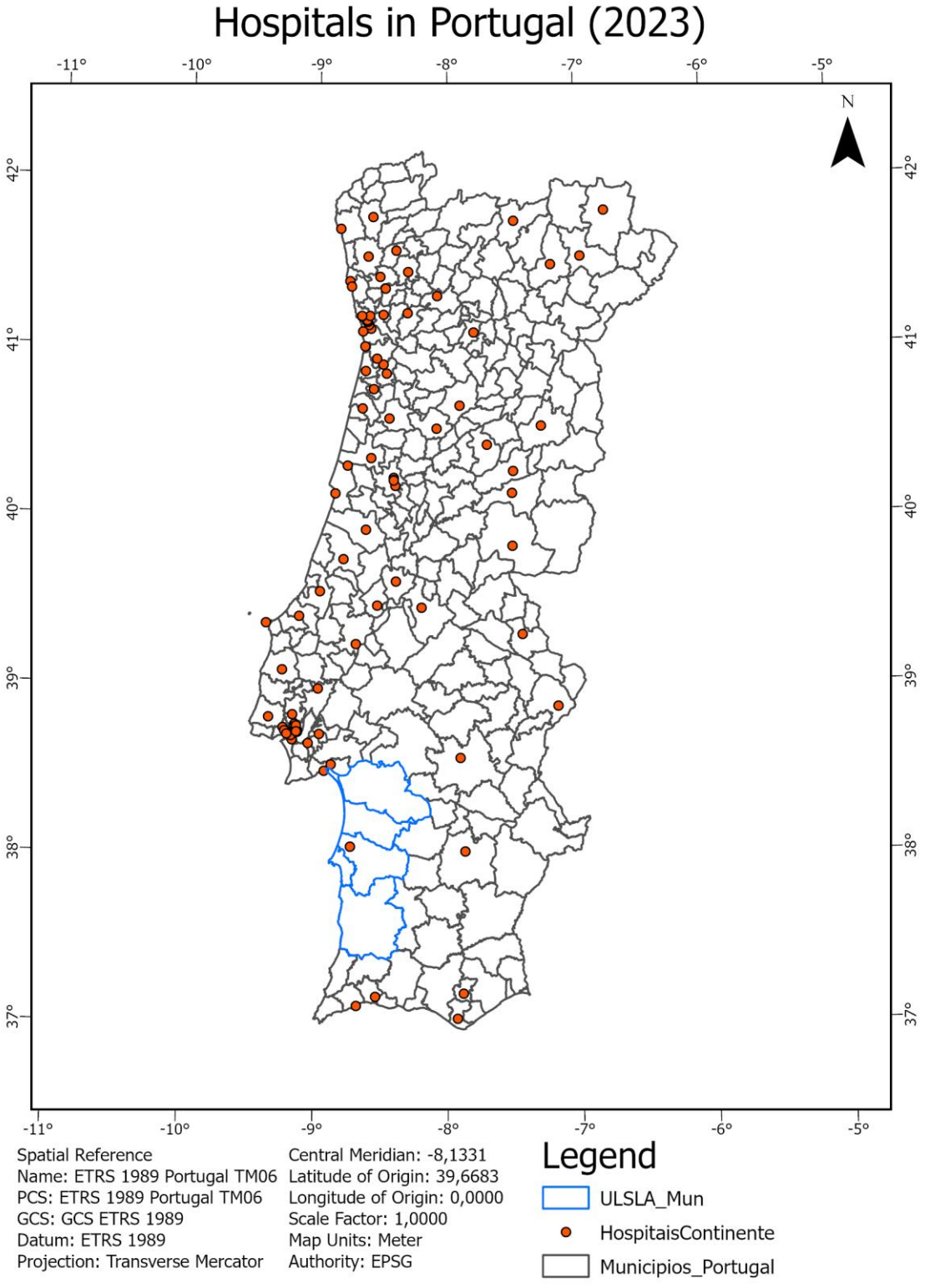
Tabela Anexo I. 14 *Estatísticas Descritivas das Variáveis Clínicas após o processamento de dados em 2023.*

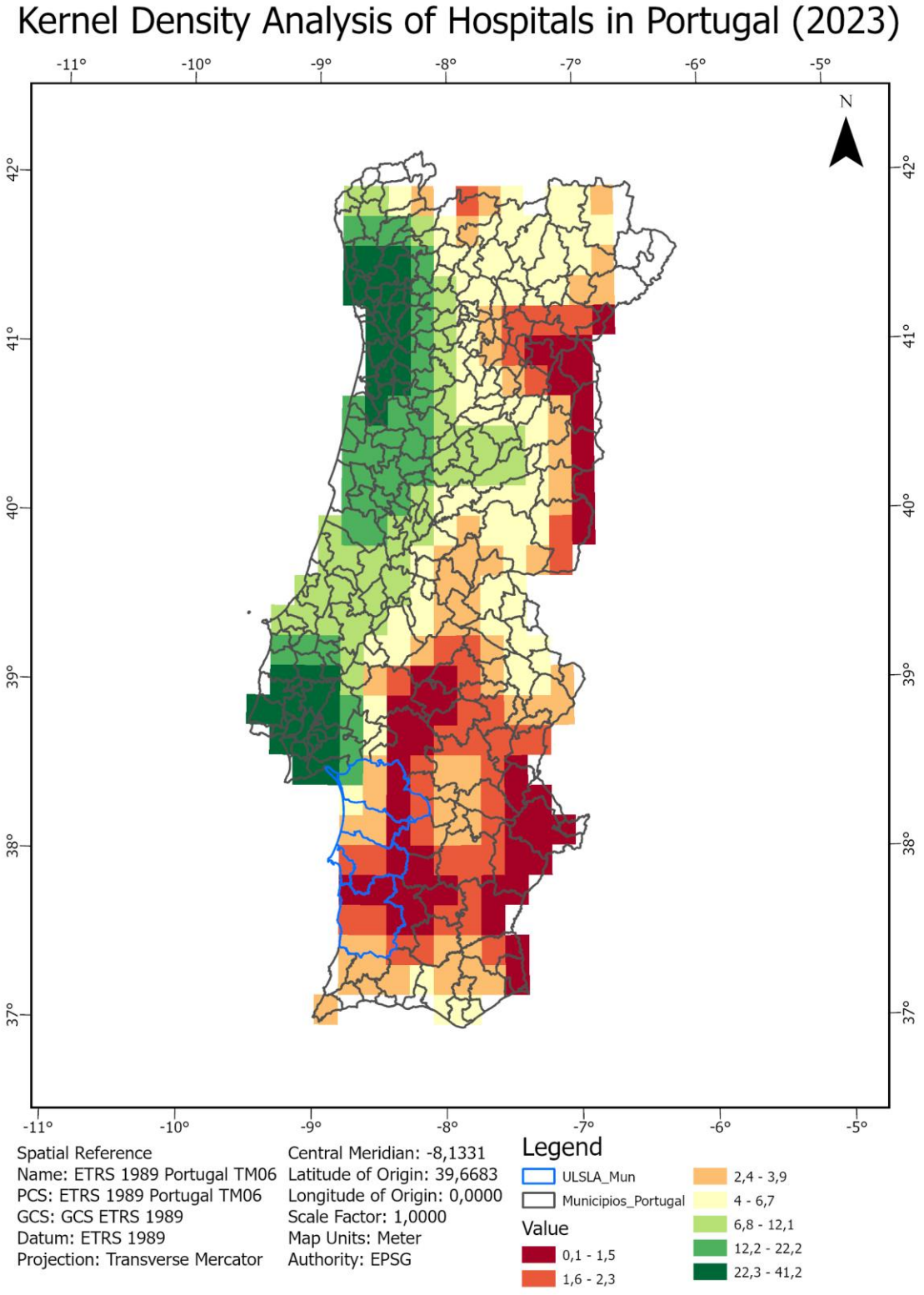
2023	Idade	Tad	Tas	Glicemia Capilar	Perimetro AB	IMC	HDL	LDL	Hba1c	Creatinina	Trigliceridos
count	9344	9341	9341	9339	9339	9318	9341	9313	9341	8933	9341
mean	70.427	75.941	136.364	227.490	139.045	28.879	51.089	93.645	6.946	0.989	134.823
std	12.108	10.149	17.708	86.841	41.419	2.940	12.426	26.654	1.145	0.474	68.782
min	17	50	60	56	31	15.63	13	10	3.4	0.5	10
25%	63	70	126	177	107	27.38	43	84.8	6.2	0.72	93
50%	71	76	136	222.78	135.61	28.69	50	94.561	6.7934	0.9	125
75%	79	82	145	245.94	144.93	29.86	57	101.704	7.3	1.13	154
max	104	135	229	599	300	49.35	100	298	14.9	10	985

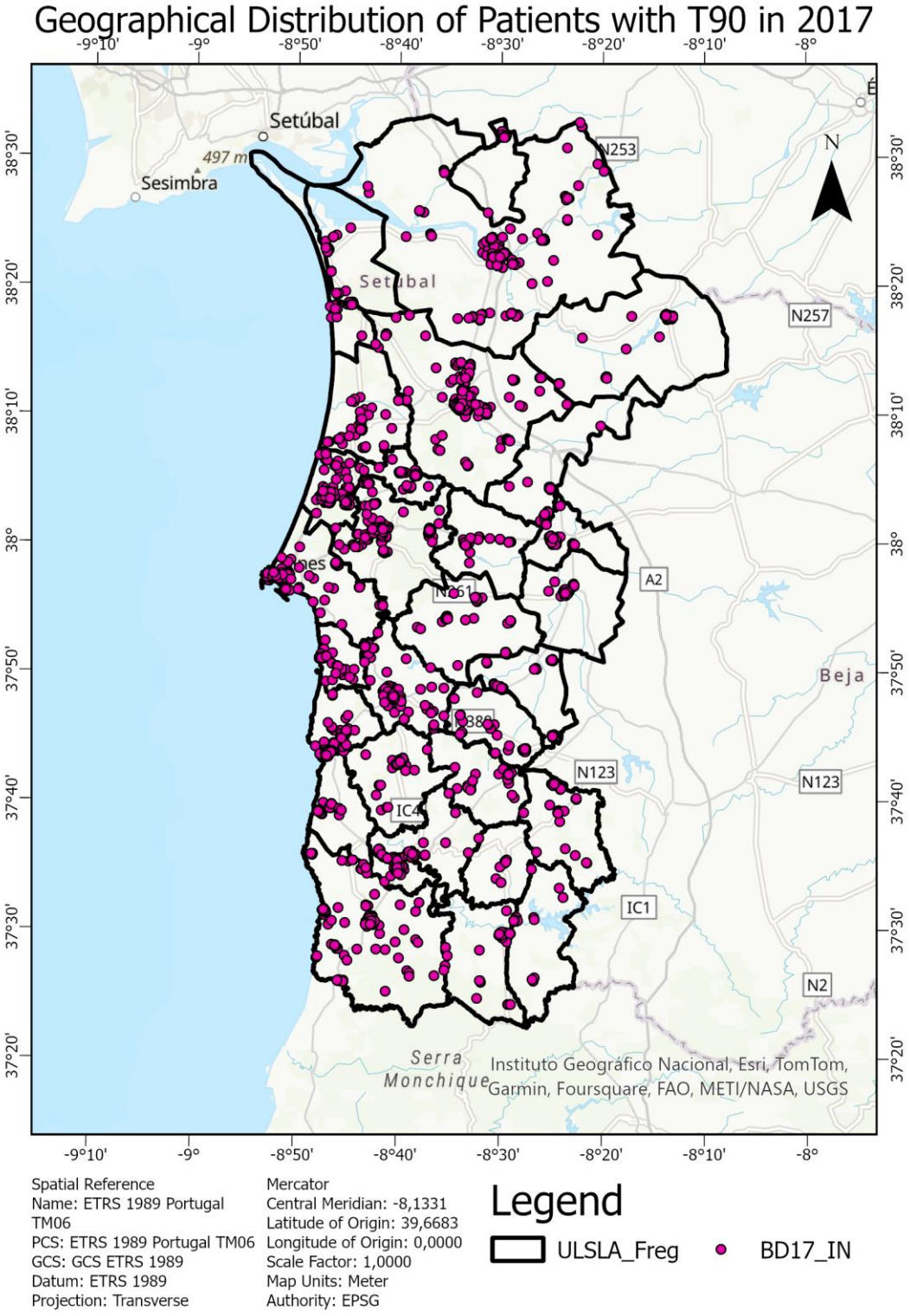
Anexo II- Visualização de Dados

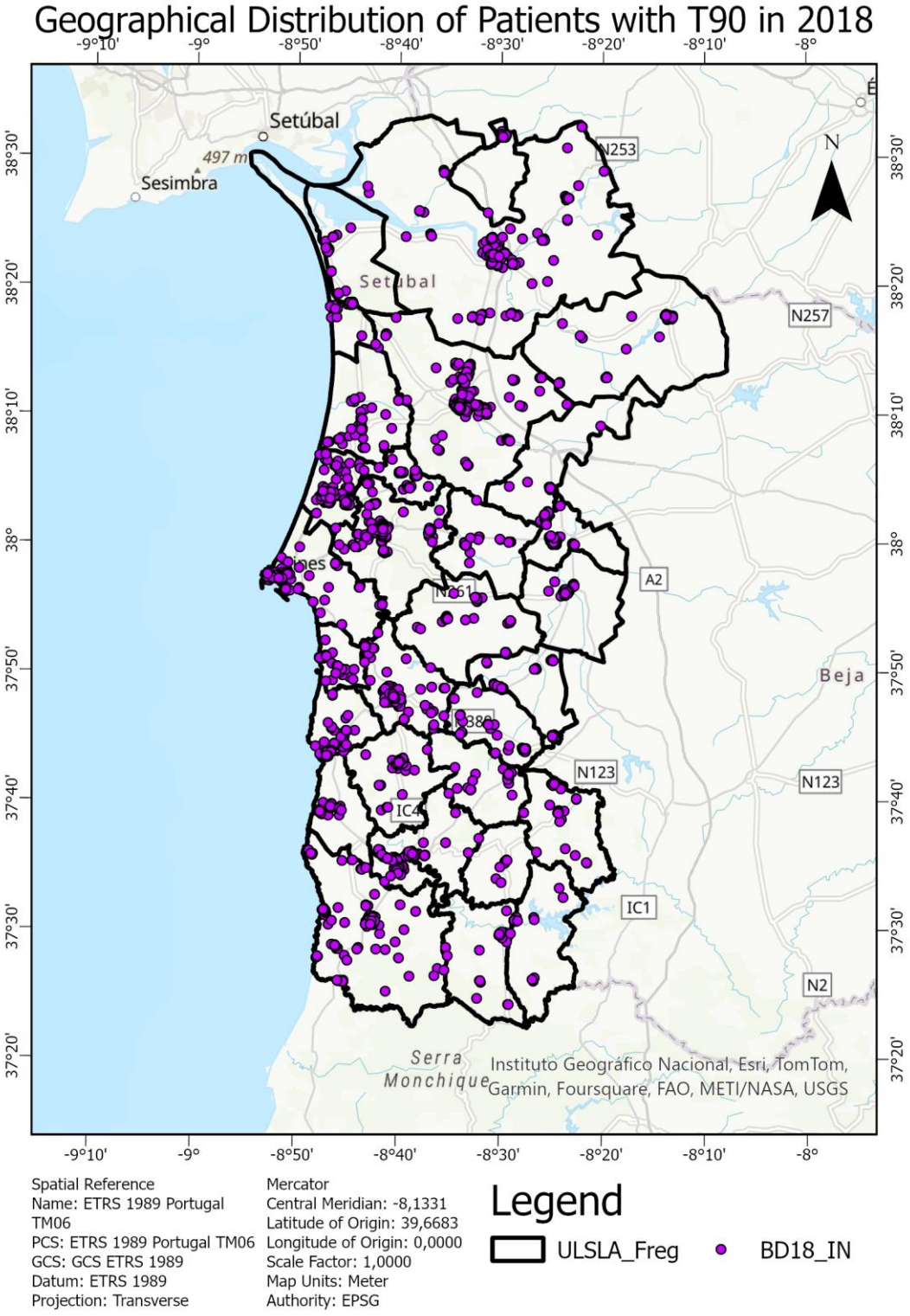


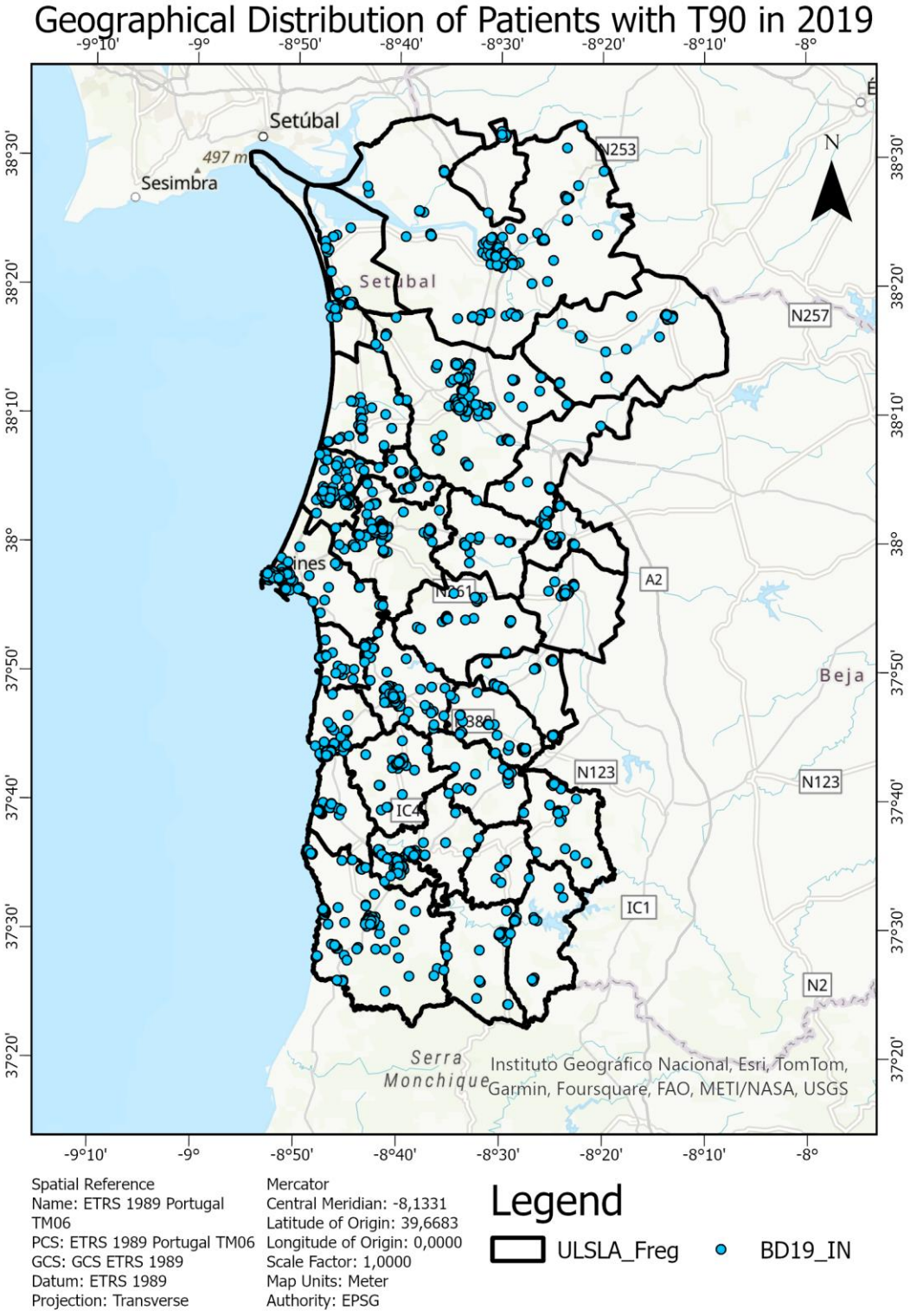

Kernel Density Analysis of Health Centers in Portugal (2023)
-11°
-10°
-9°
-8°
-7°
-6°
-5°
42°
41°
40°
39°
38°
37°
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ULSLA_Mun
Municipios_Portugal
VALUE
0,1 - 15,3
15,4 - 20,9
21 - 36,2
36,3 - 78,6
78,7 - 196,4
196,5 - 522,6
522,7 - 1 426,7

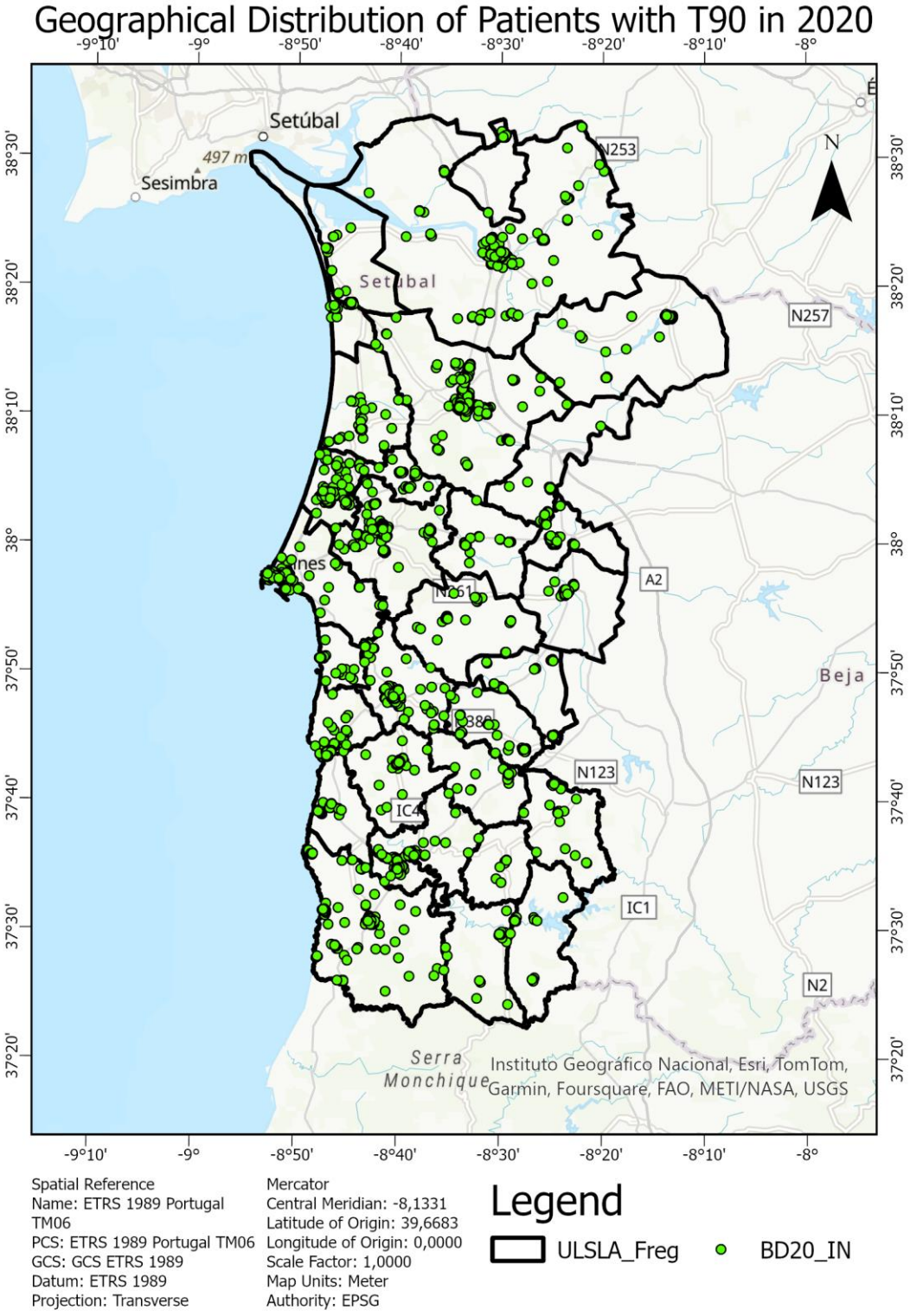

Hospitals in Portugal (2023)
-11°
-10°
-9°
-8°
-7°
-6°
-5°
42°
41°
40°
39°
38°
37°
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ULSLA_Mun
HospitaisContinente
Municipios_Portugal

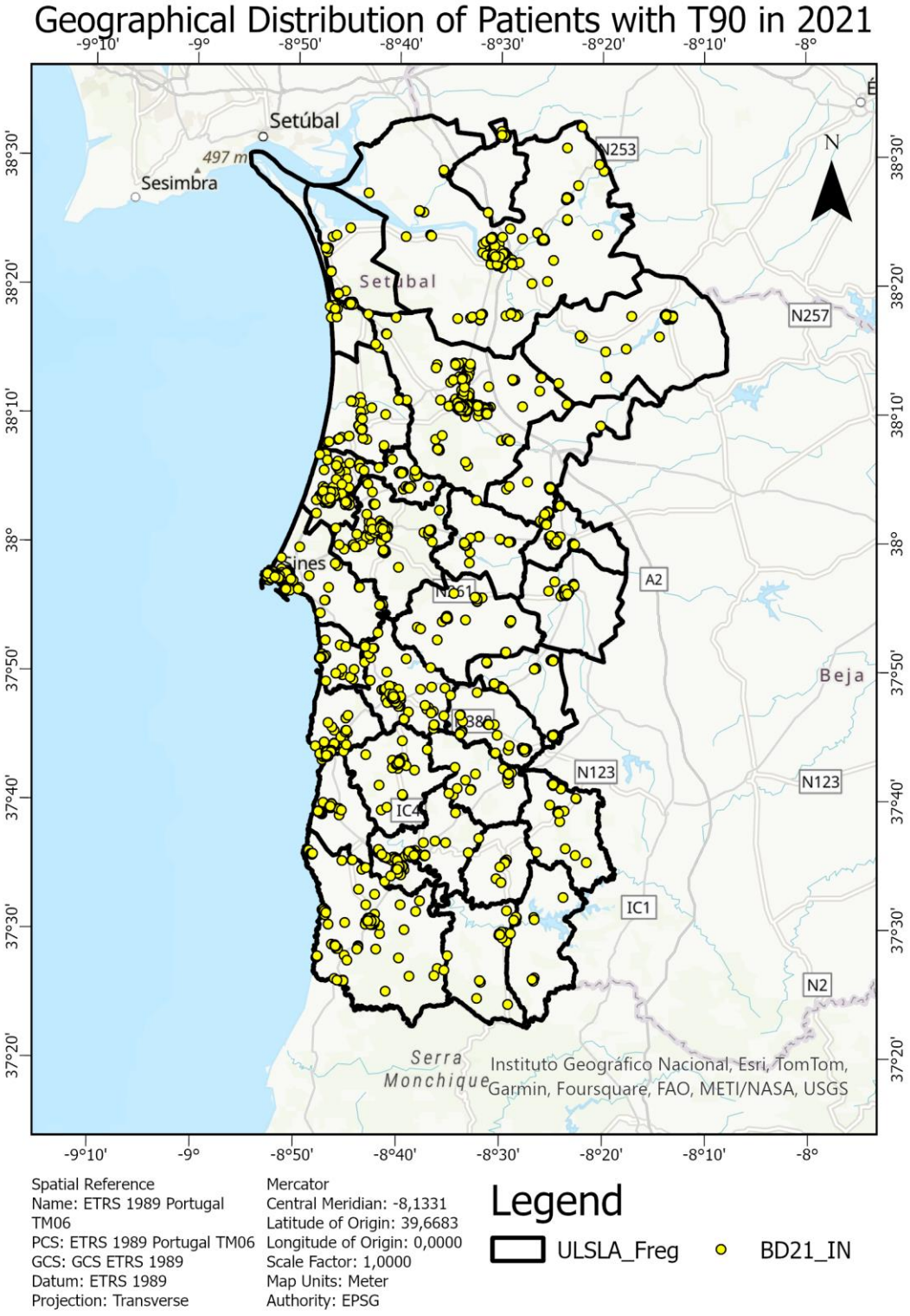

Kernel Density Analysis of Hospitals in Portugal (2023)
-11°
-10°
-9°
-8°
-7°
-6°
-5°
42°
41°
40°
39°
38°
37°
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ULSLA_Mun
Municipios_Portugal
Value
0,1 - 1,5
1,6 - 2,3
2,4 - 3,9
4 - 6,7
6,8 - 12,1
12,2 - 22,2
22,3 - 41,2

