

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E INVESTIGAÇÃO OPERACIONAL



Retrato de Risco dos Municípios Portugueses

Inês Filipa Gorjão Oliveira

Mestrado em Matemática Aplicada à Economia e Gestão

Trabalho de projeto orientado por:
Prof.^a Doutora Luísa Canto e Castro Loura

2020

AGRADECIMENTOS

A realização desta dissertação de mestrado contou com importantes apoios sem os quais o desenvolvimento da mesma teria sido impossível.

Em primeiro lugar, gostaria de agradecer à orientadora Professora Doutora Luísa Canto e Castro Lou-ra, que apesar da situação pandémica em que nos encontramos sempre se demonstrou disponível para esclarecer toda e qualquer dúvida dentro dos meios que nos eram possíveis e que sempre me acompanhou com um grande empenho e dedicação ao longo deste percurso.

Agradecer também às coordenadoras deste mestrado que me acompanharam ao longo destes últimos dois anos, a Professora Doutora Teresa Alpuim e a Professora Doutora Raquel João Espinha Fonseca que desempenharam um papel importantíssimo no que diz respeito ao meu desenvolvimento quer a nível pessoal quer a nível profissional ao longo deste percurso.

Também os meus amigos e colegas apresentaram um aspeto de relevo, motivando-me sempre para que a realização deste feito fosse hoje possível.

Seria ainda impossível alcançar a realização deste mestrado sem o apoio dos meus familiares, principalmente do meu pai, mãe e irmã, que sempre me apoiaram incondicionalmente e contribuíram para que tudo fosse possível.

A todos, o meu sincero e profundo Obrigado!

RESUMO

O presente estudo tem como objetivo traçar um perfil dos municípios portugueses, de forma a elaborar cenários prospetivos e identificar as áreas que, em cada município, poderão vir a configurar uma situação de risco e a necessitar de intervenção.

Para tal, primeiramente foi necessário estruturar, sistematizar e selecionar diversos indicadores referentes aos 308 municípios, para que, desta forma fosse possível analisar o município num contexto financeiro, demográfico, económico e social avaliando ainda as mudanças que ocorreram em alguns dos indicadores entre 2013 e 2018. Sendo que, os dados utilizados nesta análise foram recolhidos da base de dados estatísticos disponibilizada pela PORDATA. Após recolha da respetiva informação, foi traçado o perfil dos municípios através da utilização de duas metodologias de análise estatística multivariada, a Análise Fatorial e Análise de Clusters, mais especificamente, através da implementação do método *Two Step Cluster*. Numa fase inicial, a Análise Fatorial permitiu condensar a informação contida nas várias variáveis originais com uma perda mínima de informação, dando origem a novas variáveis, os *factor loadings*, que foram as variáveis utilizadas no método *Two Step Cluster*. O *Two Step Cluster*, por sua vez, permitiu traçar o perfil dos municípios, através da organização dos municípios em grupos “homogéneos”.

Por fim, foram obtidos 6 *clusters*, estes clusters permitiram avaliar o desempenho dos diversos municípios, permitindo perceber quais as características que apresentam uma maior capacidade de influenciar a competitividade de um município. Foi ainda possível, analisar possíveis situações de risco inerentes aos mesmos.

PALAVRAS-CHAVE

Municípios portugueses, Indicadores, Risco, Análise de *clusters*, Análise fatorial.

ABSTRACT

The purpose of the current study is to trace a profile of the Portuguese municipalities, in order to elaborate prospective scenarios and identify areas that, in each municipality, can become a risk situation and might need intervention.

For this purpose, firstly there was a need to structure, systematize and select the various indicators referring to the 308 municipalities, to make possible the analysis of the municipality in a financial, demographic, economic and social context evaluating also the changes that occur in many of the indicators that were used for this analysis between 2013 and 2018. The data used in this analysis was extracted from the statistical databased provided by PORDATA. After the extraction of the respective information, a profile of the municipalities has been traced, through the implementation of two methodologies of multivariate statistical analysis, the Factor Analysis and a Cluster Analysis, more specifically, through the use of the method Two Step Cluster. In an initial phase, the Factor Analysis allowed the condensation of the information contained in the original variables with a minimal loss of information, this analysis provided new variables, the factor loadings, and these were the variables that were used in the Two Step Cluster. The Two Step Cluster allowed the tracing of a profile for the different municipalities, through the organization of these municipalities in “homogeneous” groups.