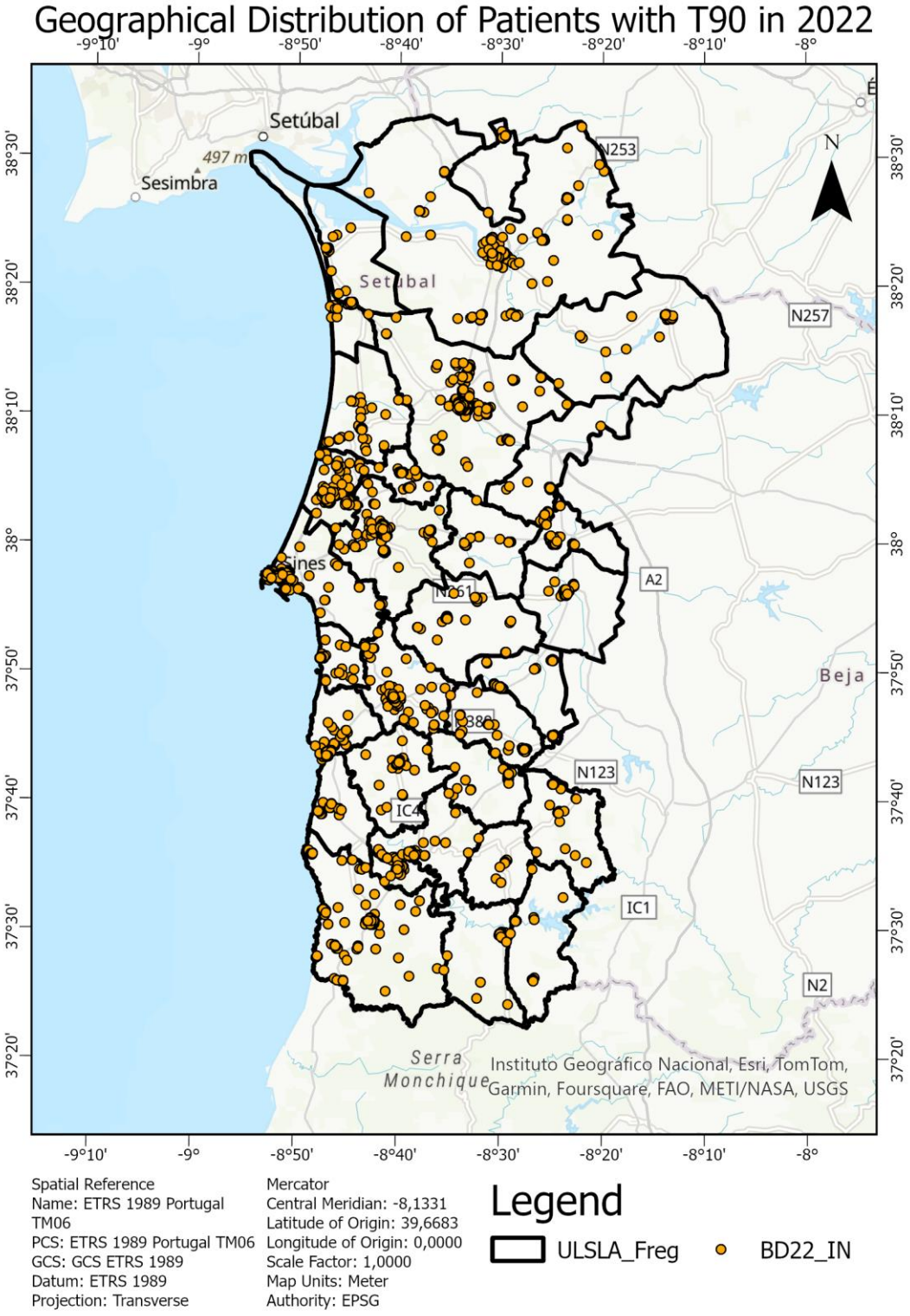

Geographical Distribution of Patients with T90 in 2017
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
Setúbal
497 m
Sesimbra
Setúbal
N253
N257
A2
Beja
N123
IC1
N2
Serra
Monchique
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ULSLA_Freg
BD17_IN

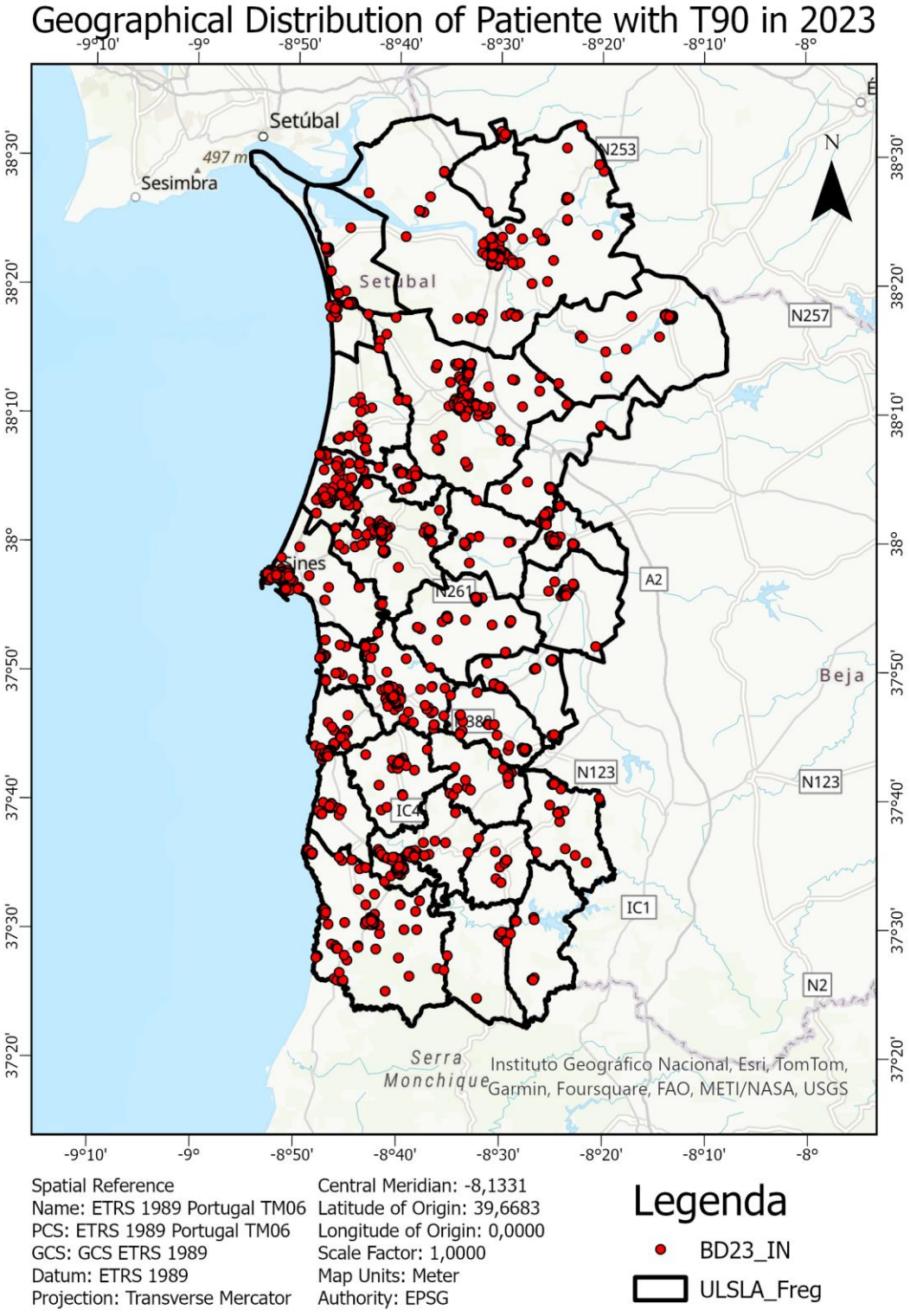

Geographical Distribution of Patients with T90 in 2018
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
Setúbal
497 m
Sesimbra
Setúbal
N253
N257
A2
N261
N123
IC4
IC1
N2
Beja
Serra
Monchique
N
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ULSLA_Freg
BD18_IN


Geographical Distribution of Patients with T90 in 2019
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
Setúbal
497 m
Sesimbra
N253
N257
A2
Beja
N123
IC1
N2
Serra
Monchique
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ULSLA_Freg
BD19_IN

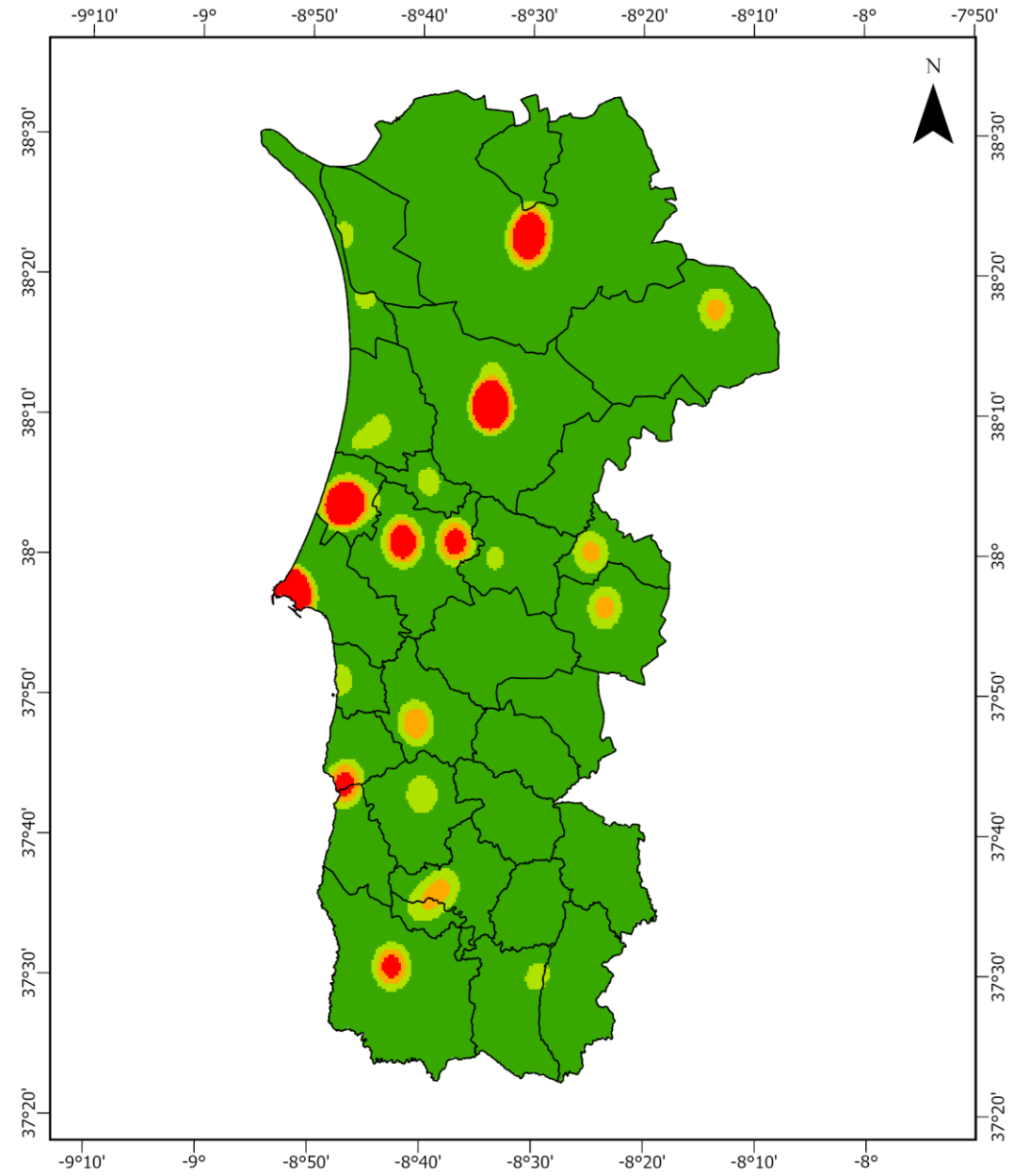

Geographical Distribution of Patients with T90 in 2020
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
Setúbal
497 m
Sesimbra
Setúbal
N253
N257
A2
N261
N123
IC4
IC1
N2
Beja
Serra
Monchique
N
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ULSLA_Freg
BD20_IN


Geographical Distribution of Patients with T90 in 2021
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
Setúbal
497 m
Sesimbra
Setúbal
Sines
N253
N257
A2
N261
N123
IC4
IC1
N2
Beja
Serra
Monchique
N
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ULSLA_Freg
BD21_IN


Geographical Distribution of Patients with T90 in 2022
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
Setúbal
497 m
Sesimbra
Setúbal
Sines
N253
N257
A2
N261
N123
IC4
IC1
N2
Beja
Serra
Monchique
N
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ULSLA_Freg
BD22_IN


Geographical Distribution of Patiente with T90 in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
Setúbal
497 m
Sesimbra
Setúbal
Sines
N253
N257
A2
N261
N123
IC4
IC1
N2
Beja
Serra
Monchique
N
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legenda
BD23_IN
ULSLA_Freg

Kernel Density Estimation of Patients with T90 in 2017



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

Legend

ULSLA_Freg

VALUE

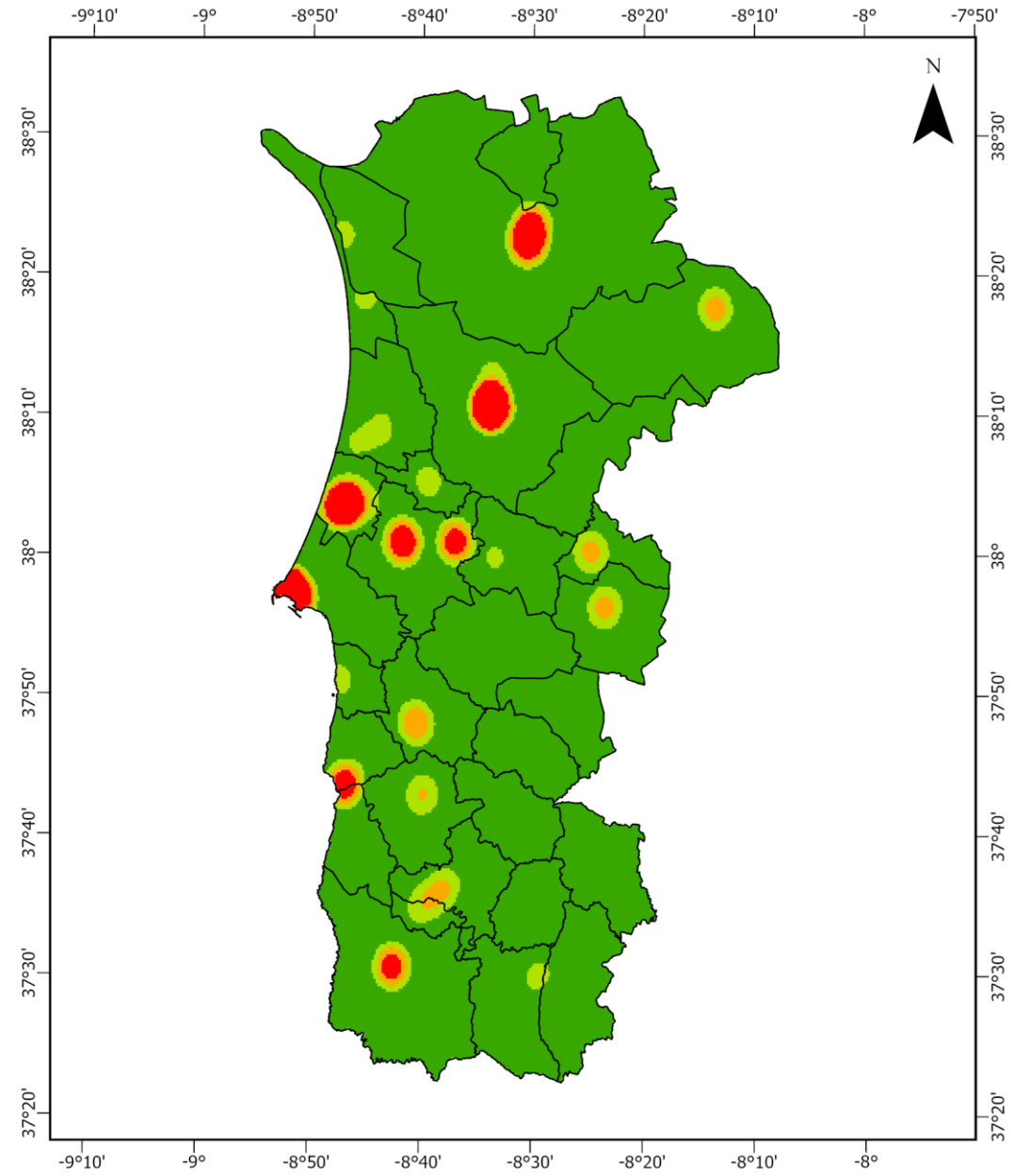
0 - 38 715

38 716 - 85 053

85 054 - 131 390

131 391 - 590 048

Kernel Density Estimation of Patients with T90 in 2018



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

Legend

ULSLA_Freg

VALUE

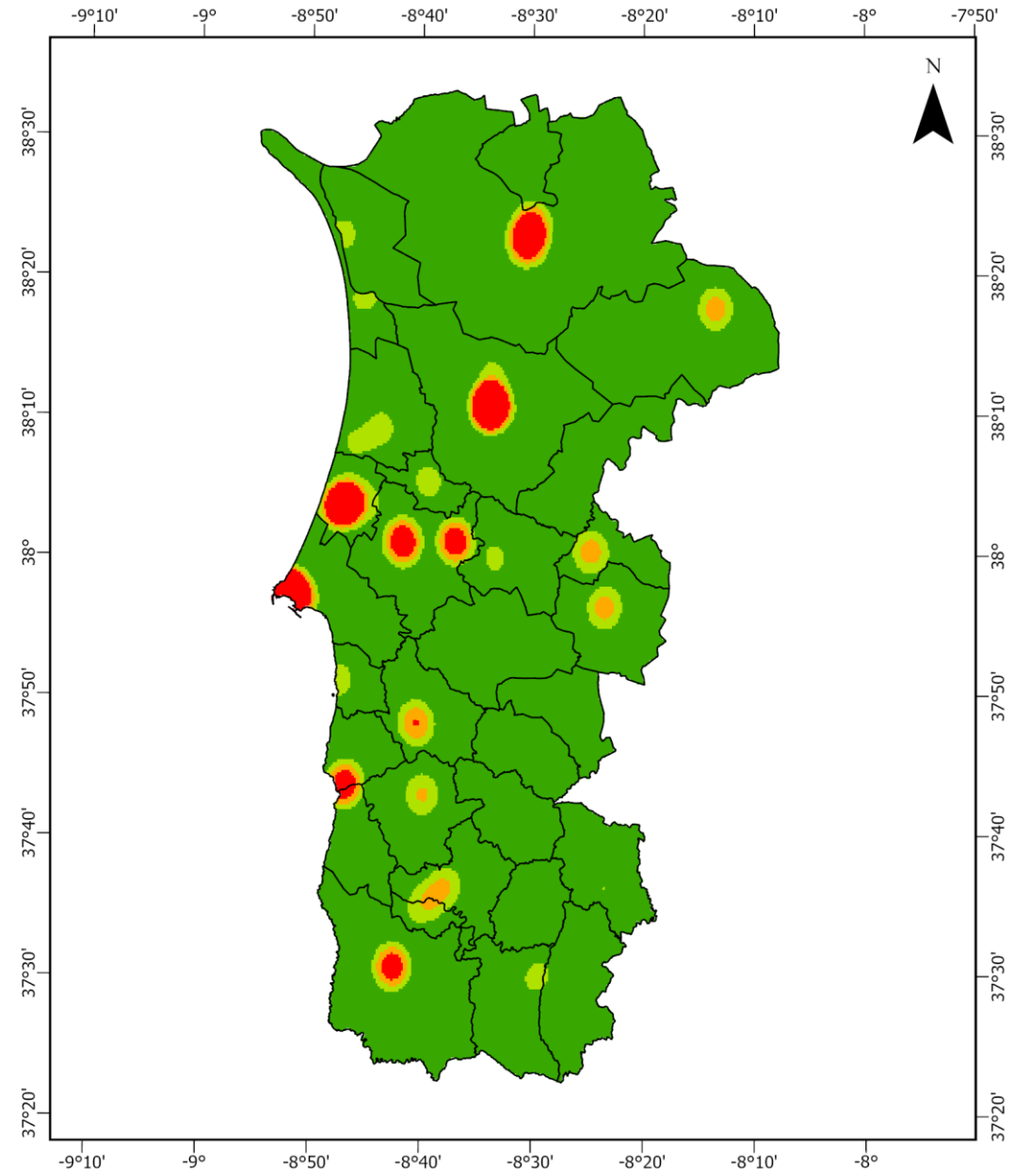
0 - 38 715

38 716 - 85 053

85 054 - 131 390

131 391 - 590 048

Kernel Density Estimation of Patients with T90 in 2019



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

Legend

ULSLA_Freg

VALUE

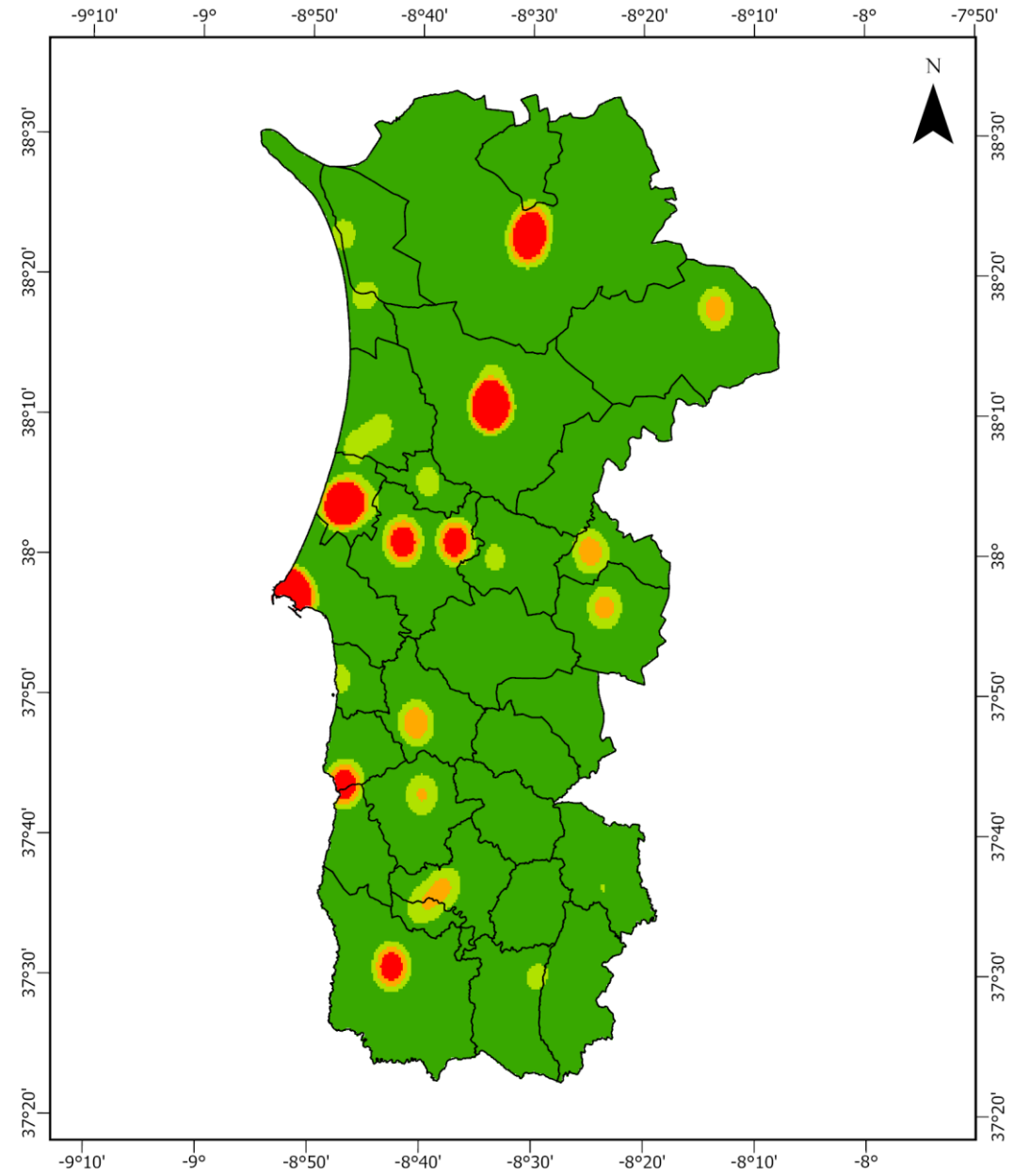
0 - 38 715

38 716 - 85 053

85 054 - 131 390

131 391 - 590 048

Kernel Density Estimation of Patients with T90 in 2020



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

Legend

ULSLA_Freg

VALUE

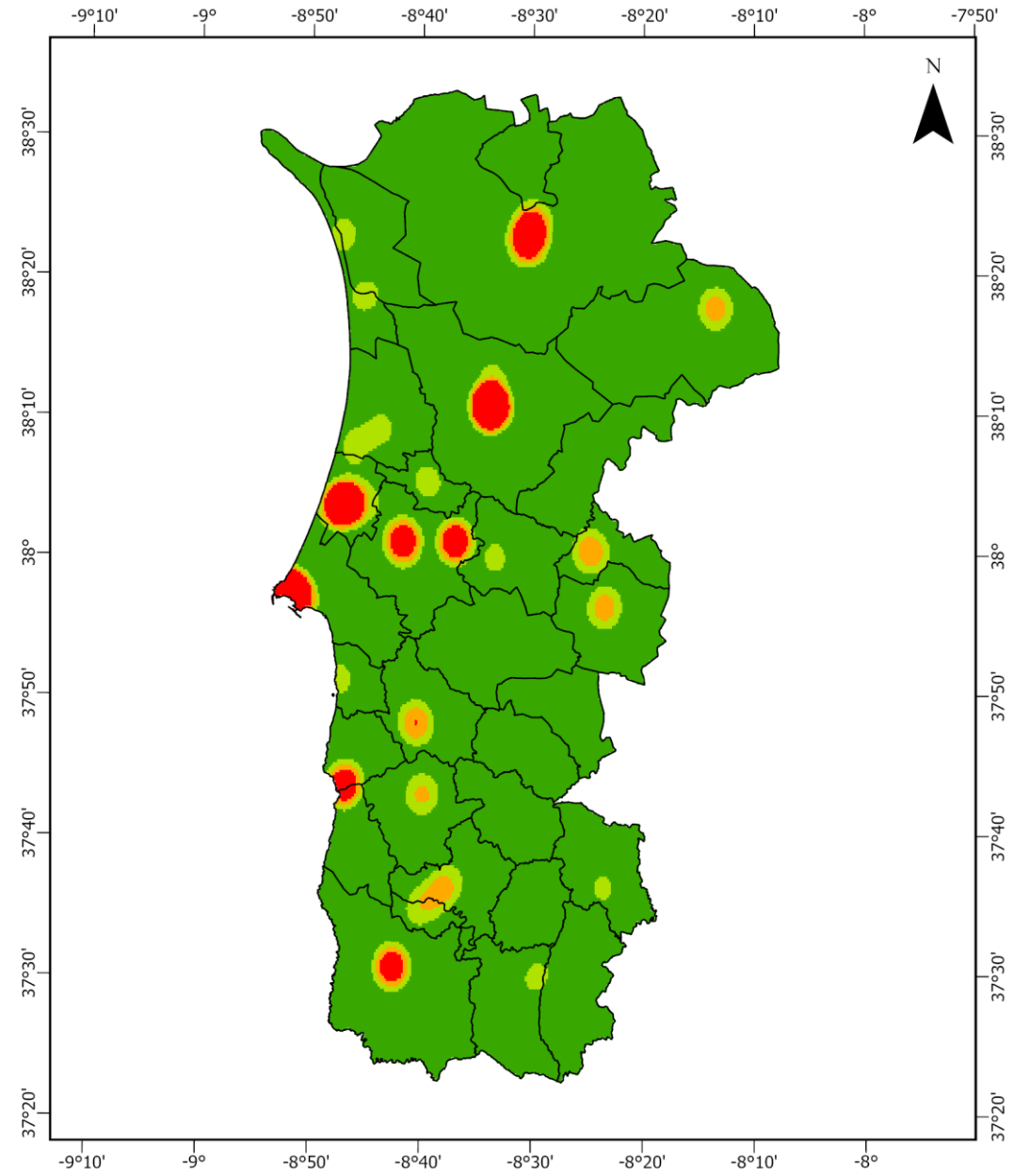
0 - 38 715

38 716 - 85 053

85 054 - 131 390

131 391 - 590 048

Kernel Density Estimation of Patients with T90 in 2021



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

Legend

ULSLA_Freg

VALUE

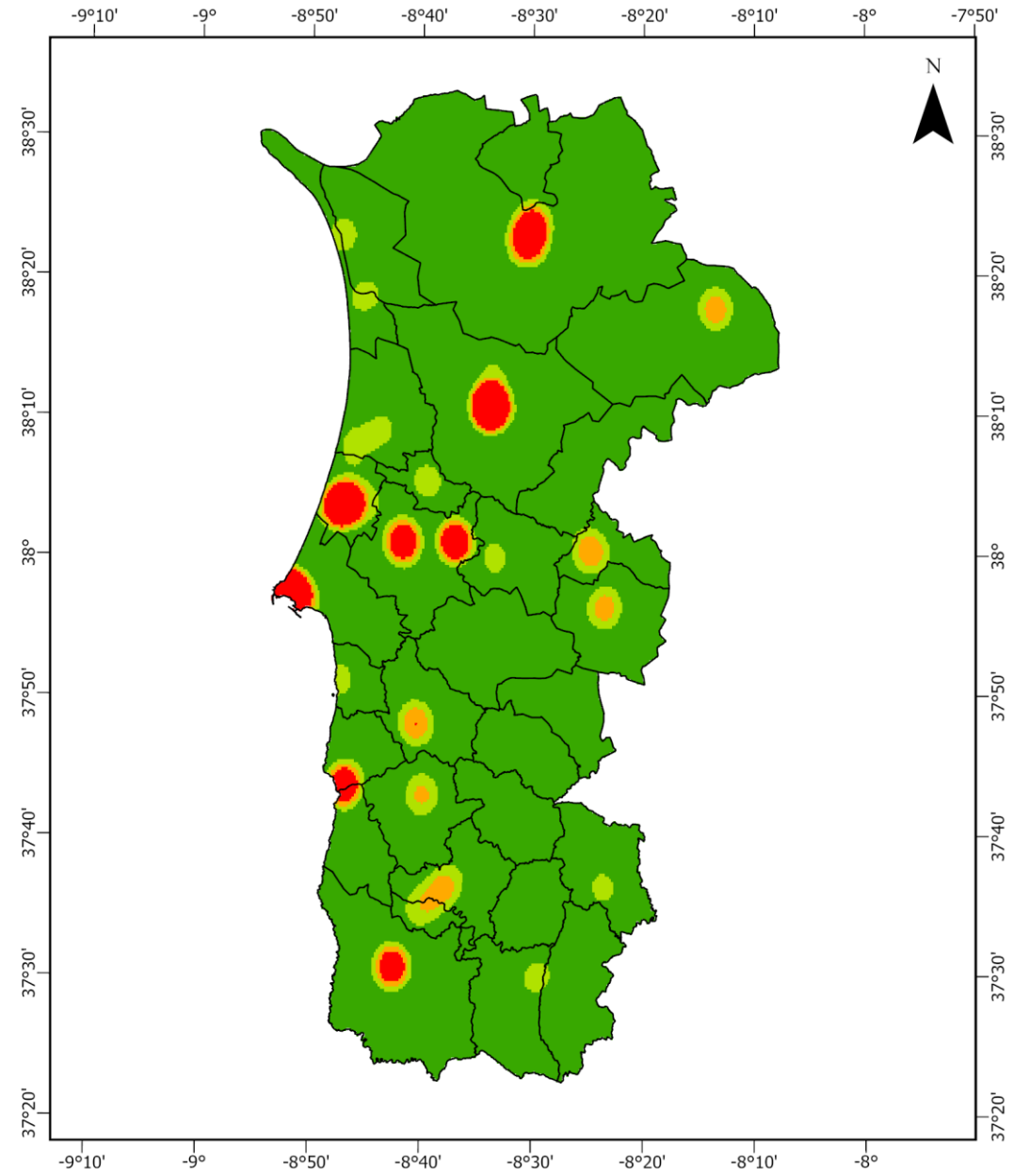
0 - 38 715

38 716 - 85 053

85 054 - 131 390

131 391 - 590 048

Kernel Density Estimation of Patients with T90 in 2022



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

Legend

ULSLA_Freg

VALUE

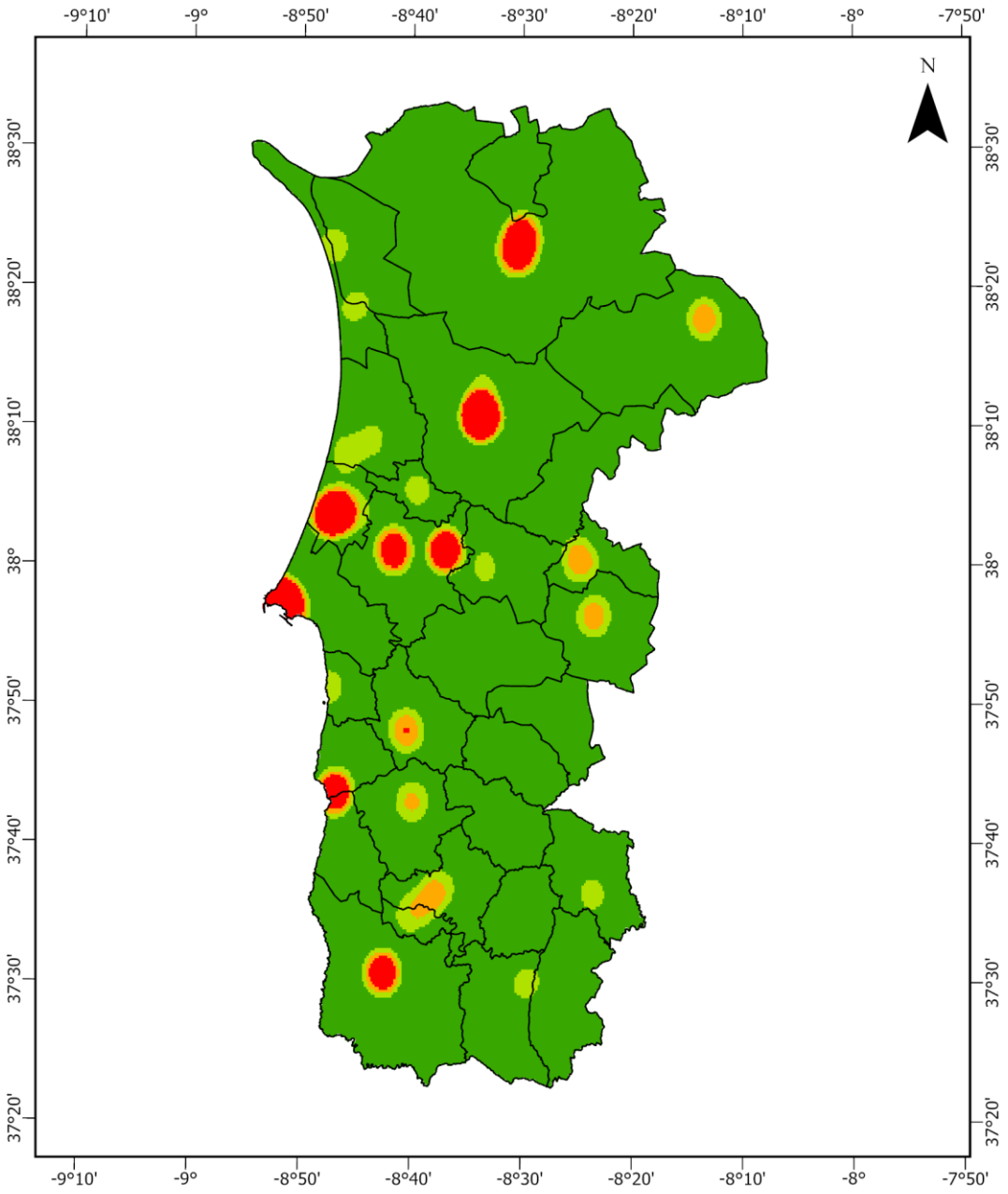
0 - 38 175

38 176 - 85 053

85 056 - 131 390

131 391 - 590 048

Kernel Density Estimation of Patients with T90 in 2023



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

Legend

ULSLA_Freg

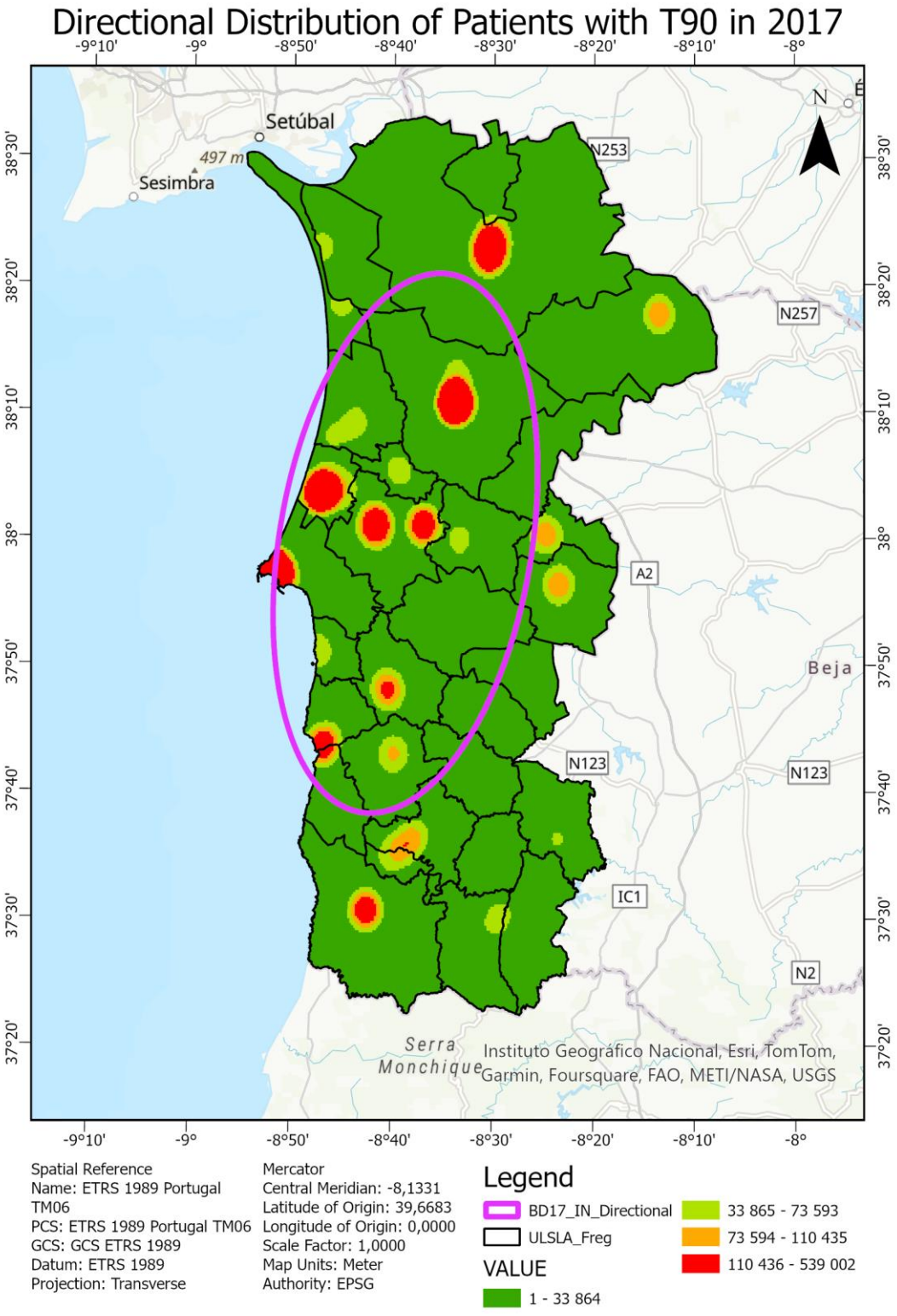
VALUE

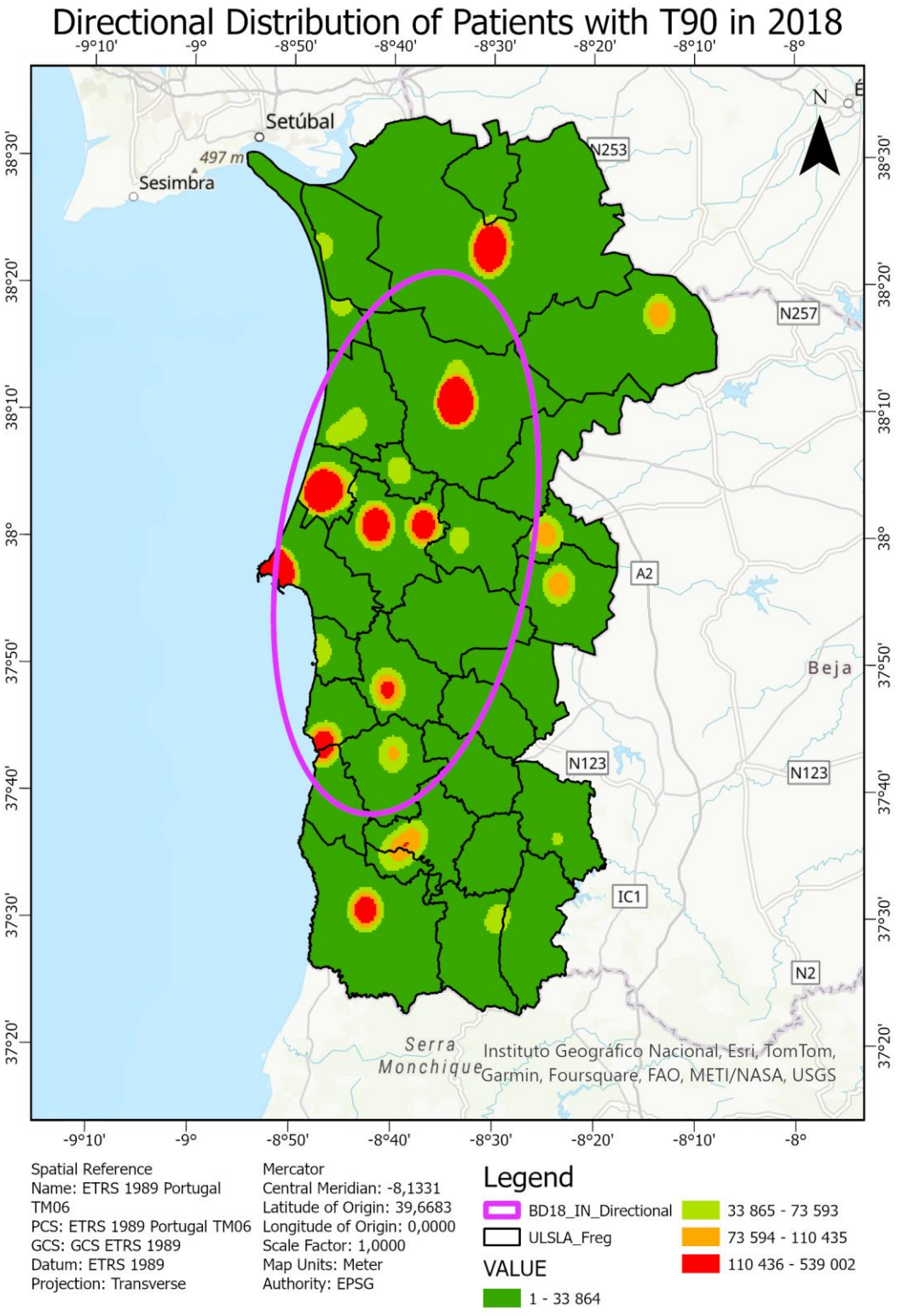
1 - 38 715

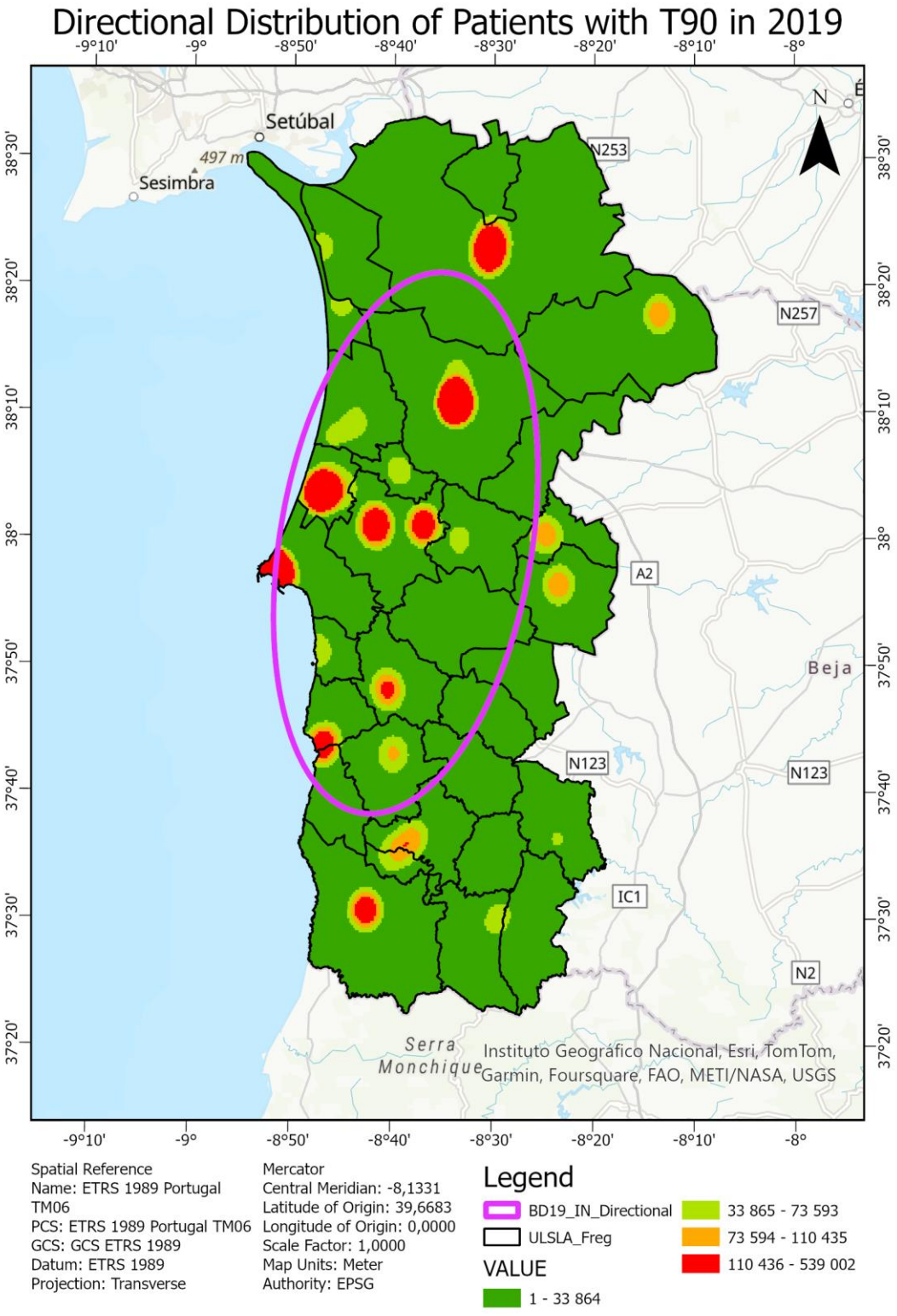
38 716 - 85 053

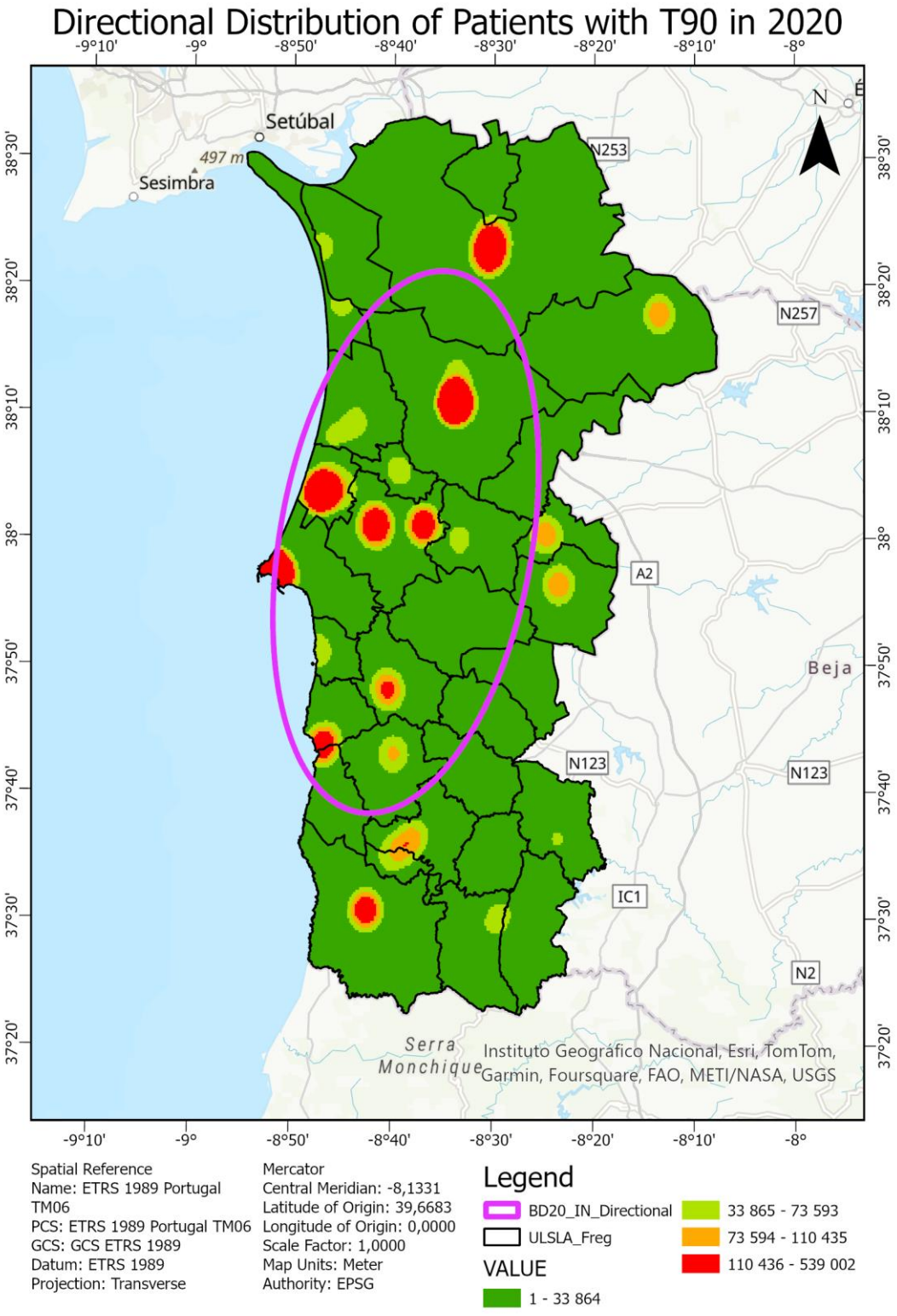
85 054 - 131 390

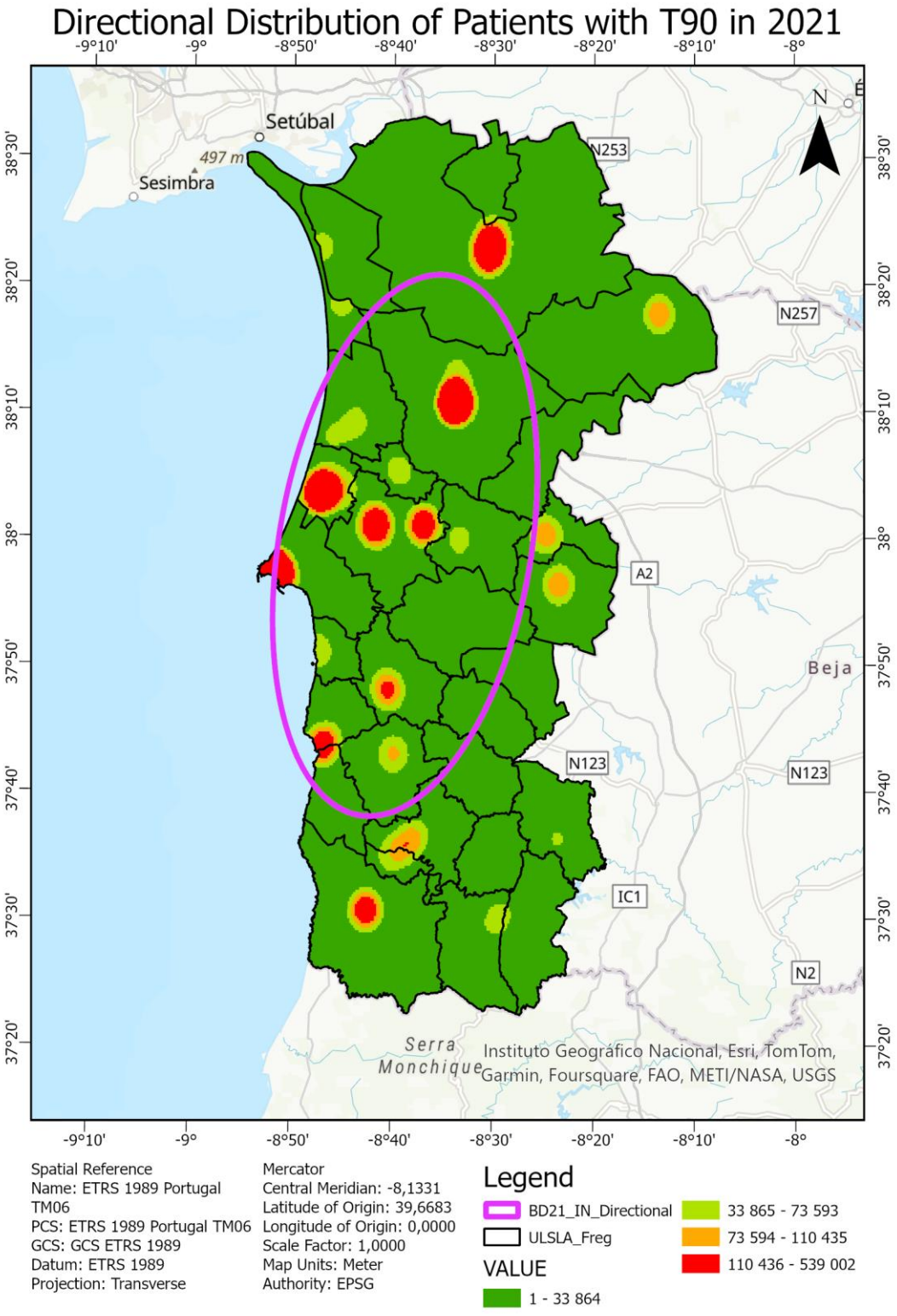
131 391 - 590 048


Directional Distribution of Patients with T90 in 2017
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Setúbal
497 m
Sesimbra
N253
N257
A2
Beja
N123
IC1
N2
Serra Monchique
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
BD17_IN_Directional
ULSLA_Freg
VALUE
1 - 33 864
33 865 - 73 593
73 594 - 110 435
110 436 - 539 002

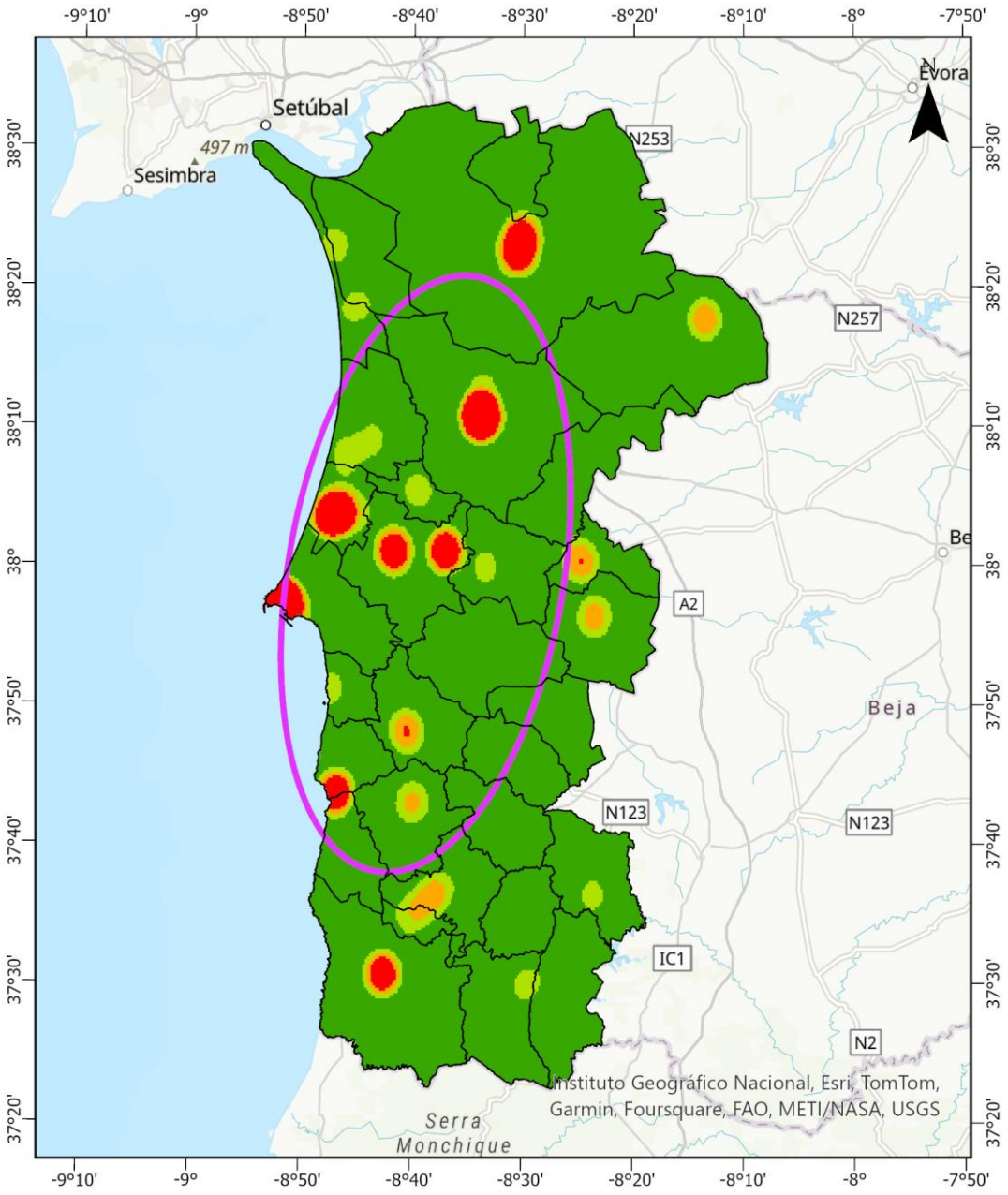

Directional Distribution of Patients with T90 in 2018
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Setúbal
497 m
Sesimbra
N253
N257
A2
Beja
N123
IC1
N2
Serra Monchique
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
BD18_IN_Directional
ULSLA_Freg
VALUE
1 - 33 864
33 865 - 73 593
73 594 - 110 435
110 436 - 539 002


Directional Distribution of Patients with T90 in 2019
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Setúbal
497 m
Sesimbra
N253
N257
A2
Beja
N123
IC1
N2
Serra Monchique
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
BD19_IN_Directional
ULSLA_Freg
VALUE
1 - 33 864
33 865 - 73 593
73 594 - 110 435
110 436 - 539 002