In conclusion, 6 clusters were obtained, these clusters allowed to evaluate the performance of the municipalities, allowing as well the perception of which characteristics have a greater impact in the competitiveness of a municipality. It was also possible to analyse possible risk situations inherent to them.

KEYWORDS

Portuguese municipalities, Indicators, Risk, Cluster Analysis, Factor Analysis.

ÍNDICE

INTRODUÇÃO	1
METODOLOGIA	3
2.1 Análise Fatorial	3
2.2 Análise de Clusters	7
2.2.1 Método <i>Two Step Cluster</i>	8
APRESENTAÇÃO, ANÁLISE E DISCUSSÃO DOS RESULTADOS	10
3.1 Resultados obtidos da Análise Fatorial	10
3.1.1 Teste de KMO e Bartlett	10
3.1.2 Comunalidades	10
3.1.3 Variância total explicada	12
3.1.4 Matriz de componentes	13
3.2 Normalização das variáveis	17
3.3 Resultados obtidos do <i>Two Step Cluster</i>	18
3.3.1 Descrição dos Clusters	21
3.3.2 <i>Cluster 1</i>	22
3.3.3 <i>Cluster 2</i>	25
3.3.4 <i>Cluster 3</i>	27
3.3.5 <i>Cluster 4</i>	30
3.3.6 <i>Cluster 5</i>	34
3.3.7 <i>Cluster 6</i>	36
CONCLUSÃO	40
BIBLIOGRAFIA	44
ANEXOS	50

LISTA DE TABELAS

Tabela 2.1 - Valores KMO	5
Tabela 3.1 – Resultado obtido no Teste de KMO e Bartlett em SPSS	10
Tabela 3.2 – Tabela das Comunalidades obtida em SPSS	11
Tabela 3.3 – Tabela da Variância total explicada obtida em SPSS	12
Tabela 3.4 – Tabela da Matriz de componentes obtida em SPSS	13
Tabela 3.5 – Tabela da Matriz de componentes rotacionada obtida em SPSS.....	14
Tabela 3.6 - Valores mais positivos e negativos nos fatores	17
Tabela 3.7 - Interpretação dos valores do coeficiente de silhueta	19
Tabela 3.8 - Tamanho do maior e menor cluster obtido em SPSS	20
Tabela 3.9 - Município mais característico do Cluster 1	22
Tabela 3.10 - Municípios característicos do Cluster 2	26
Tabela 3.11 - Município mais característico do Cluster 3	28
Tabela 3.12 - Município mais característico do cluster 4.....	31
Tabela 3.13 - Município mais característico do Cluster 5	34
Tabela 3.14 - Município mais característico do Cluster 6.....	37
Tabela 4.1 - Os 10 Municípios onde o valor médio de avaliação bancária dos alojamentos é mais elevado	42
Tabela 4.2 - Refugiados em Portugal, entre 2013 e 2018.....	42
Tabela 1 - Municípios pertencentes ao Cluster 1	50
Tabela 2 - Municípios pertencentes ao Cluster 3	50
Tabela 3 - Municípios pertencentes ao Cluster 4	51
Tabela 4 - Municípios pertencentes ao Cluster 5	51
Tabela 5 - Municípios pertencentes ao Cluster 6	51

LISTA DE FIGURAS

Figura 3.1 - Resumo do modelo e Qualidade do cluster obtido em SPSS	18
Figura 3.2 - Tamanhos de cluster obtido em SPSS	20
Figura 3.3 - Distribuição geográfica dos municípios elaborado no Flourish	20
Figura 3.4 - Exemplo de caixa de bigodes/boxplot	21
Figura 3.5 - Cluster 1 obtido em SPSS.....	22
Figura 3.6 - Cluster 2 obtido em SPSS.....	25
Figura 3.7 - Cluster 3 obtido em SPSS.....	28
Figura 3.8 - Cluster 4 obtido em SPSS.....	31
Figura 3.9 - Cluster 5 obtido em SPSS.....	34
Figura 3.10 - Cluster 6 obtido em SPSS.....	36