Directional Distribution of Patients with T90 in 2020
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Setúbal
497 m
Sesimbra
N253
N257
A2
Beja
N123
IC1
N2
Serra Monchique
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
BD20_IN_Directional
ULSLA_Freg
VALUE
1 - 33 864
33 865 - 73 593
73 594 - 110 435
110 436 - 539 002


Directional Distribution of Patients with T90 in 2021
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Setúbal
497 m
Sesimbra
N253
N257
A2
Beja
N123
IC1
N2
Serra Monchique
Instituto Geográfico Nacional, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
BD21_IN_Directional
ULSLA_Freg
VALUE
1 - 33 864
33 865 - 73 593
73 594 - 110 435
110 436 - 539 002

Directional Distribution of Patients with T90 in 2022

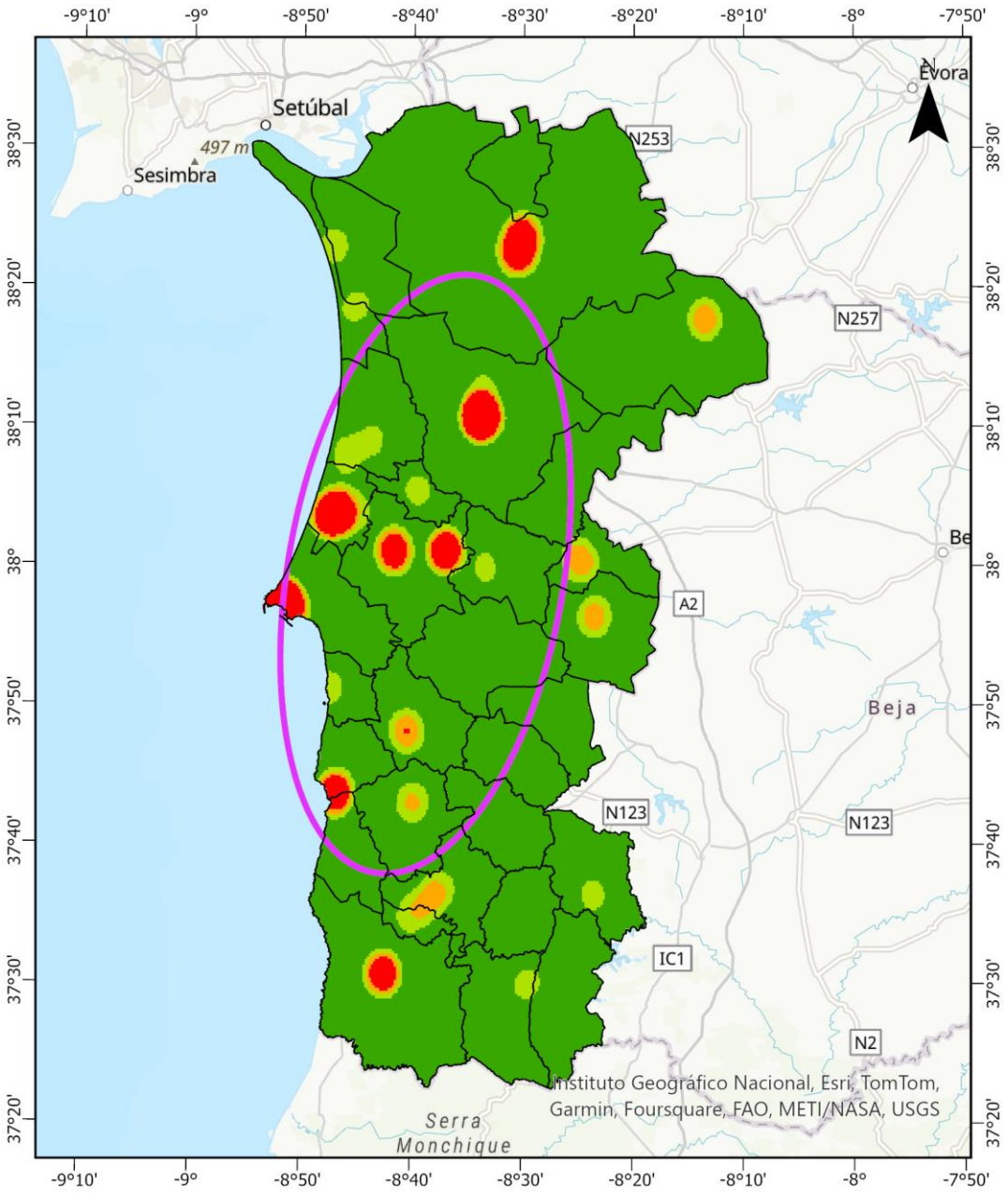


Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

Legend
ULSLA_Freg
BD22_IN_Directional
VALUE
1 - 37 328
37 329 - 81 737
81 738 - 126 146
126 147 - 564 643

Directional Distribution of Patients with T90 in 2023



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

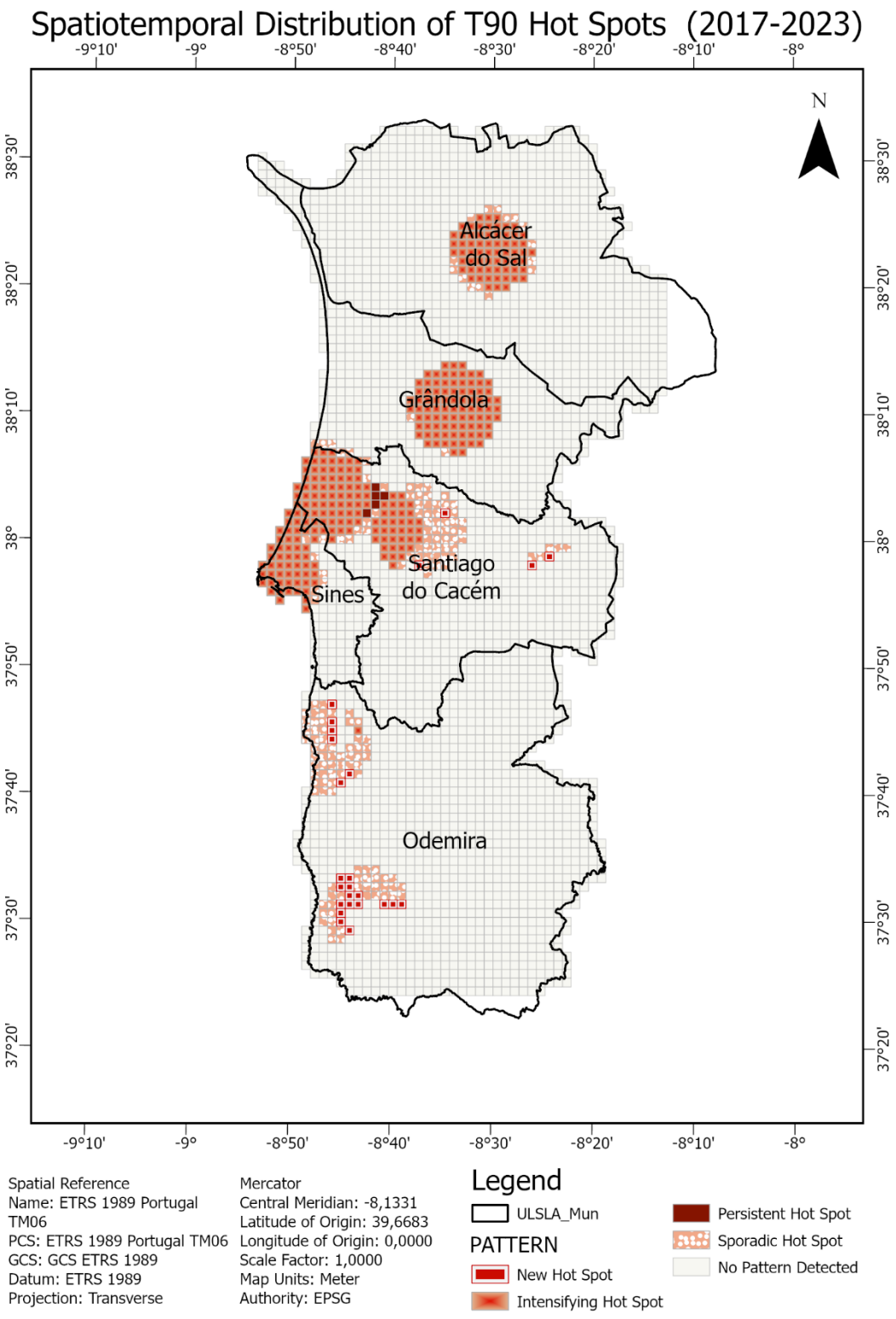
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

Legend

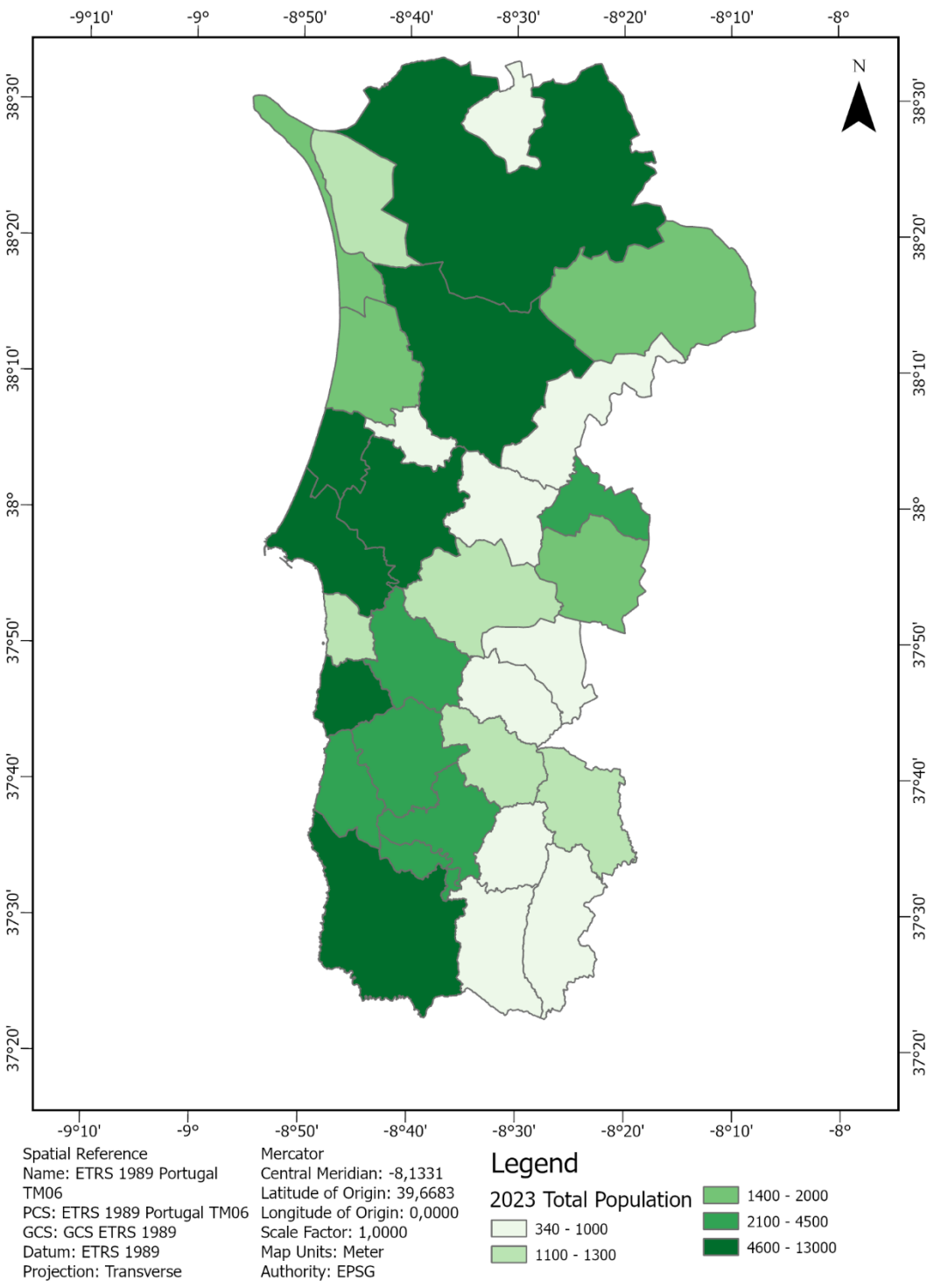
ULSLA_Freg
BD23_IN_Directional

VALUE

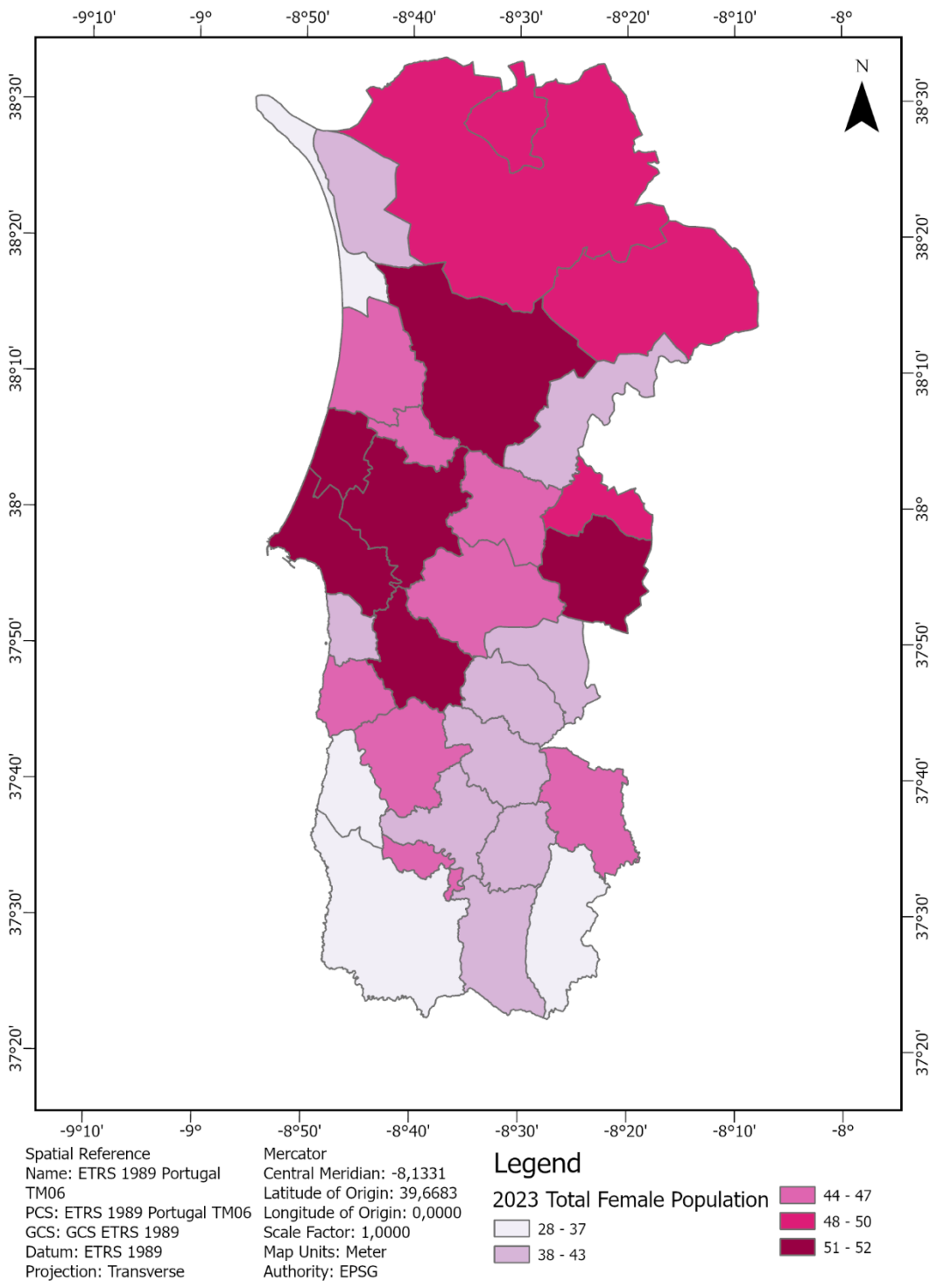
1 - 38 715
38 716 - 85 053
85 054 - 131 390
131 391 - 590 048


Spatiotemporal Distribution of T90 Hot Spots (2017-2023)
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
N
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
Alcácer do Sal
Grândola
Santiago do Cacém
Sines
Odemira
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ULSLA_Mun
PATTERN
New Hot Spot
Intensifying Hot Spot
Persistent Hot Spot
Sporadic Hot Spot
No Pattern Detected

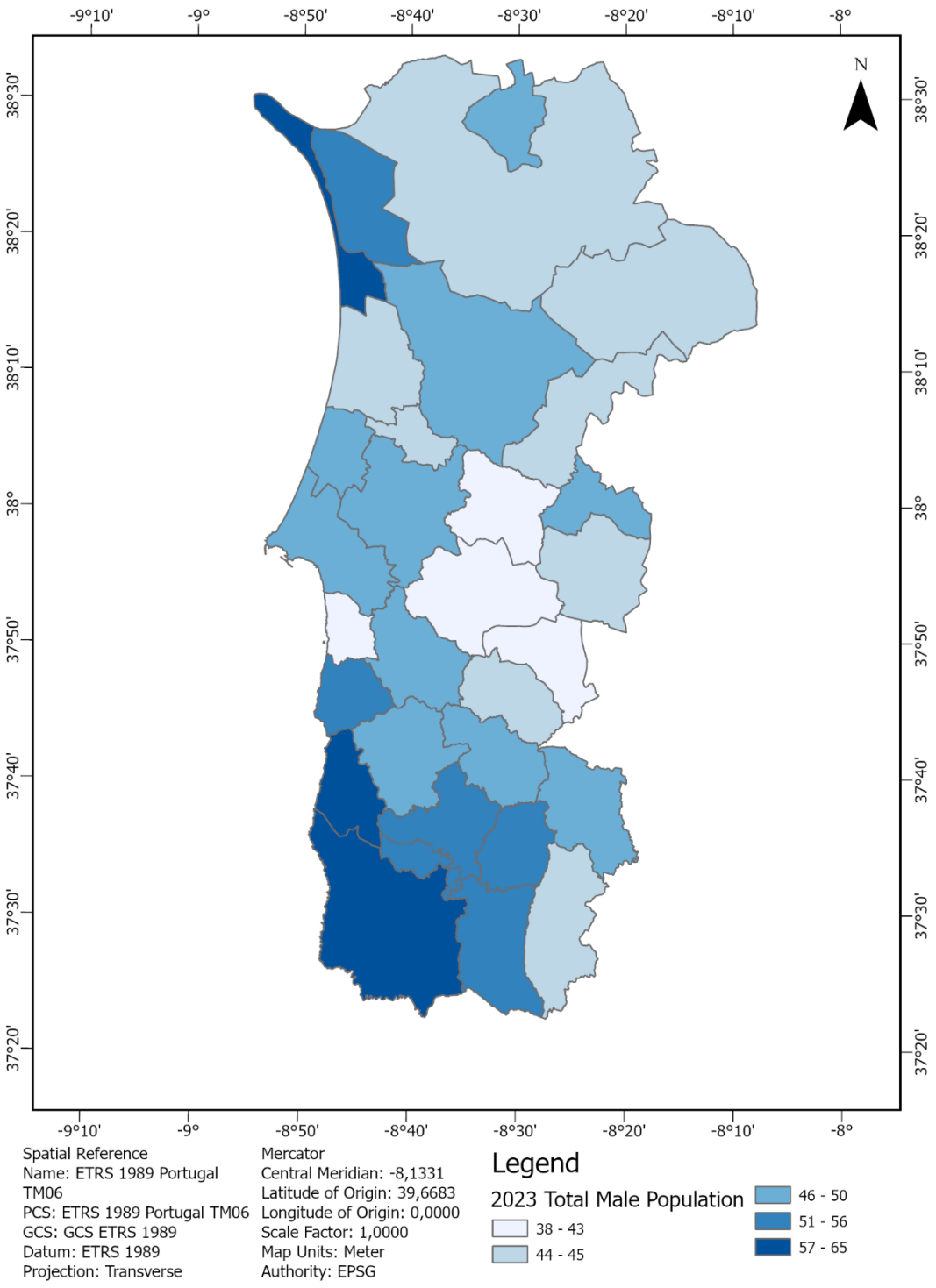
Total Population by Parish in 2023



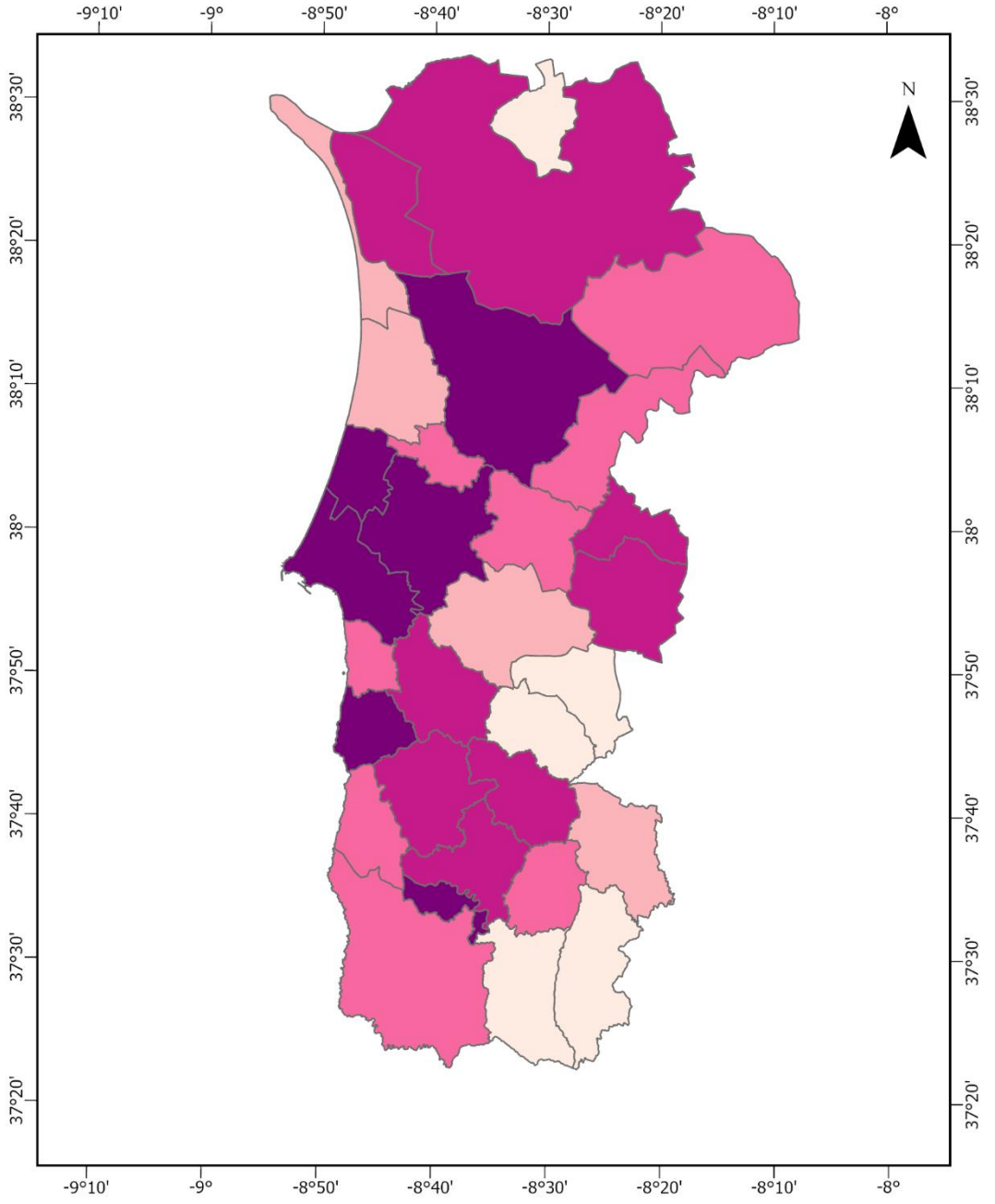
Percentage of Female Population by Parish in 2023



Percentage of Male Population by Parish in 2023



Percentage of Population Aged 0-14 by Parish in 2023



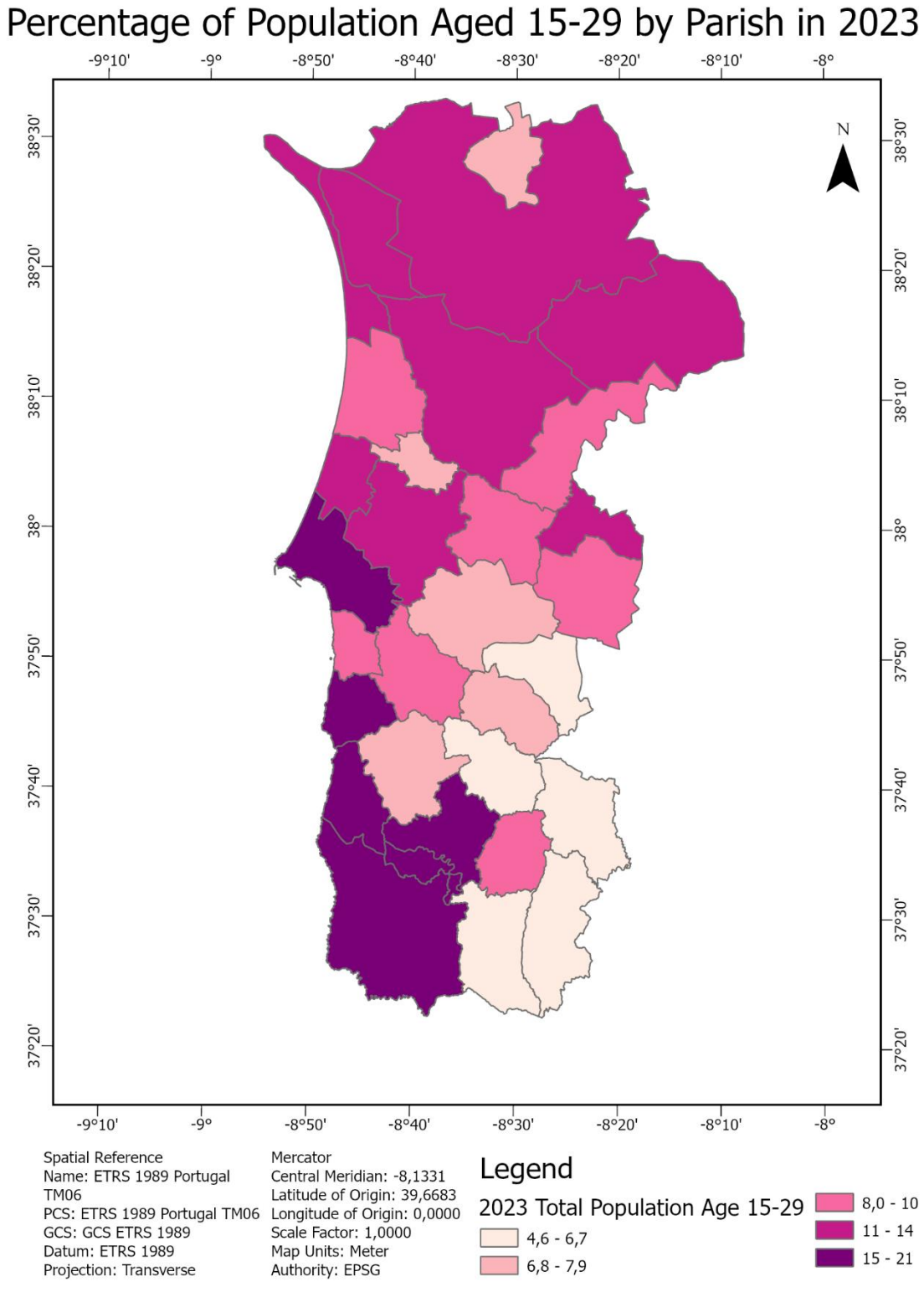
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

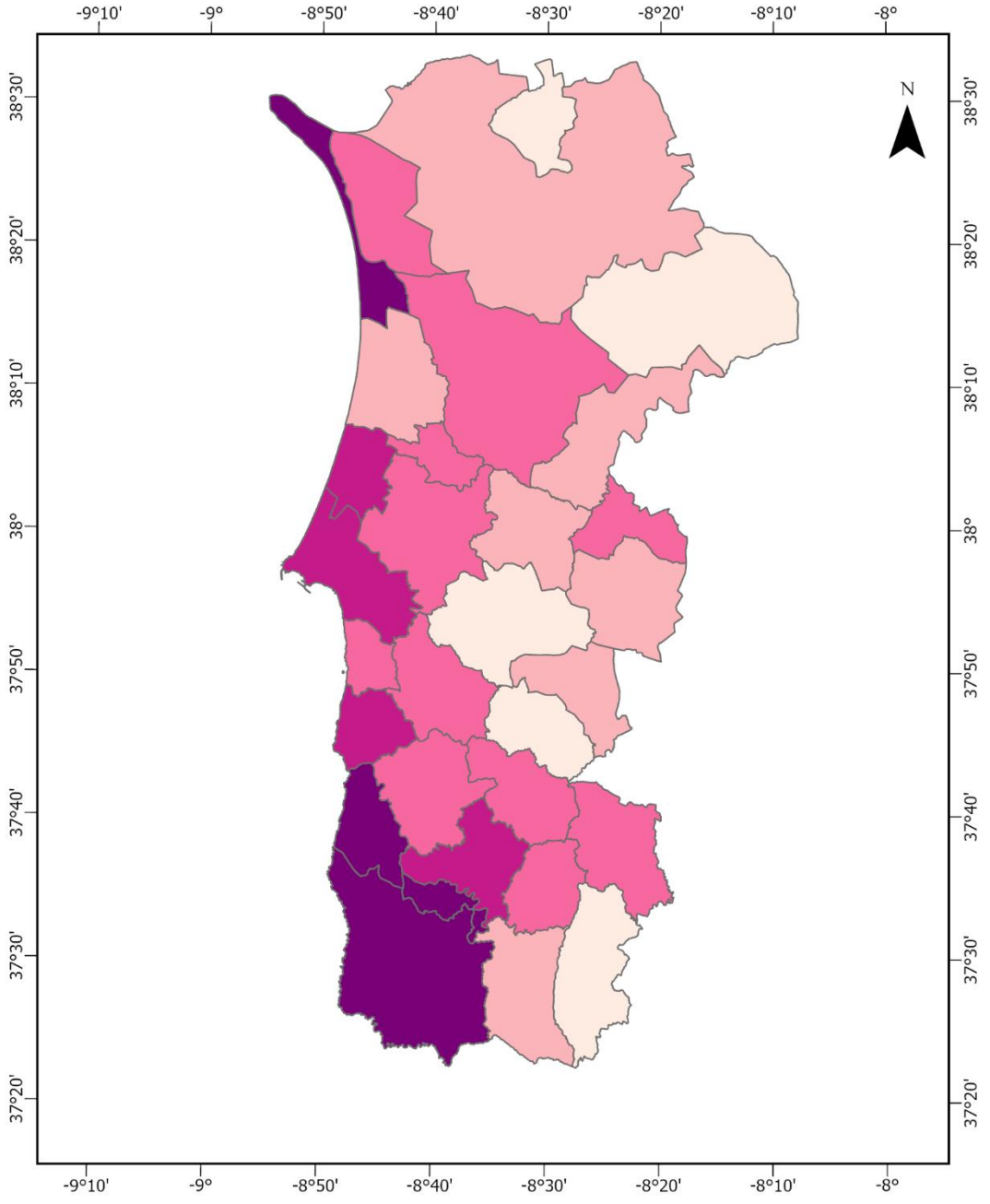
Legend

2023 Total Population Age 0-14

- 4,1 - 5,7
- 5,8 - 6,9
- 7,0 - 8,5
- 8,6 - 11
- 12 - 14


Percentage of Population Aged 15-29 by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
2023 Total Population Age 15-29
4,6 - 6,7
6,8 - 7,9
8,0 - 10
11 - 14
15 - 21

Percentage of Population Aged 30-44 by Parish in 2023

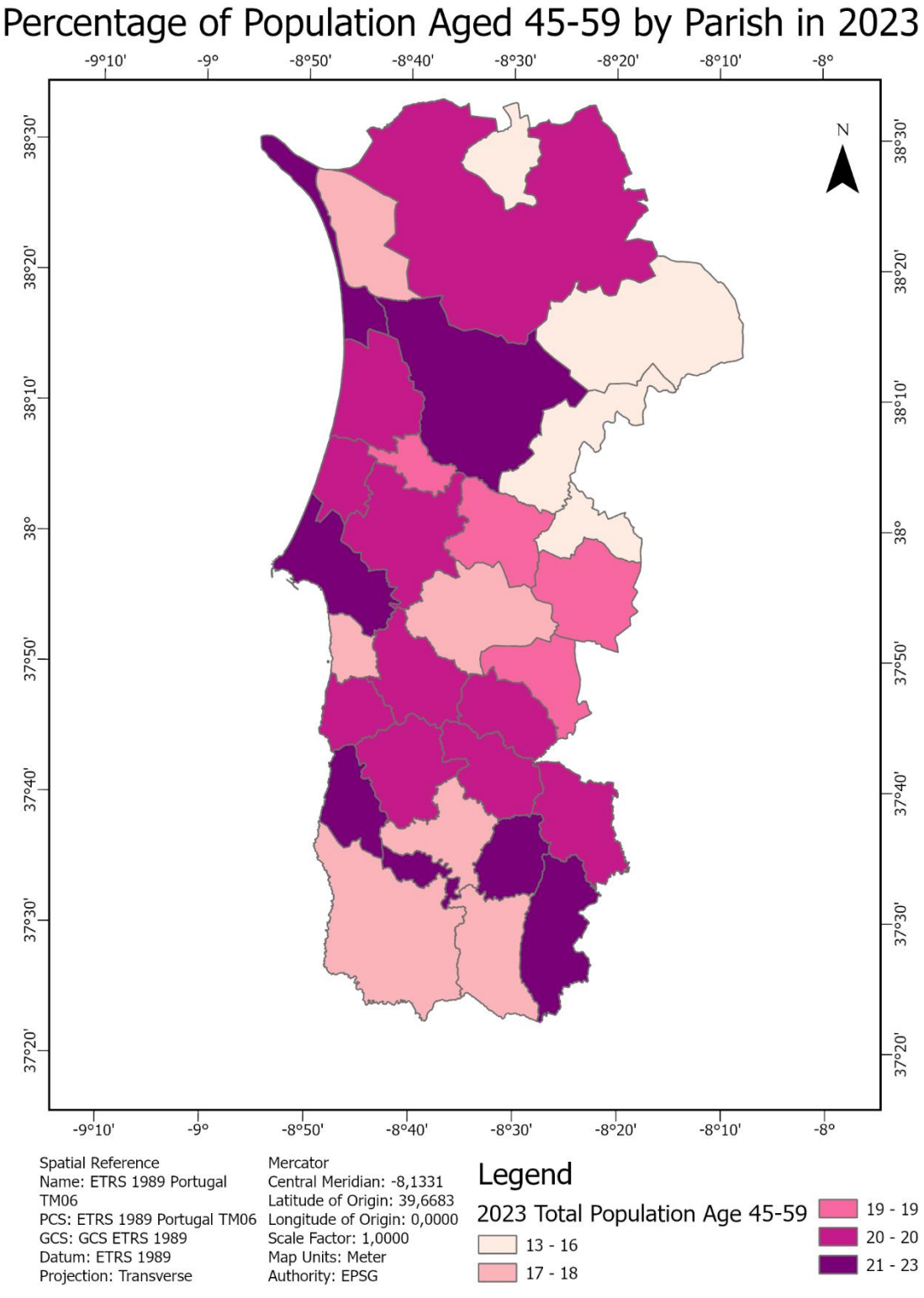


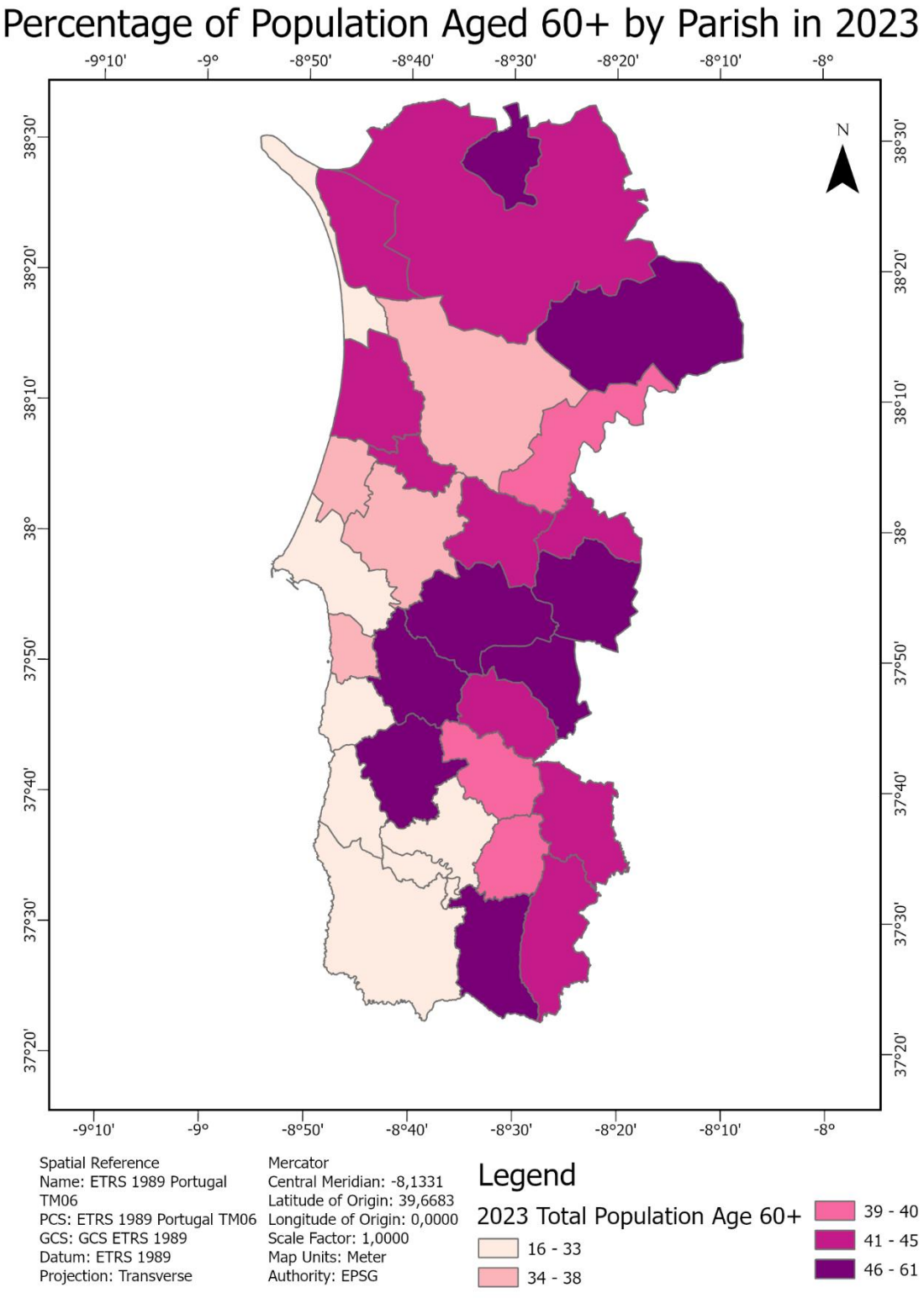
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

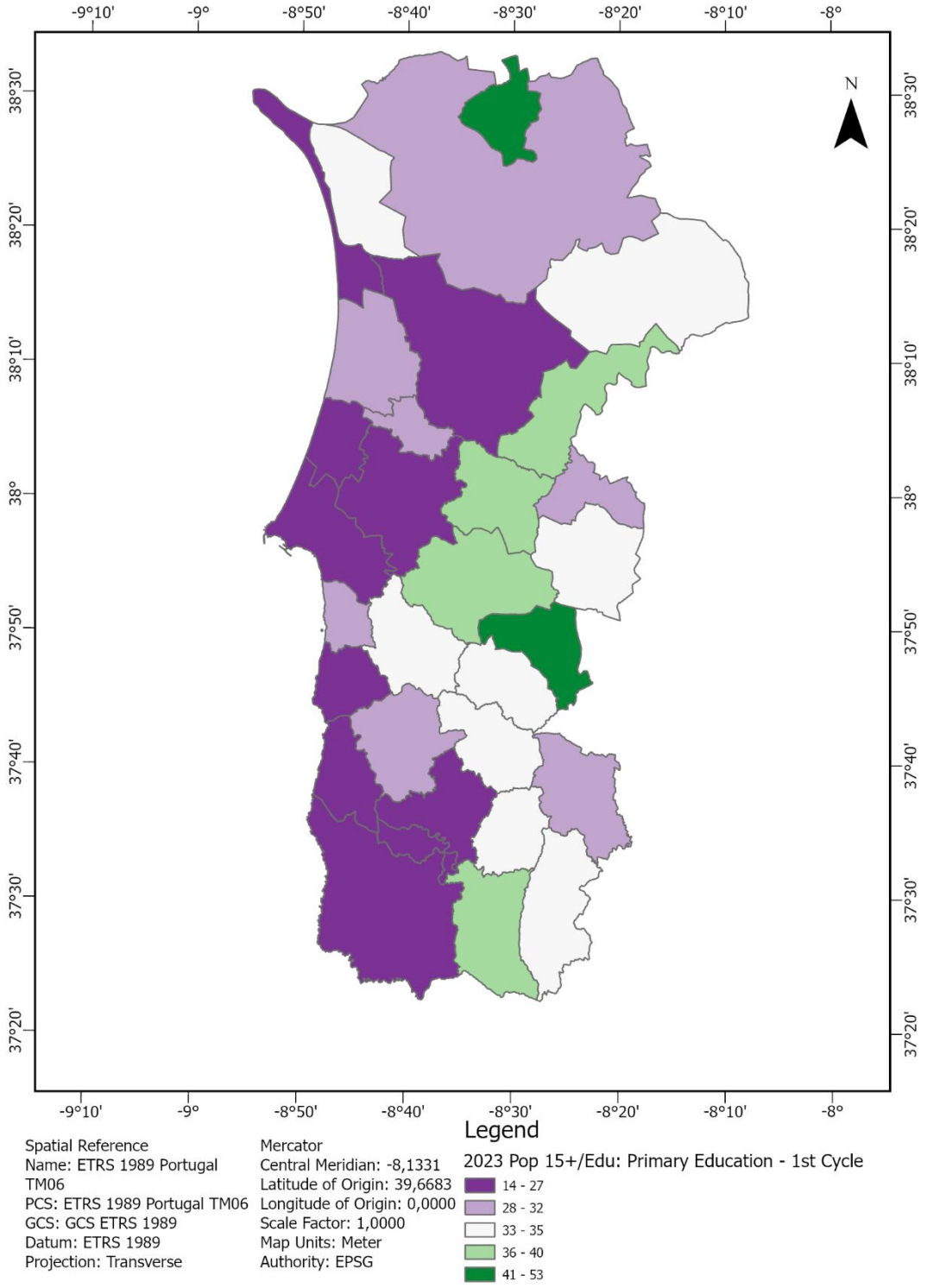
Legend

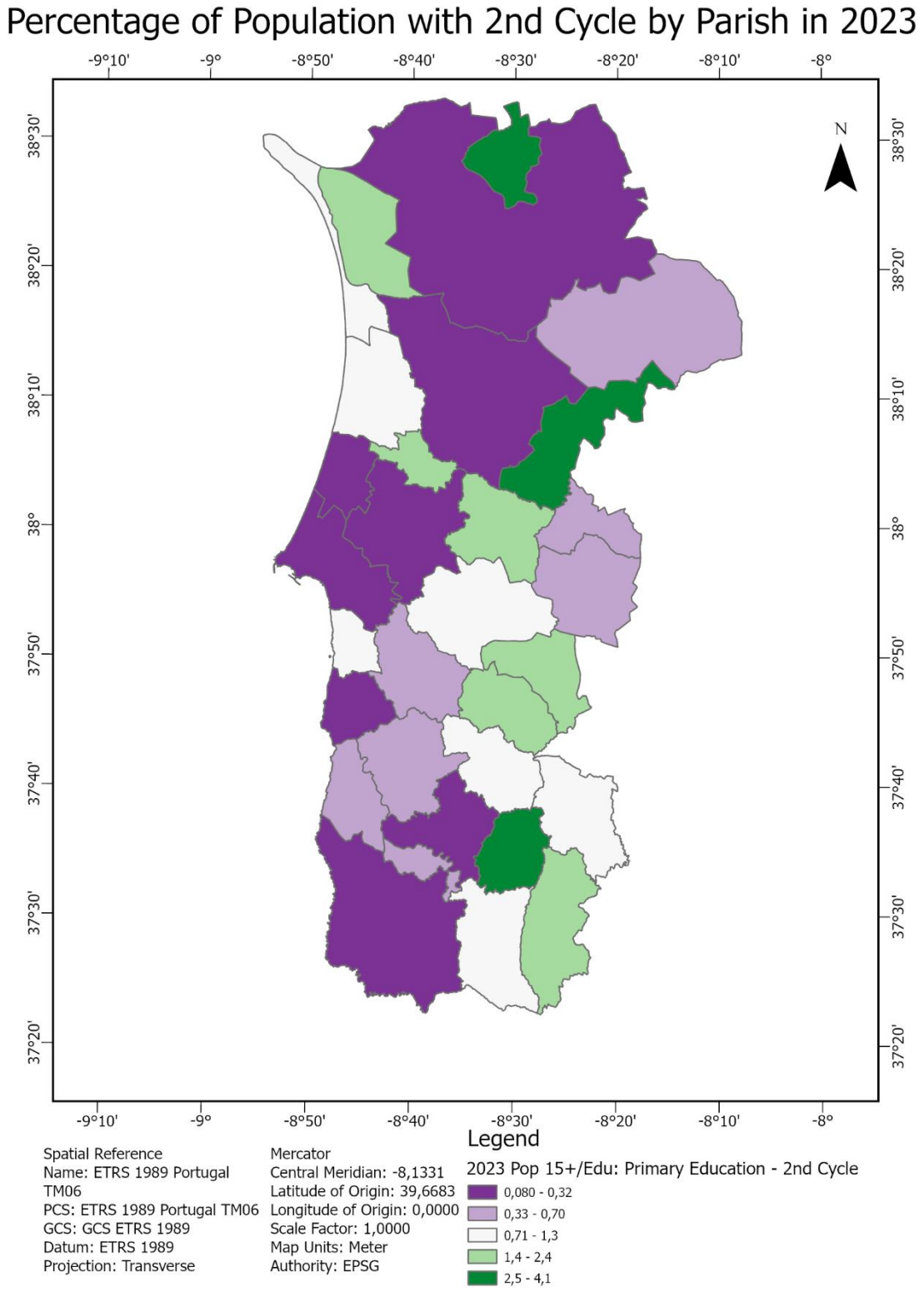
2023 Total Population Age 30-44
- 3,5 - 9,6
- 9,7 - 13
- 14 - 15
- 16 - 18
- 19 - 24
- 25 - 35

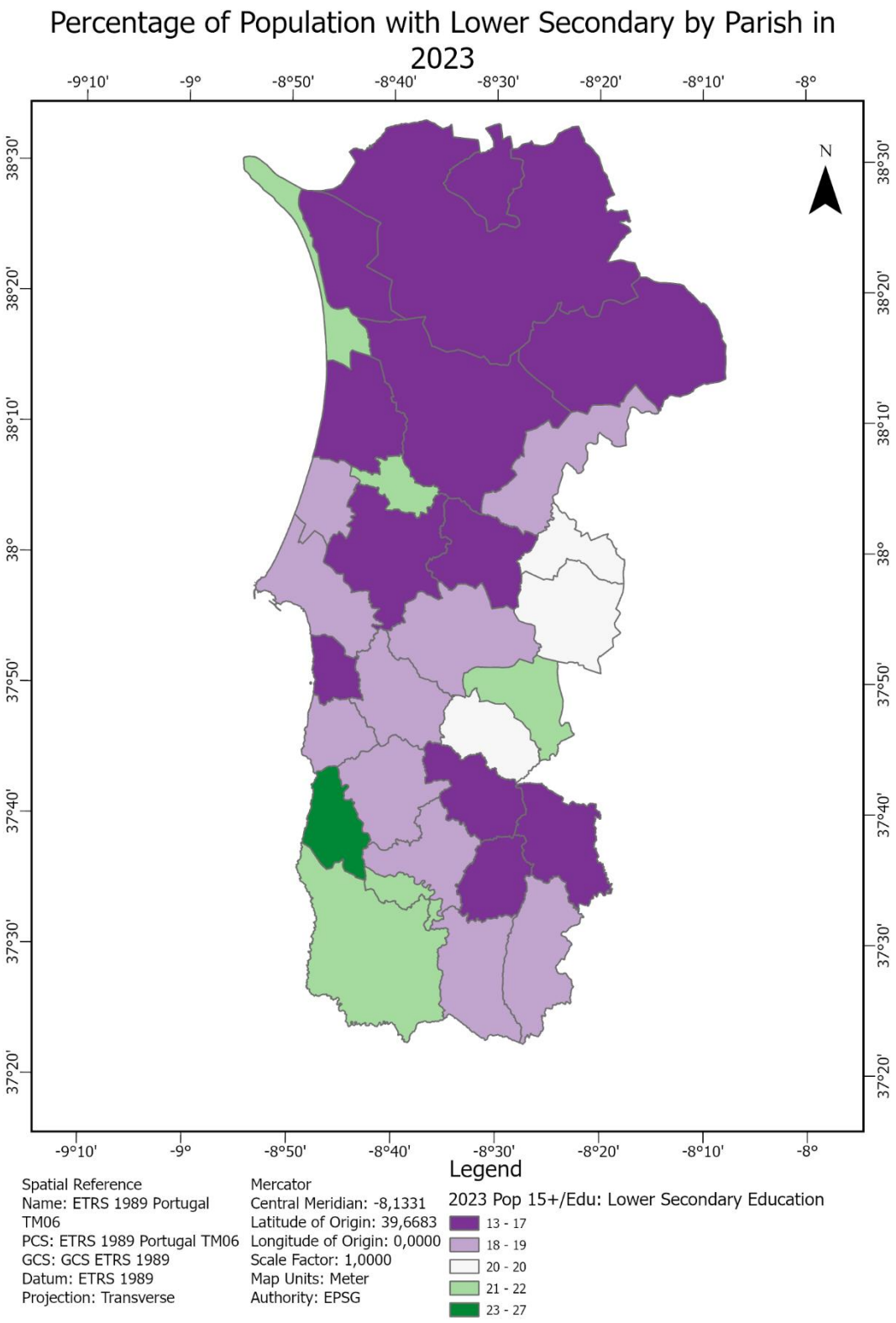

Percentage of Population Aged 45-59 by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
2023 Total Population Age 45-59
13 - 16
17 - 18
19 - 19
20 - 20
21 - 23

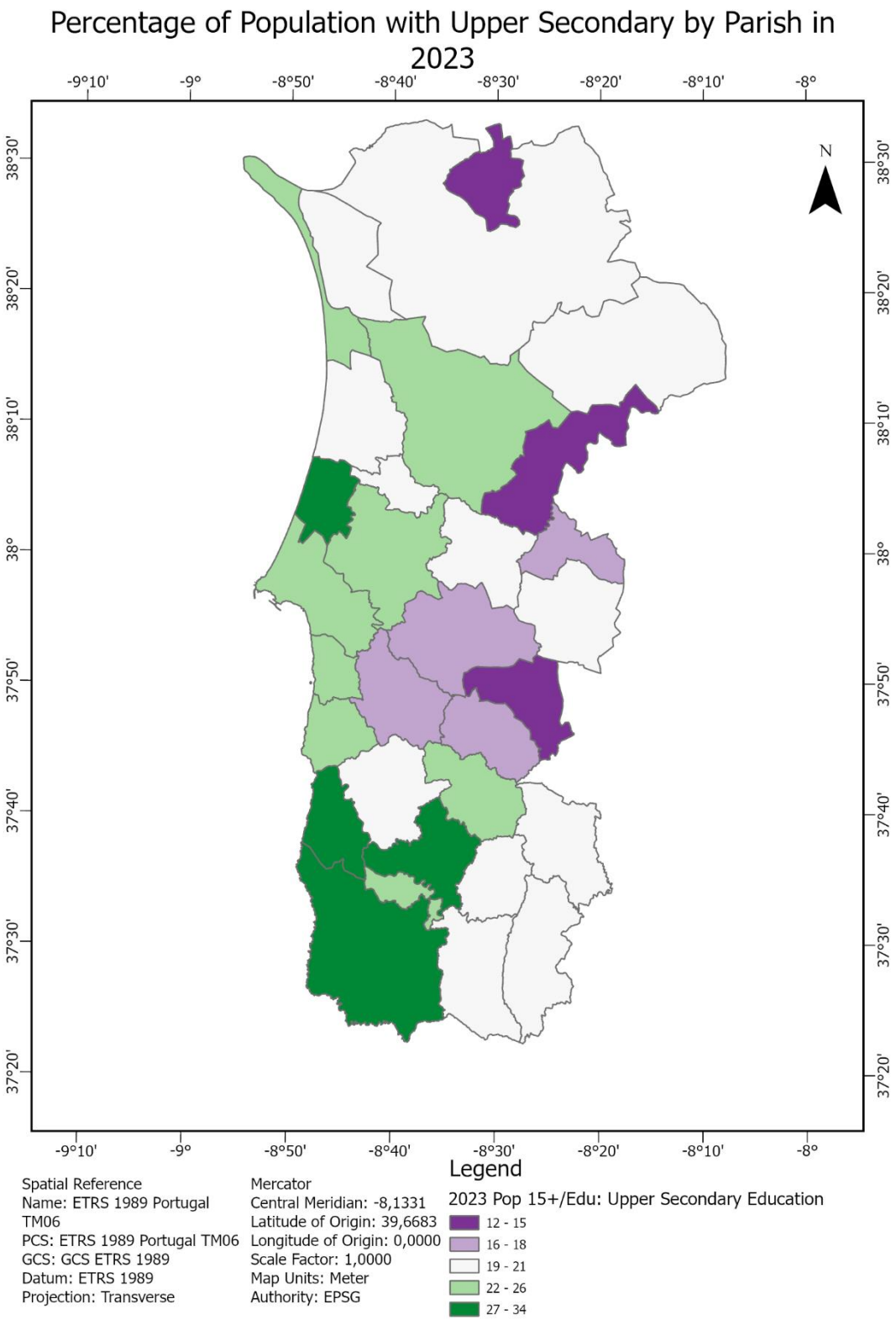

Percentage of Population Aged 60+ by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
2023 Total Population Age 60+
16 - 33
34 - 38
39 - 40
41 - 45
46 - 61

Percentage of Population with 1st Cycle by Parish in 2023

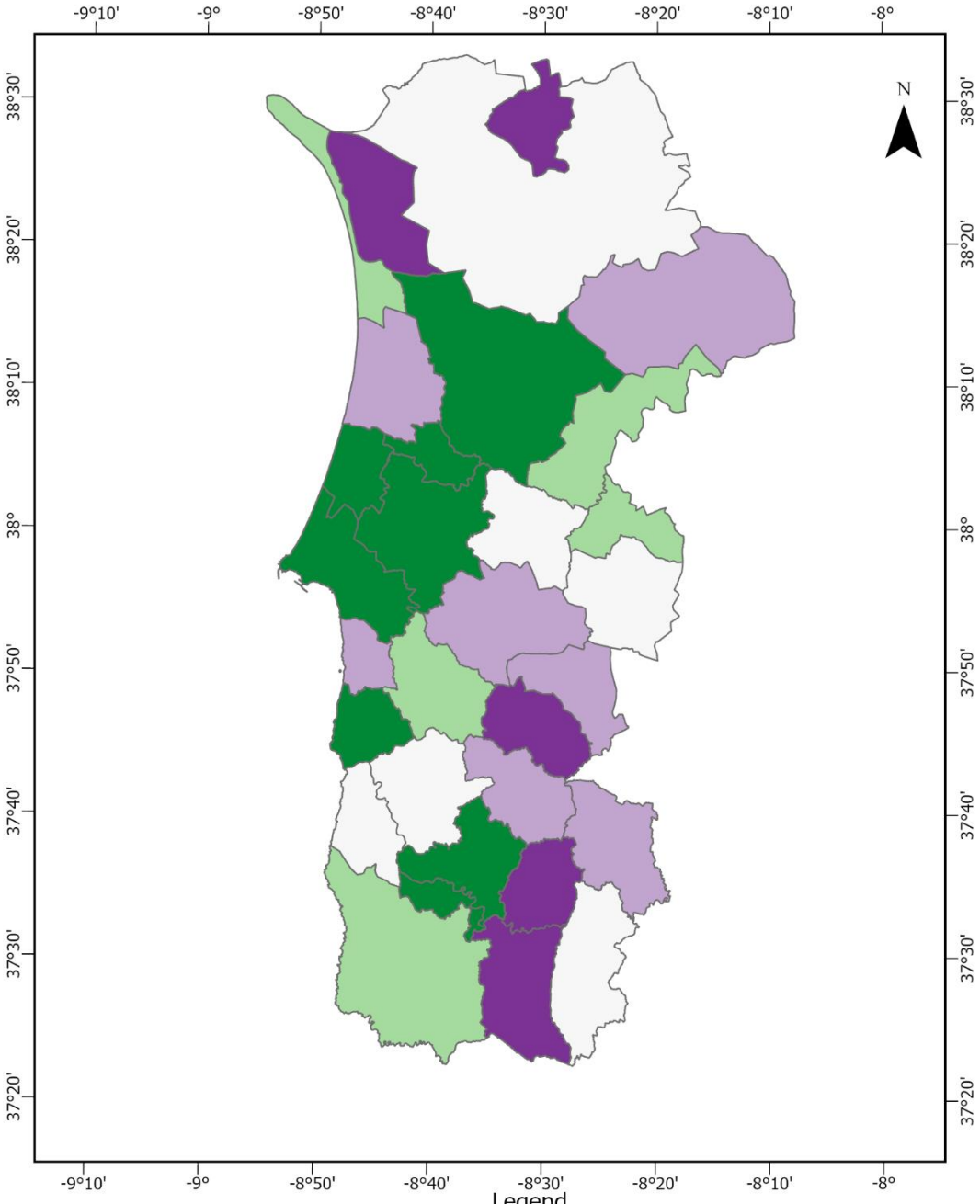



Percentage of Population with 2nd Cycle by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Legend
2023 Pop 15+/Edu: Primary Education - 2nd Cycle
0,080 - 0,32
0,33 - 0,70
0,71 - 1,3
1,4 - 2,4
2,5 - 4,1
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG


Percentage of Population with Lower Secondary by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Legend
2023 Pop 15+/Edu: Lower Secondary Education
13 - 17
18 - 19
20 - 20
21 - 22
23 - 27
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG


Percentage of Population with Upper Secondary by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Legend
2023 Pop 15+/Edu: Upper Secondary Education
12 - 15
16 - 18
19 - 21
22 - 26
27 - 34
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

Percentage of Population with Post Secondary by Parish in 2023

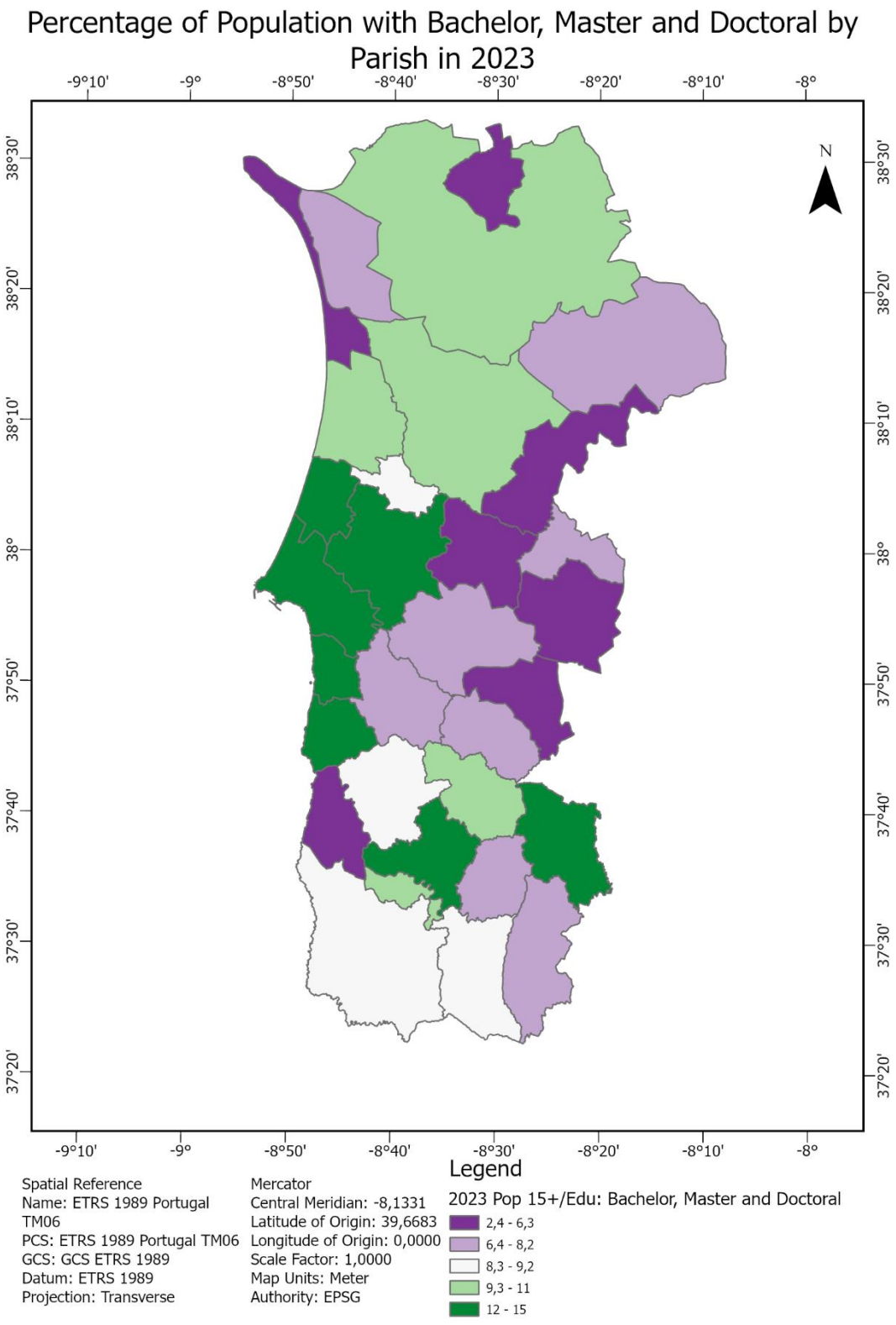


Legend

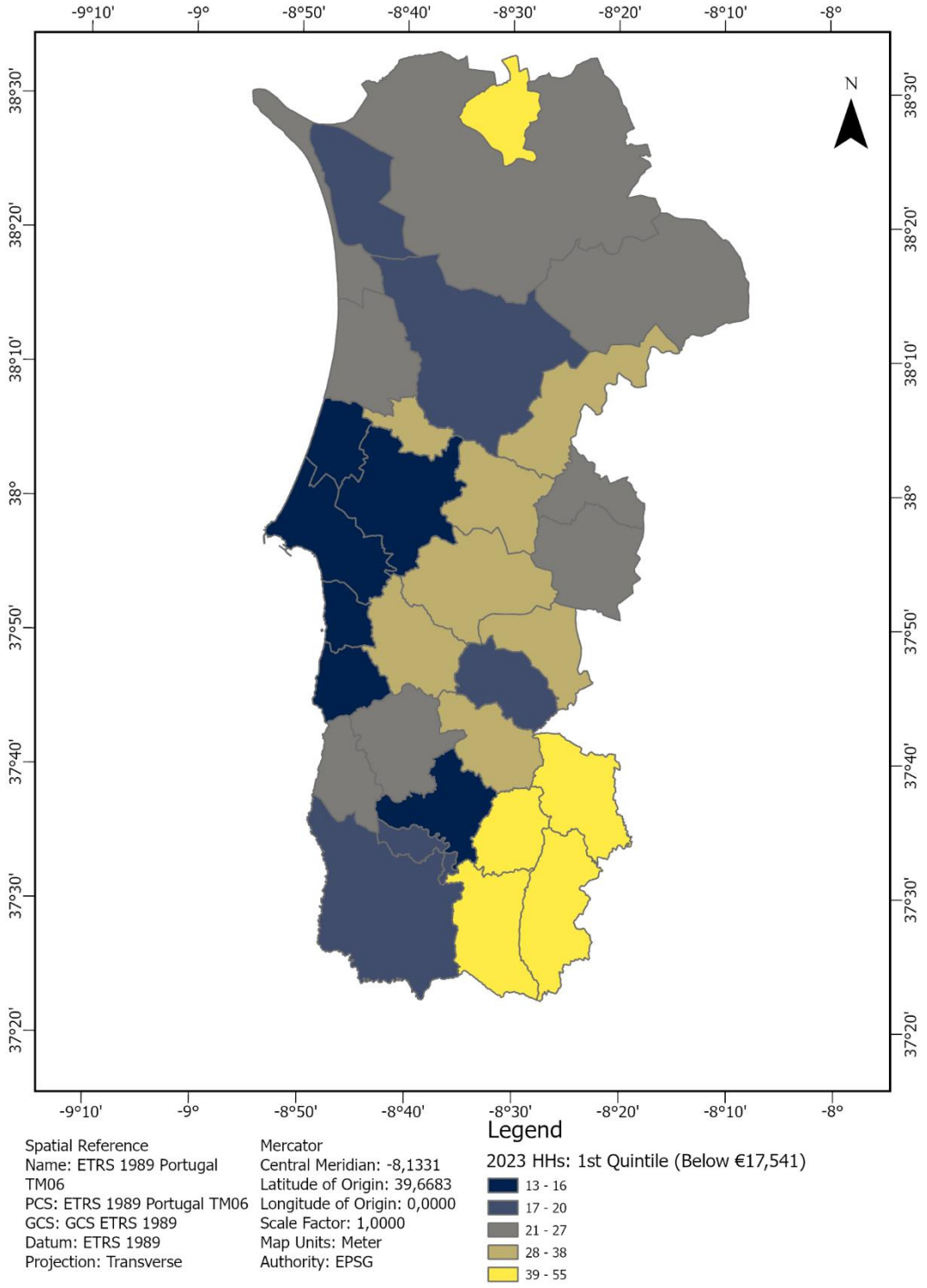
2023 Pop 15+/Edu: Post Secondary Non-Tertiary Education

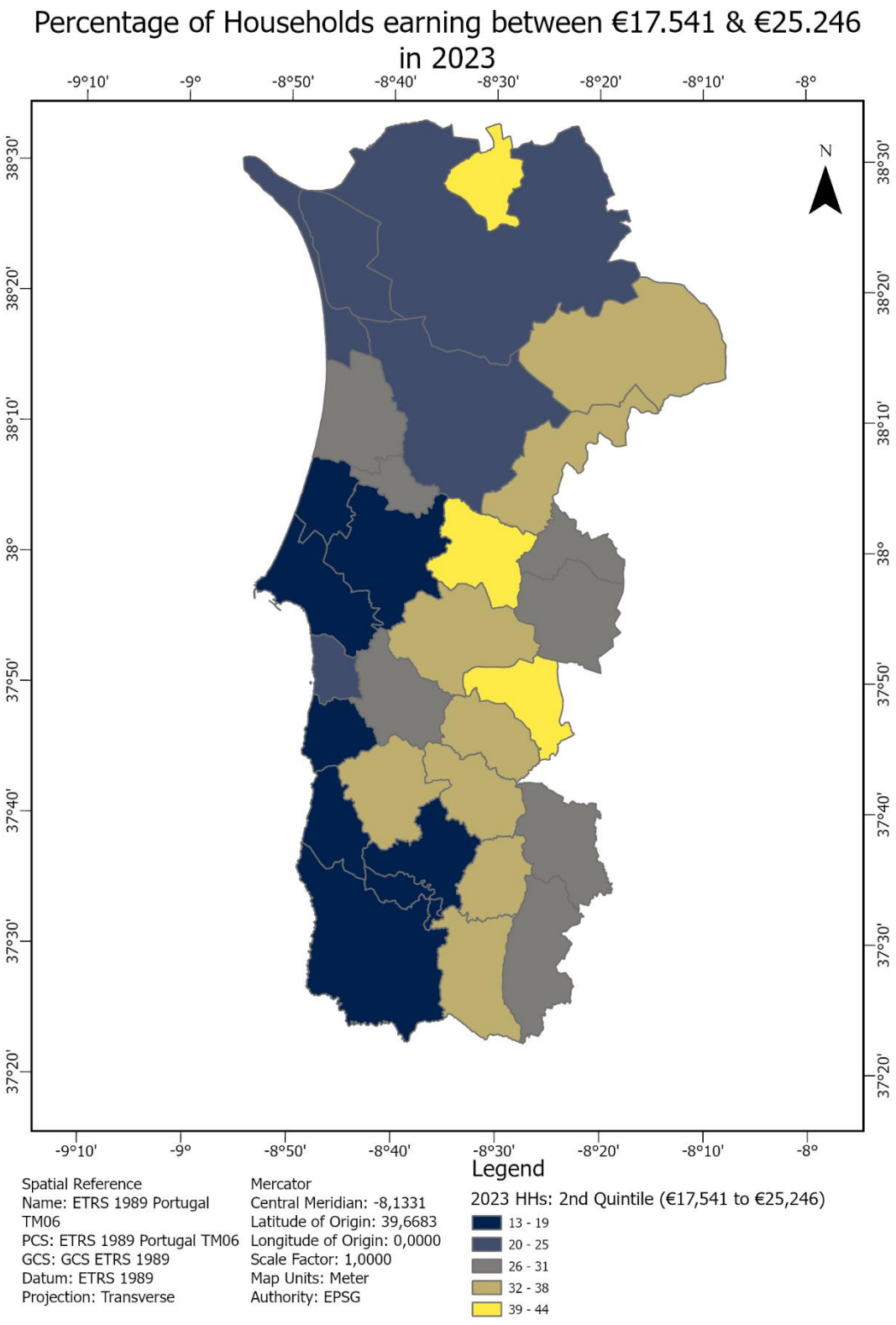
- 0,0 - 0,044
- 0,045 - 0,13
- 0,14 - 0,32
- 0,33 - 0,70
- 0,71 - 1,5

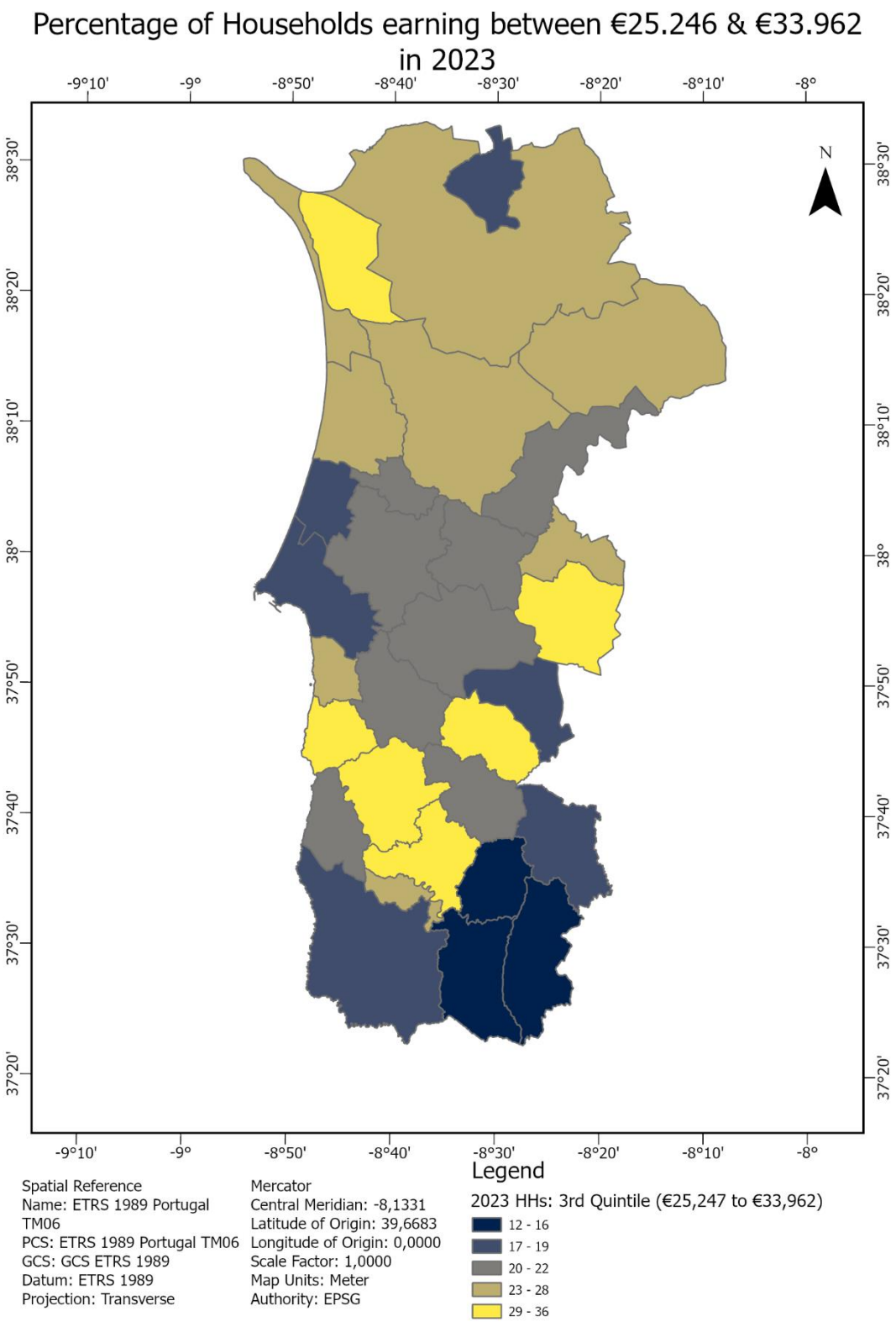
Spatial Reference	Mercator
Name: ETRS 1989 Portugal TM06	Central Meridian: -8,1331
PCS: ETRS 1989 Portugal TM06	Latitude of Origin: 39,6683
GCS: GCS ETRS 1989	Longitude of Origin: 0,0000
Datum: ETRS 1989	Scale Factor: 1,0000
Projection: Transverse	Map Units: Meter
	Authority: EPSG

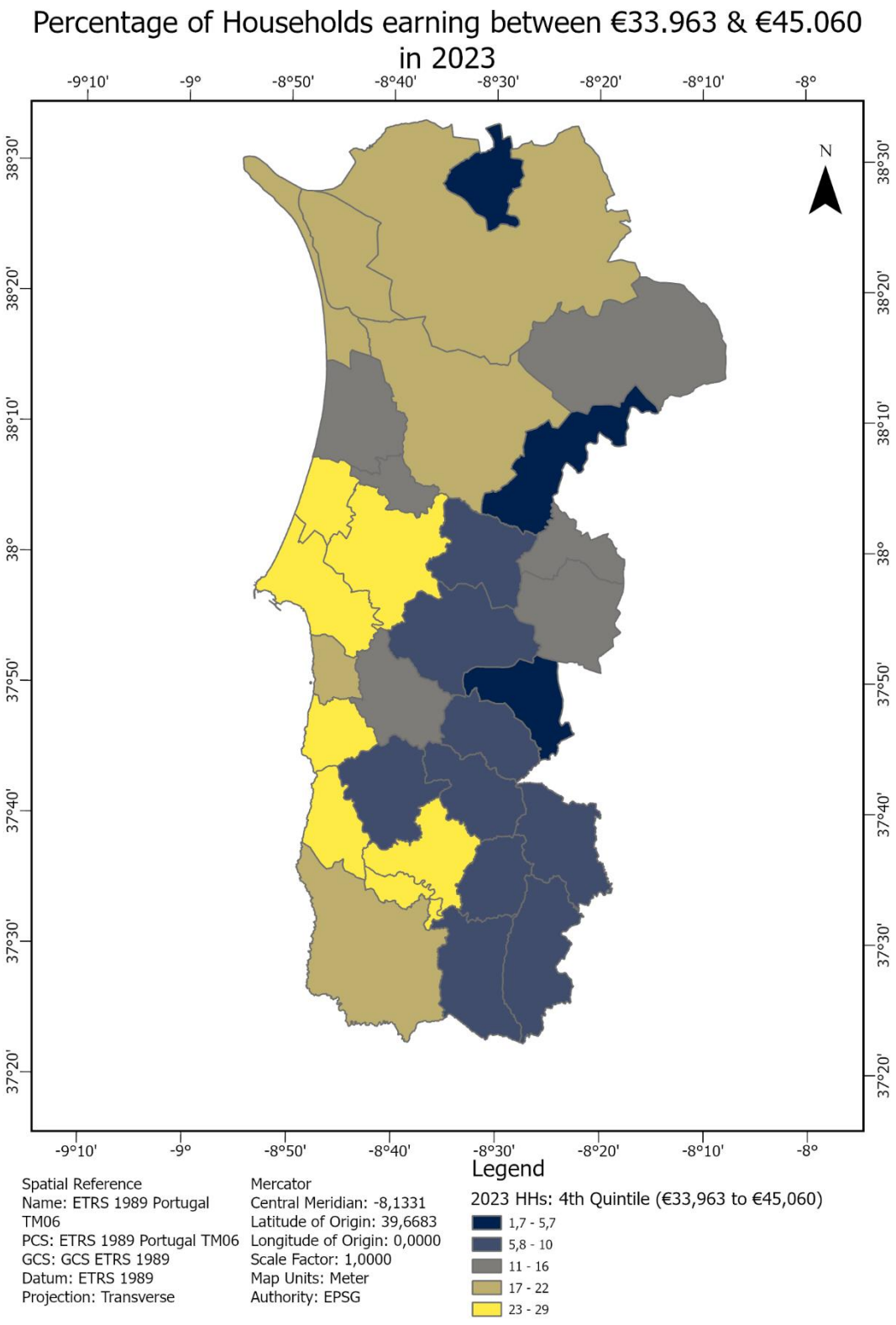

Percentage of Population with Bachelor, Master and Doctoral by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Legend
2023 Pop 15+/Edu: Bachelor, Master and Doctoral
2,4 - 6,3
6,4 - 8,2
8,3 - 9,2
9,3 - 11
12 - 15
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

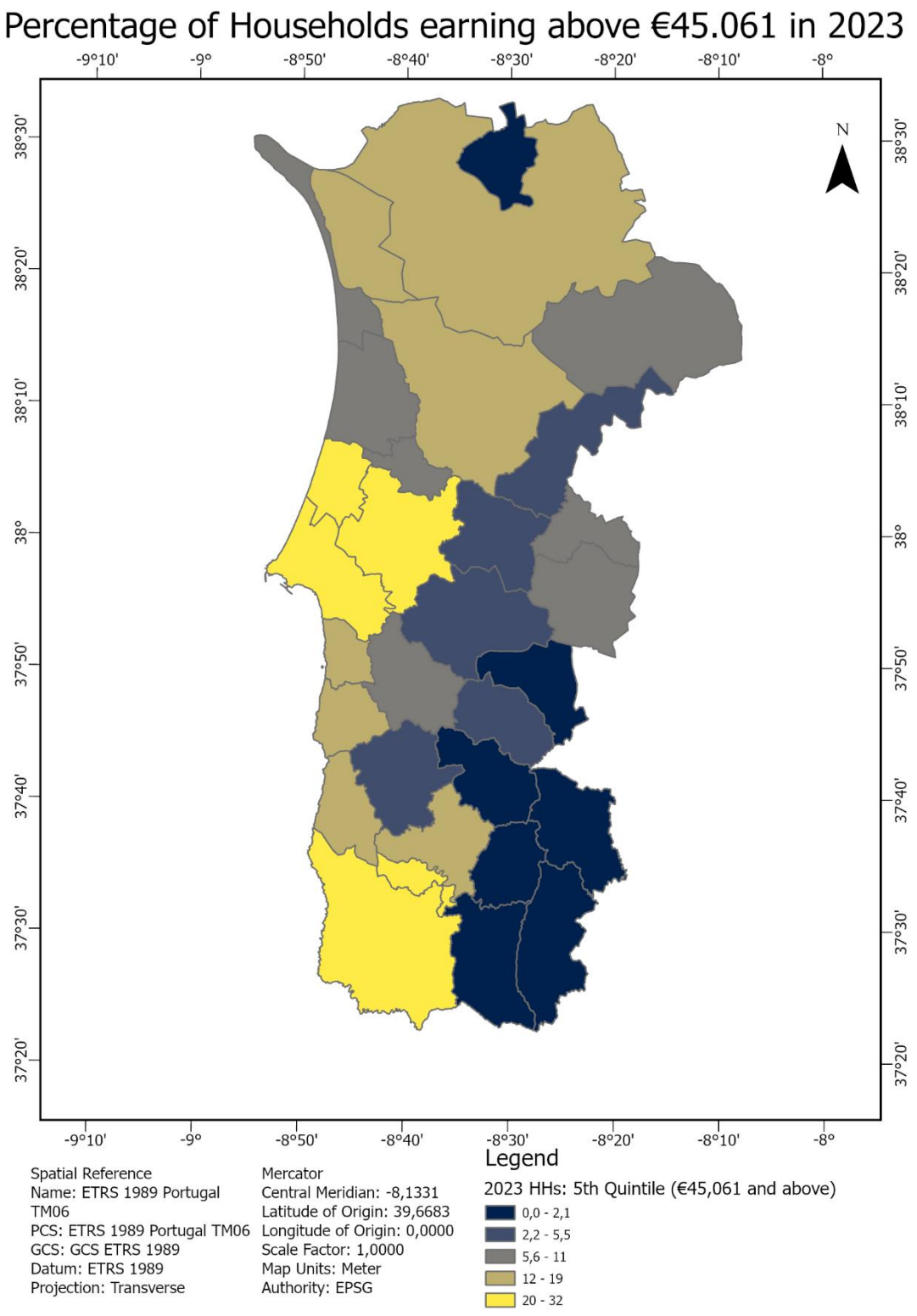
Percentage of Households earning below €17.541 in 2023

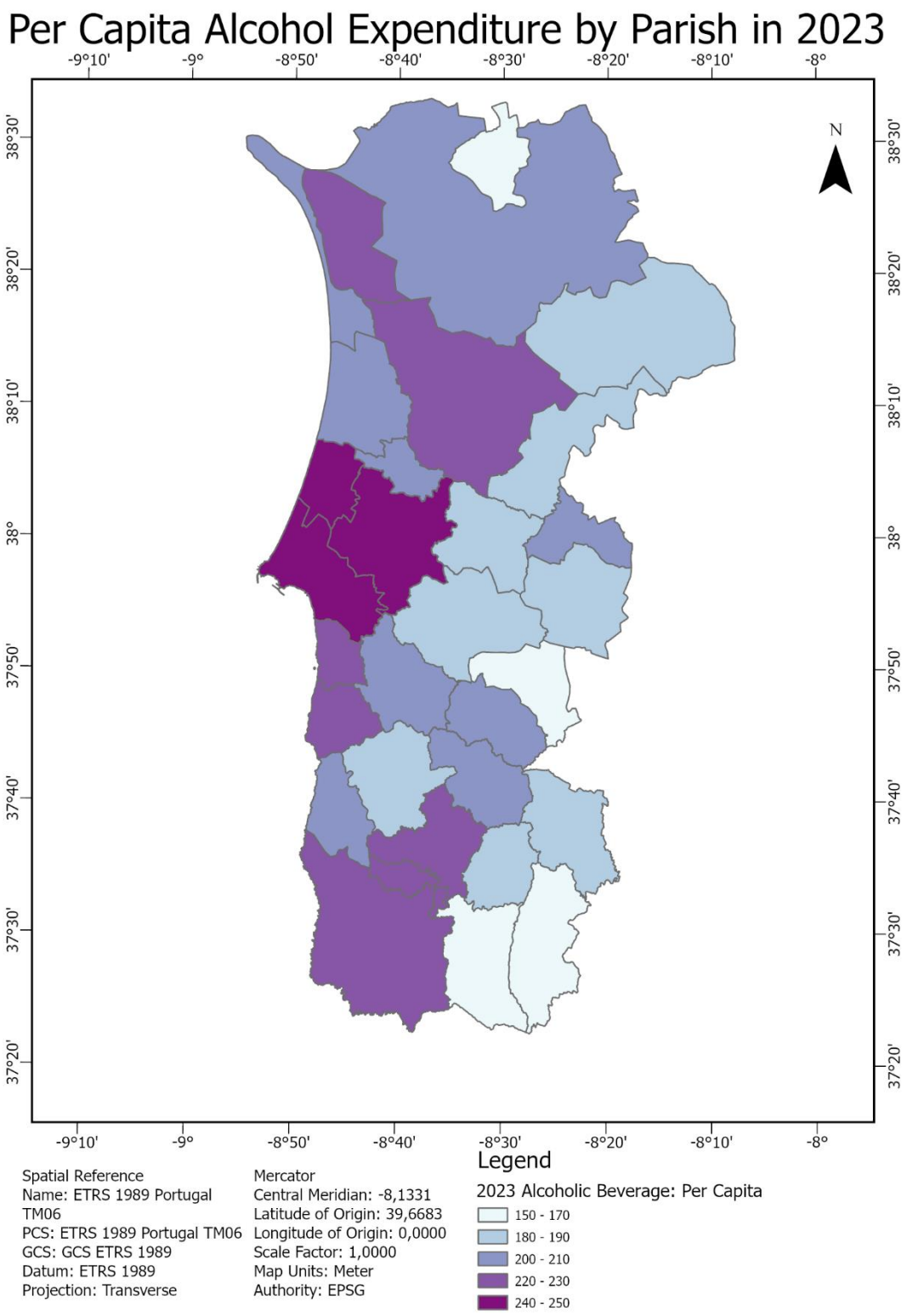



Percentage of Households earning between €17.541 & €25.246 in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
2023 HHs: 2nd Quintile (€17,541 to €25,246)
13 - 19
20 - 25
26 - 31
32 - 38
39 - 44

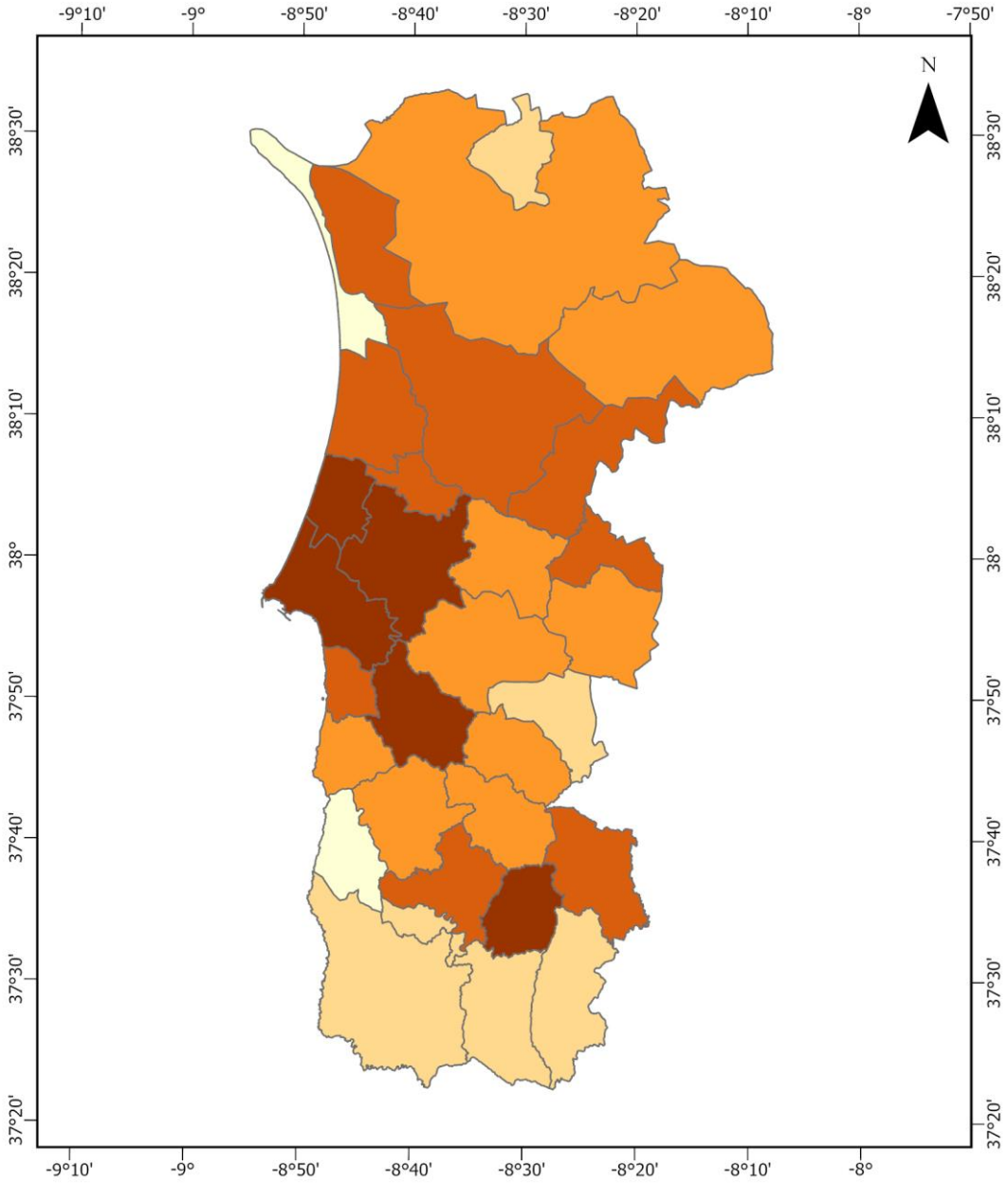

Percentage of Households earning between €25.246 & €33.962 in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
2023 HHs: 3rd Quintile (€25,247 to €33,962)
12 - 16
17 - 19
20 - 22
23 - 28
29 - 36


Percentage of Households earning between €33.963 & €45.060 in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
2023 HHs: 4th Quintile (€33,963 to €45,060)
1,7 - 5,7
5,8 - 10
11 - 16
17 - 22
23 - 29


Percentage of Households earning above €45.061 in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
2023 HHs: 5th Quintile (€45,061 and above)
0,0 - 2,1
2,2 - 5,5
5,6 - 11
12 - 19
20 - 32


Per Capita Alcohol Expenditure by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
2023 Alcoholic Beverage: Per Capita
150 - 170
180 - 190
200 - 210
220 - 230
240 - 250

Per Capita Tobacco Expenditure by Parish in 2023



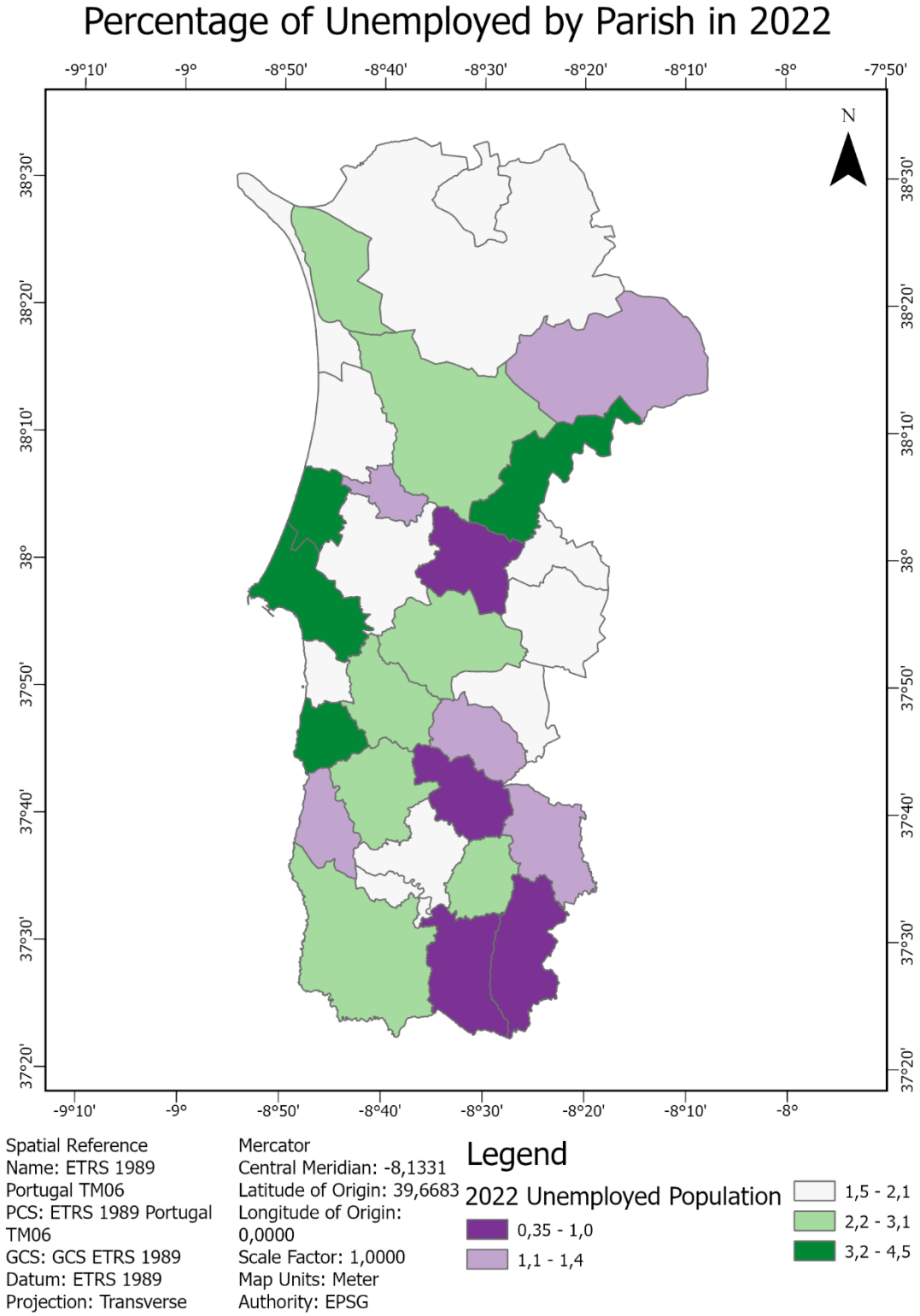
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

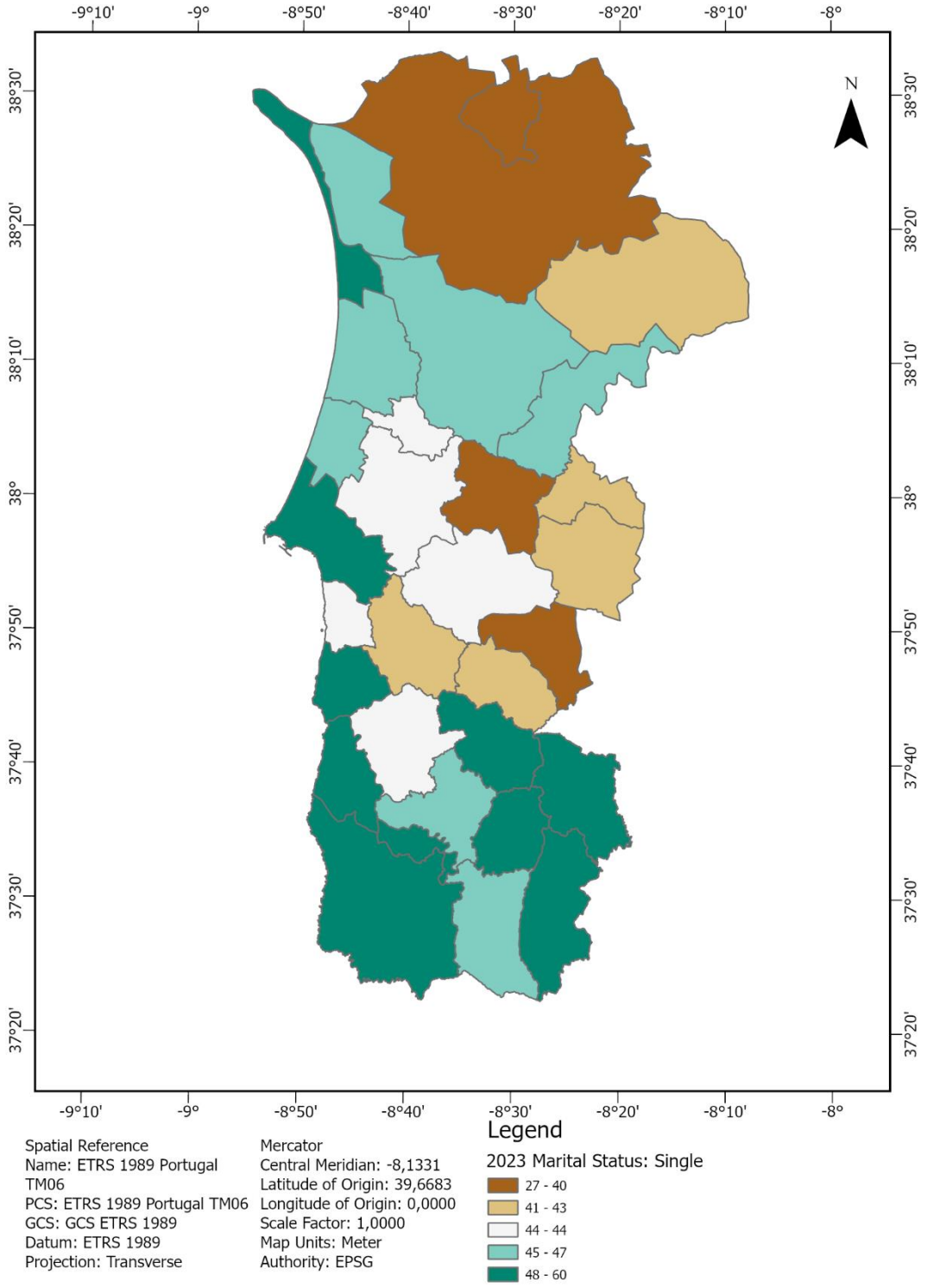
Legend

2023 Tobacco: Per Capita

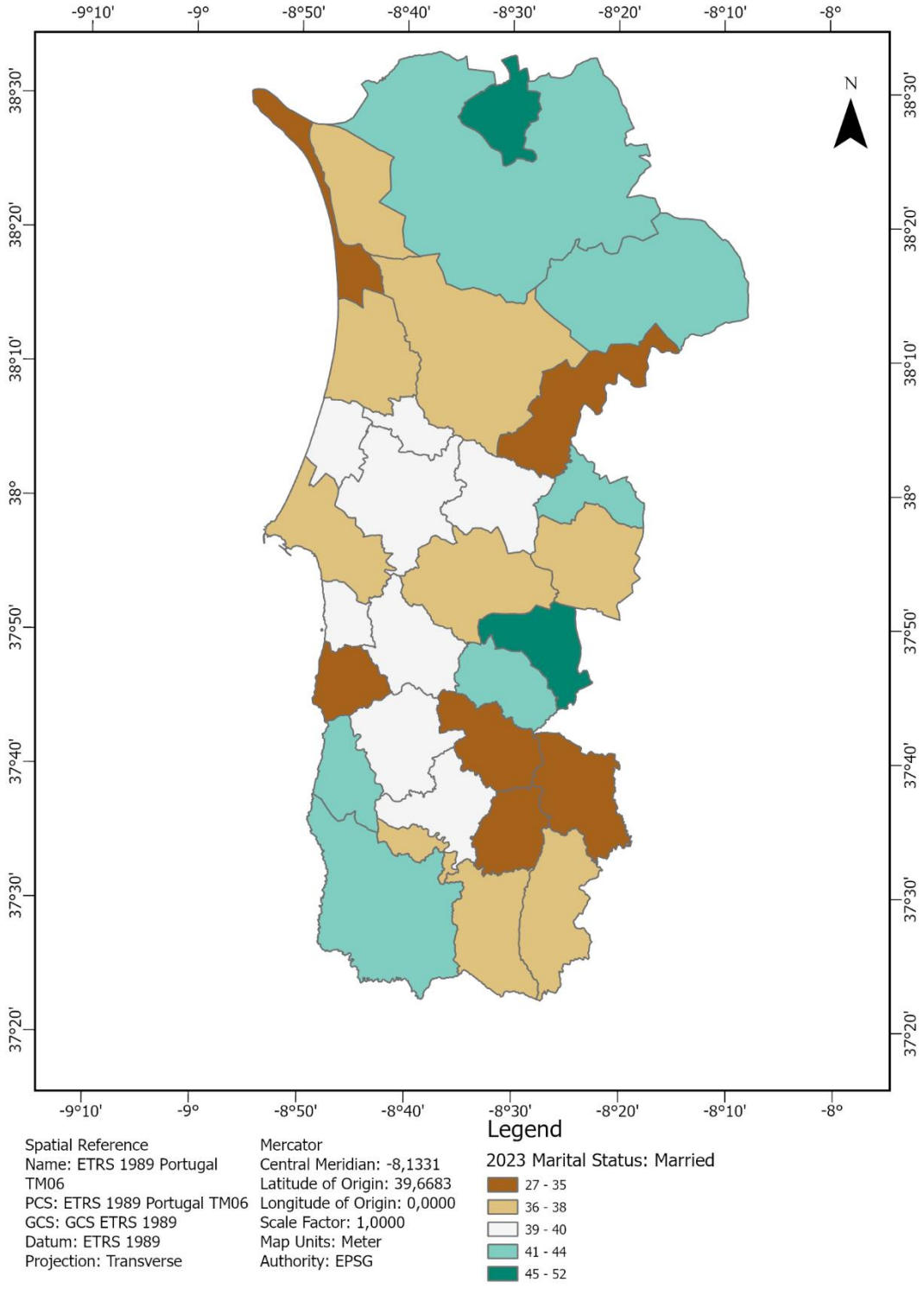
- 210 - 230
- 231 - 247
- 248 - 253
- 254 - 257
- 258 - 264


Percentage of Unemployed by Parish in 2022
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
-7°50'
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
2022 Unemployed Population
0,35 - 1,0
1,1 - 1,4
1,5 - 2,1
2,2 - 3,1
3,2 - 4,5

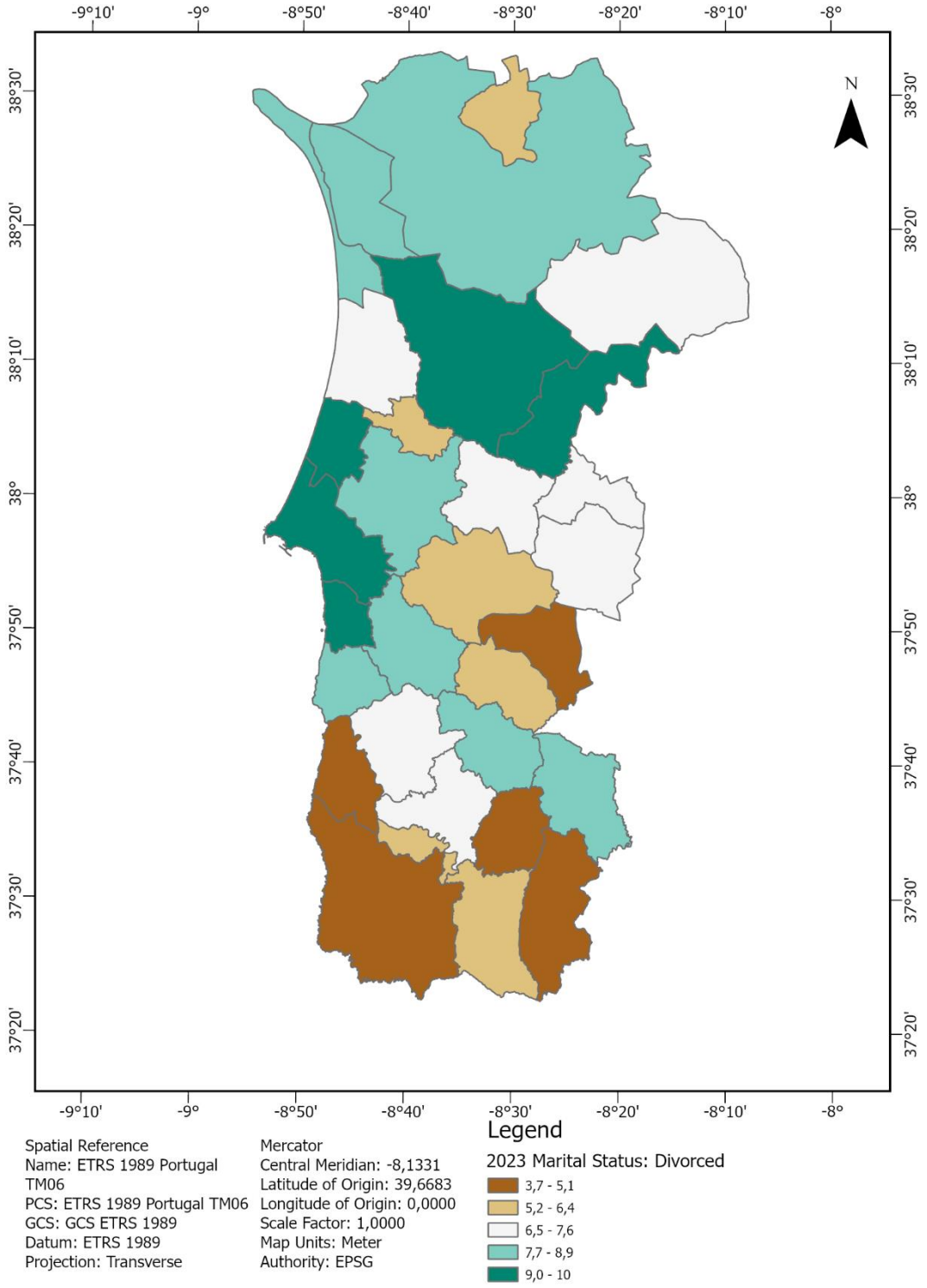
Percentage of Population that is Single by Parish in 2023

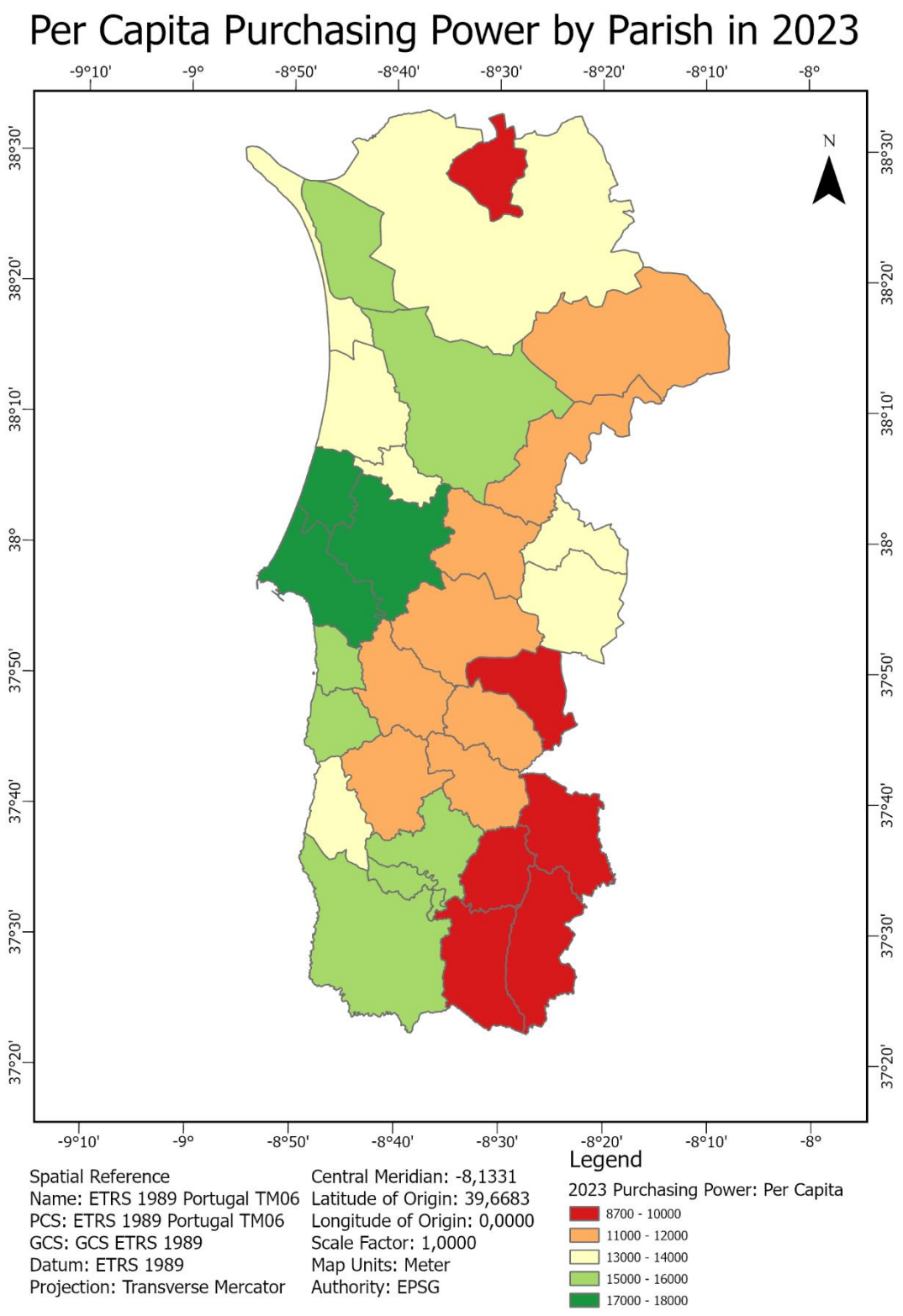


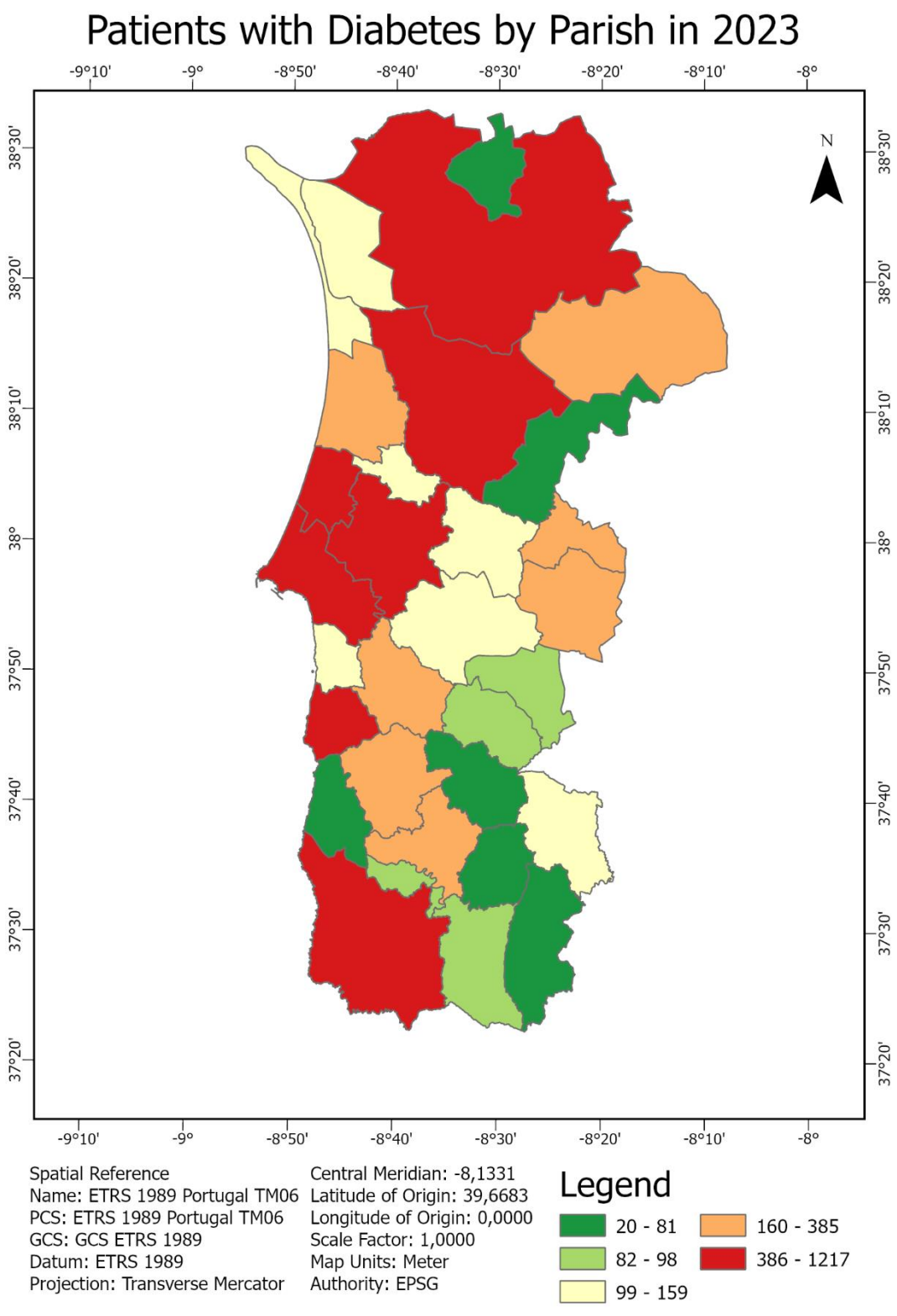
Percentage of Population that is Married by Parish in 2023

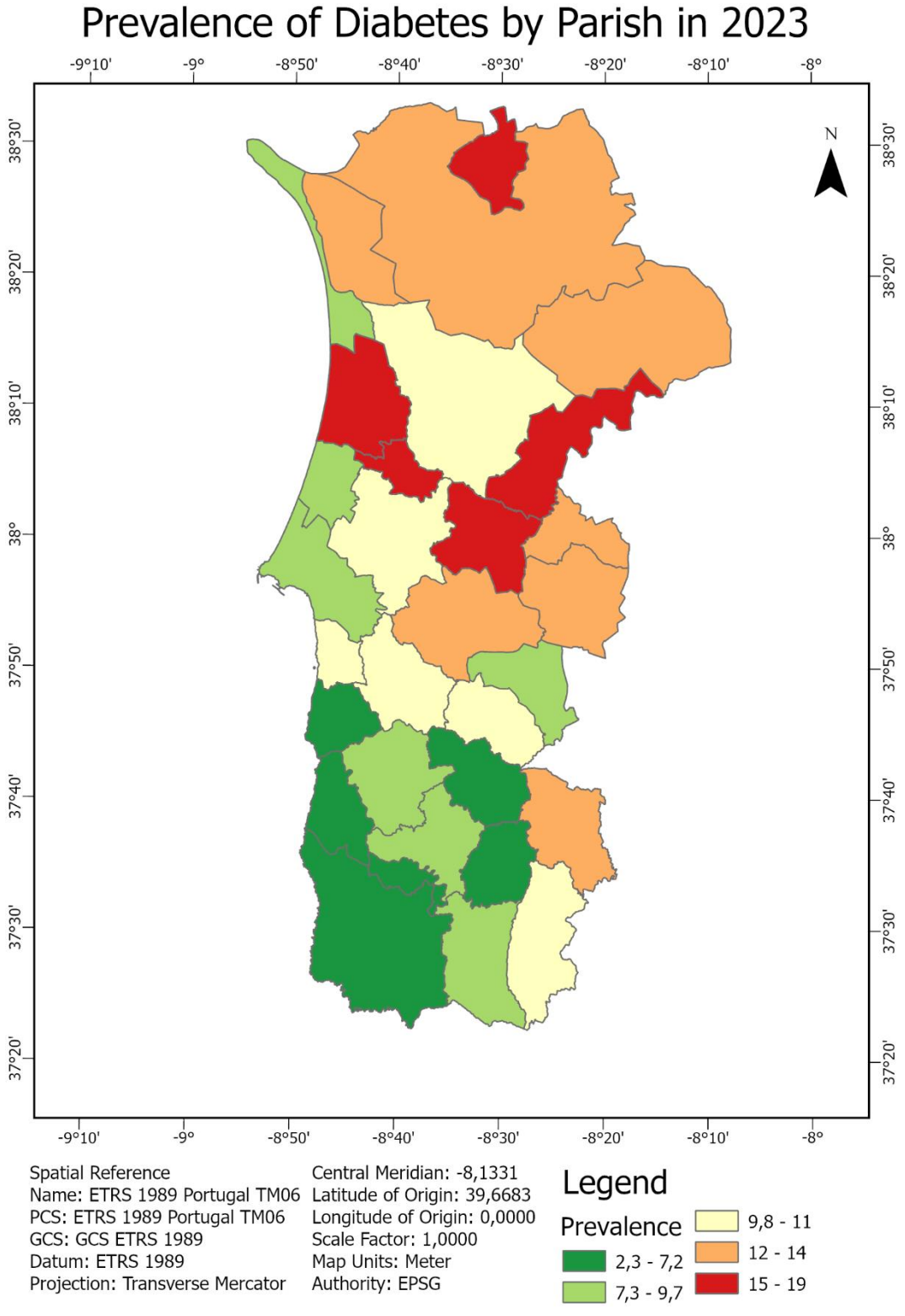


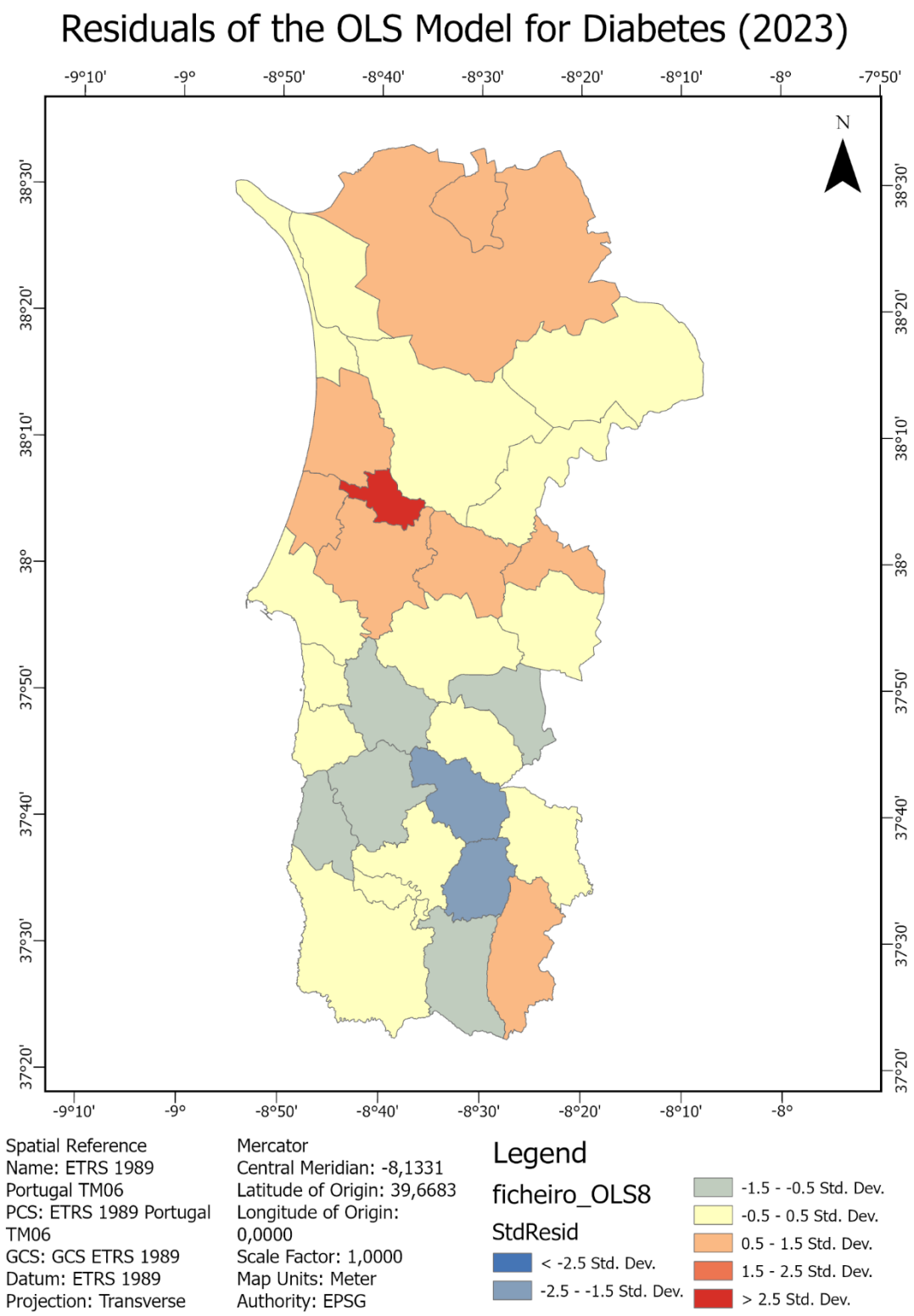
Percentage of Population that is Divorced by Parish in 2023

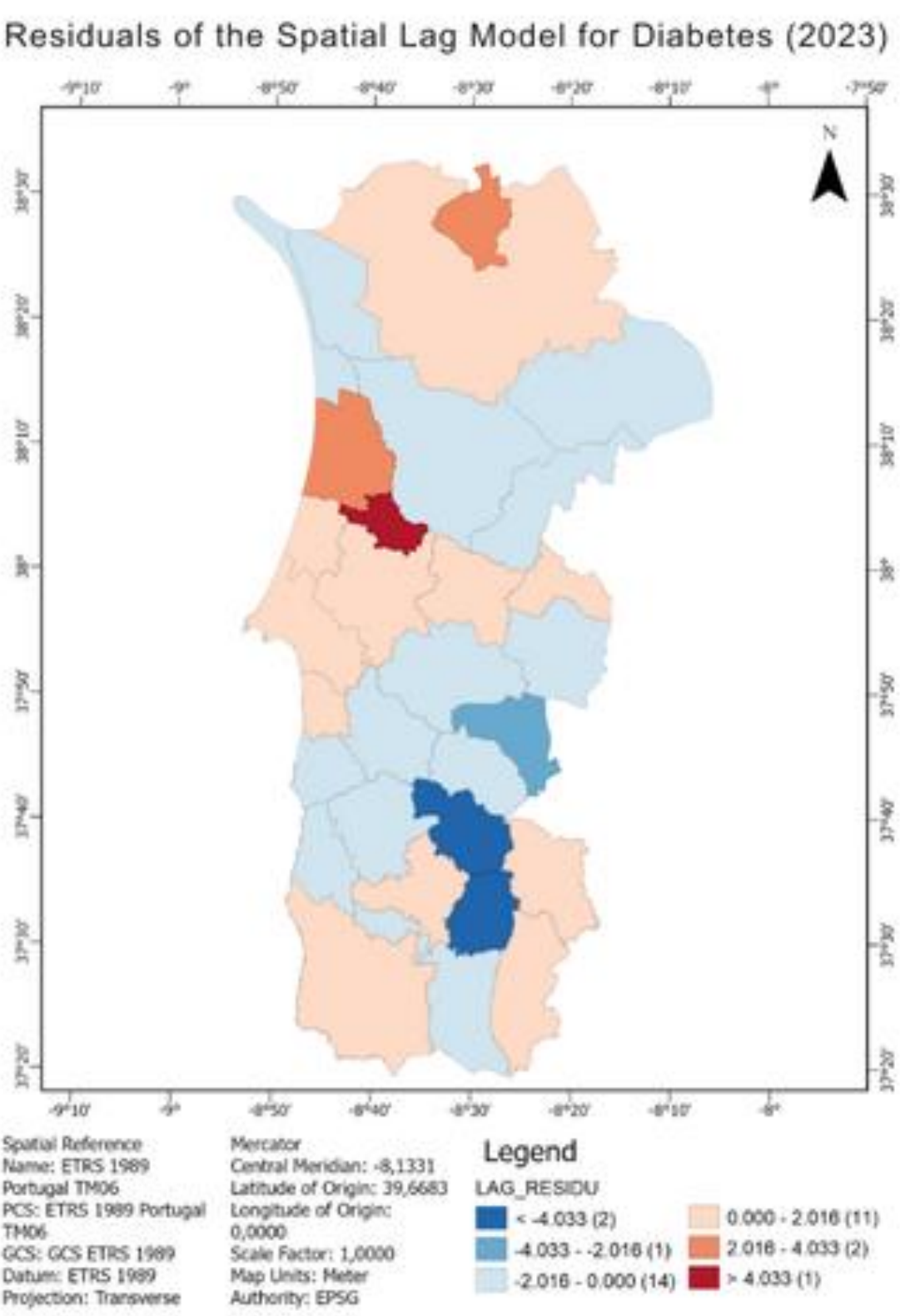



Per Capita Purchasing Power by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
2023 Purchasing Power: Per Capita
8700 - 10000
11000 - 12000
13000 - 14000
15000 - 16000
17000 - 18000

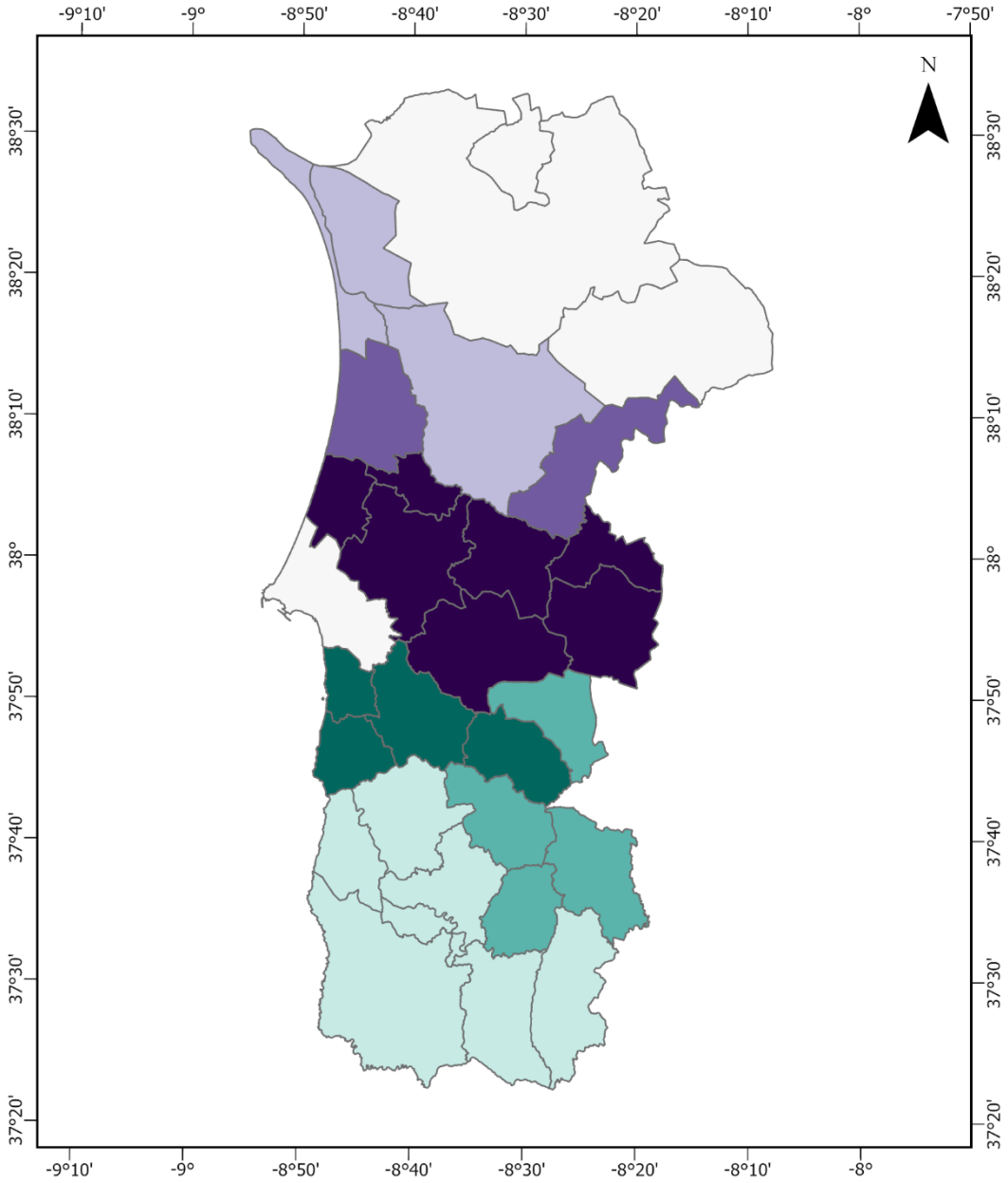

Patients with Diabetes by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
20 - 81
82 - 98
99 - 159
160 - 385
386 - 1217


Prevalence of Diabetes by Parish in 2023
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
Prevalence
2,3 - 7,2
7,3 - 9,7
9,8 - 11
12 - 14
15 - 19


Residuals of the OLS Model for Diabetes (2023)
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
-7°50'
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
ficheiro_OLS8
StdResid
< -2.5 Std. Dev.
-2.5 - -1.5 Std. Dev.
-1.5 - -0.5 Std. Dev.
-0.5 - 0.5 Std. Dev.
0.5 - 1.5 Std. Dev.
1.5 - 2.5 Std. Dev.
> 2.5 Std. Dev.


Residuals of the Spatial Lag Model for Diabetes (2023)
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
LAG_RESIDU
< -4.033 (2)
-4.033 - -2.016 (1)
-2.016 - 0.000 (14)
0.000 - 2.016 (11)
2.016 - 4.033 (2)
> 4.033 (1)

Spatial Distribution of GWR Coefficients for PER1EDU



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

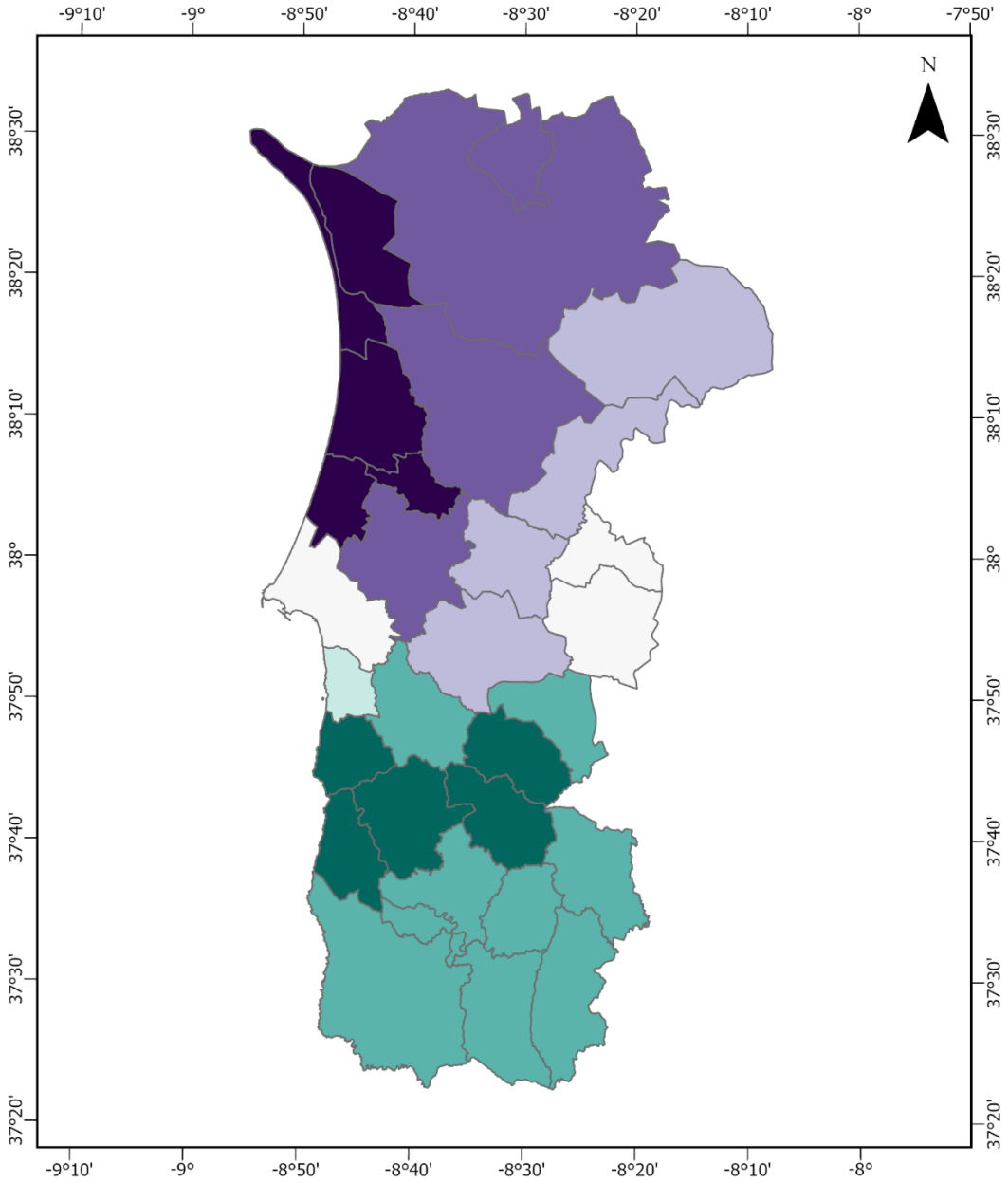
Legend

GWR

Coefficient (Scaled PER1EDU)

- 0,32 - 0,37
- 0,38 - 0,45
- 0,46 - 0,52
- 0,53 - 0,57
- 0,58 - 0,59
- 0,60 - 0,60
- 0,61 - 0,65

Spatial Distribution of GWR Coefficients for PERDIVORC



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

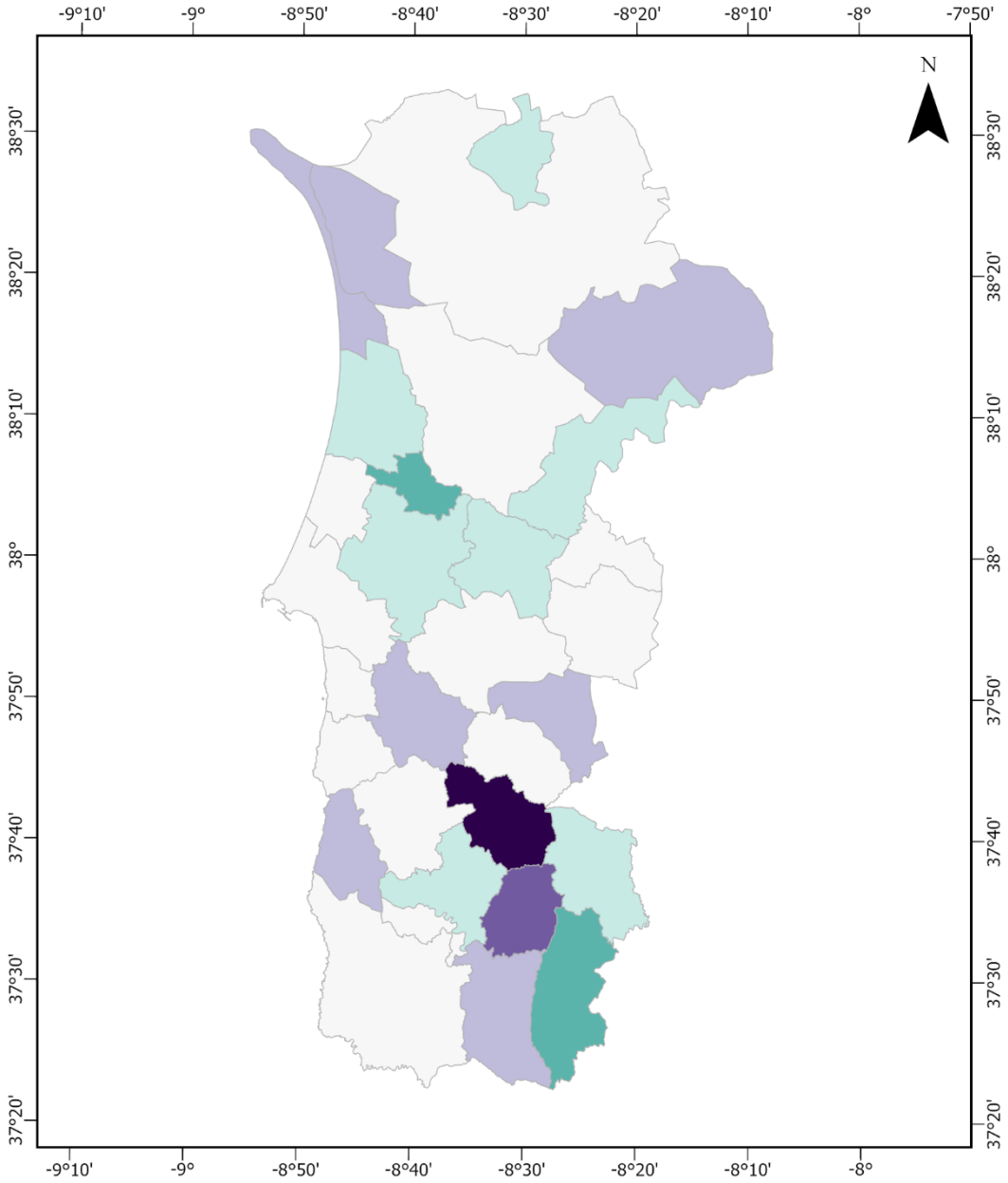
Legend

GWR

Coefficient (Scaled PERDIVORC)

- -0,35 - -0,22
- -0,21 - -0,12
- -0,11 - 0,031
- 0,032 - 0,18
- 0,19 - 0,32
- 0,33 - 0,39
- 0,40 - 0,42

Residuals of the GWR Model for Diabetes (2023)



Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse

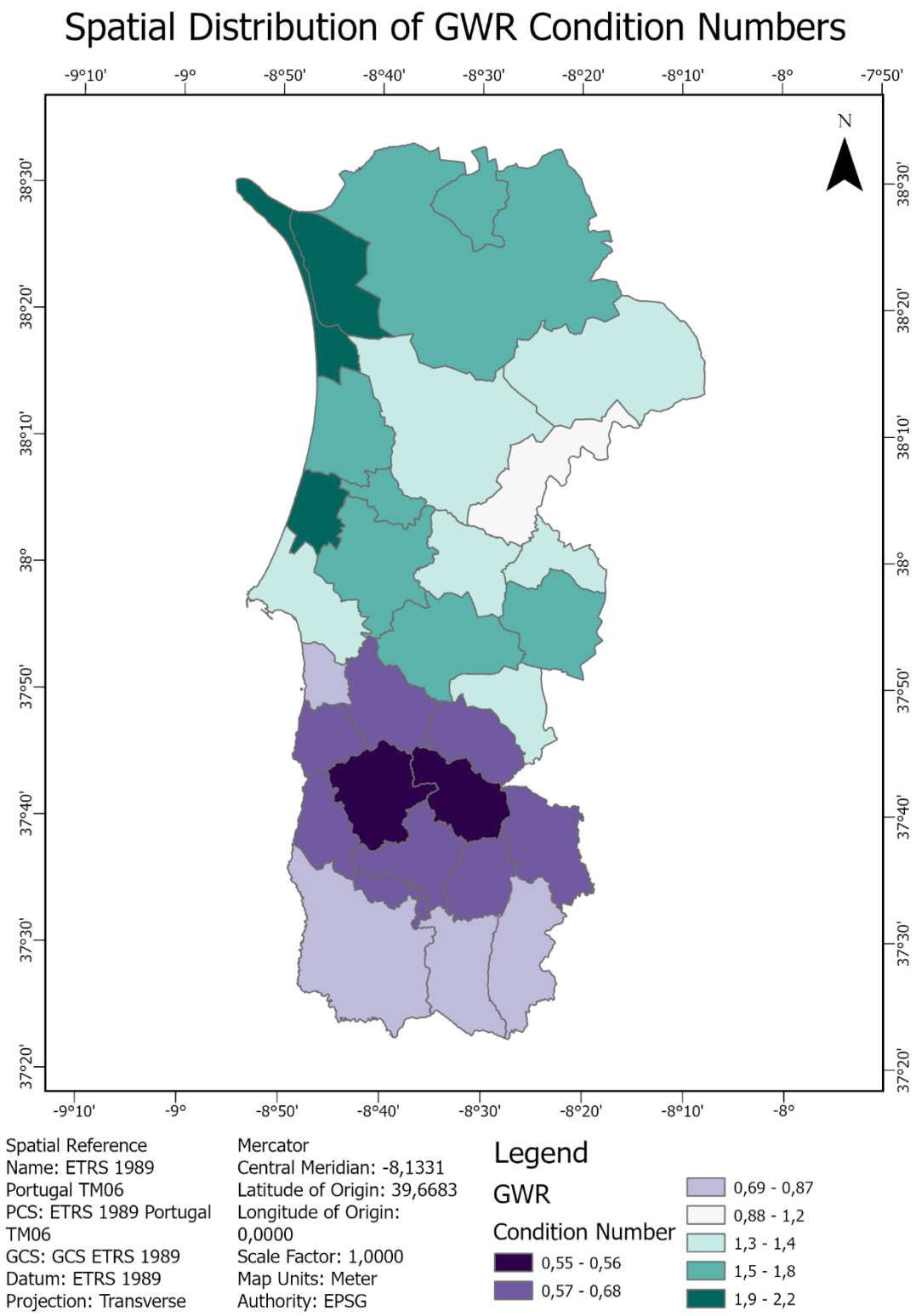
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG

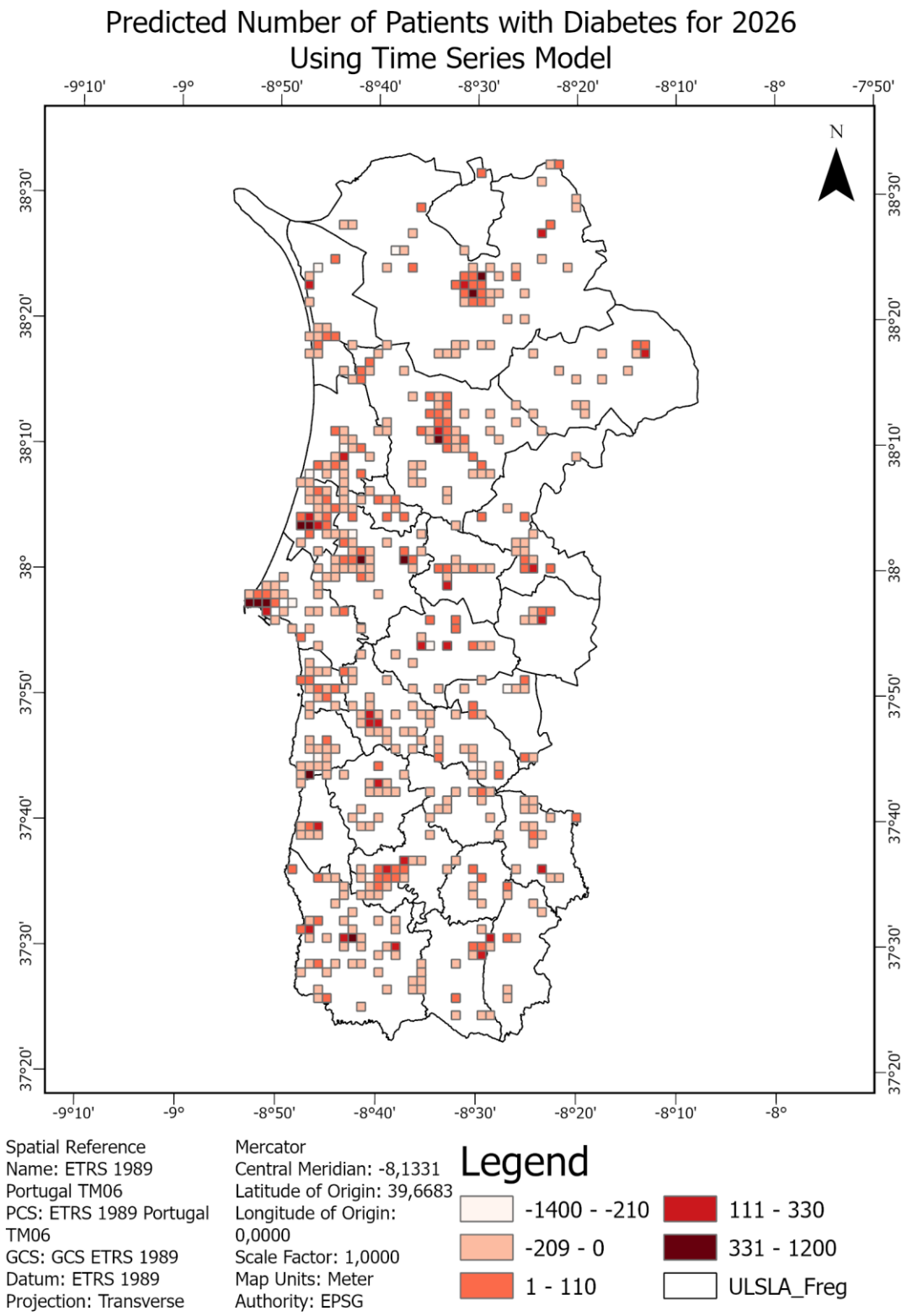
Legend

GWR

Standardized Residual

- < -2,5 Std. Dev.
- -2,5 - -1,5 Std. Dev.
- -1,5 - -0,5 Std. Dev.
- -0,5 - 0,5 Std. Dev.
- 0,5 - 1,5 Std. Dev.
- 1,5 - 2,5 Std. Dev.
- > 2,5 Std. Dev.


Spatial Distribution of GWR Condition Numbers
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
-7°50'
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
GWR
Condition Number
0,55 - 0,56
0,57 - 0,68
0,69 - 0,87
0,88 - 1,2
1,3 - 1,4
1,5 - 1,8
1,9 - 2,2


Predicted Number of Patients with Diabetes for 2026
Using Time Series Model
-9°10'
-9°
-8°50'
-8°40'
-8°30'
-8°20'
-8°10'
-8°
-7°50'
38°30'
38°20'
38°10'
38°
37°50'
37°40'
37°30'
37°20'
N
Spatial Reference
Name: ETRS 1989 Portugal TM06
PCS: ETRS 1989 Portugal TM06
GCS: GCS ETRS 1989
Datum: ETRS 1989
Projection: Transverse
Mercator
Central Meridian: -8,1331
Latitude of Origin: 39,6683
Longitude of Origin: 0,0000
Scale Factor: 1,0000
Map Units: Meter
Authority: EPSG
Legend
-1400 - -210
-209 - 0
1 - 110
111 - 330
331 - 1200
ULSLA_Freg