LISTA DE ABREVIATURAS

AIC – Critério de Informação de Akaike

BIC - Critério Bayesiano de Schwarz

IMD – Institute for Management Development

ONU – Organização das Nações Unidas

PDR – Programa de Desenvolvimento Rural

PIB – Produto Interno Bruto

SI2E - Sistema de Incentivos ao Empreendedorismo e ao Emprego

CAPÍTULO 1

INTRODUÇÃO

A competitividade regional corresponde à capacidade que determinada região apresenta no sentido de conseguir conciliar condições de sucesso empresarial com a capacidade de proporcionar uma boa qualidade de vida à população (Fagerberg *et al.*, 2004).

Segundo o relatório do *IMD's World Competitiveness*¹ realizado este ano, existem diversos fatores que demonstraram apresentar uma importância extrema para que um país seja considerado competitivo. Foram eles: uma forte performance económica, relacionada com um comércio e investimento internacional robusto, um forte desempenho relativamente ao sistema educativo e relativamente ao desenvolvimento tecnológico, como: telecomunicações, velocidade de internet e exportações de alta tecnologia. Outro fator de relevo está relacionado com a existência de um bom sistema de saúde.

De acordo com o artigo da Porto *Business School*, “*Portugal alcança o 37º lugar no ranking da competitividade*”, os principais fatores que tornam a economia nacional mais atrativa dizem respeito à mão de obra qualificada, custo de oportunidade e estabilidade das infraestruturas. Ao longo dos últimos anos, de acordo com os dados obtidos no ranking “*IMD's World Competitiveness*”, Portugal tem apresentado uma economia mais sólida. Tal é principalmente devido a receitas turísticas, exportação de produtos e à elevada inflação dos preços ao consumidor. A estrutura social e legislação empresarial que Portugal apresenta é também motivo de destaque, assim como o nível de infraestruturas educativas, de saúde e ambiente. Relativo ainda à sua performance económica, Portugal destaca-se também na categoria de preços e no comércio externo.

No entanto, existem fatores que conferem uma situação de risco e que como tal deveriam ser mitigados de forma a melhorar o desempenho de Portugal e aumentar a sua competitividade, são eles: competência do Estado, eficácia do ambiente legal e regime fiscal. Pelo que, as metas definidas para 2020 consistem em conseguir garantir um crescimento estável e alto do PIB num contexto de restrições económicas internacionais, e na adoção de um quadro fiscal favorável às empresas e ao investimento. Algo que é evidenciado também por Patrícia Teixeira Lopes, *Associate Dean* da Porto *Business School*, no artigo “*A competitividade da Economia. Mind the Gap!*” onde afirma que “Para aumentar a competitividade da nossa economia será necessário estabilizar as políticas públicas, reduzir e estabilizar a carga fiscal, reduzir a burocracia e melhorar a eficiência do sistema legal”.

Com o objetivo de compreender melhor este desempenho, na sua total envolvente, foi feita uma análise relativa aos 308 municípios de Portugal para que, desta forma, fosse possível identificar os municípios que de facto apresentam um melhor desempenho tornando Portugal mais competitivo, assim como aqueles municípios que podem apresentar fatores de risco pondo em causa esta mesma competitividade.

A metodologia utilizada no desenvolvimento desta tese é semelhante àquela que foi abordada no artigo de Soares *et al.* (2003), publicado no *European Journal of Operational Research*. Nesse estudo, primeiramente, foi feita uma seleção das variáveis que os autores consideraram ser aquelas que iriam permitir analisar de uma forma geral o perfil dos diversos municípios. Posteriormente, esses dados

¹ Relatório elaborado pelo International Institute for Management Development (IMD), Lausanne, Suíça

foram submetidos a duas técnicas estatísticas: a Análise Fatorial seguida de uma Análise de Clusters, com o objetivo de condensar e sintetizar a informação permitindo criar perfis associados aos diversos municípios. O contributo adicional do presente estudo, prende-se principalmente com o facto de ter em consideração indicadores que traduzem a evolução que ocorreu ao longo dos últimos anos, mais especificamente, entre 2013 e 2018, nas variáveis que foram tidas em consideração para os 308 municípios. Desta forma, é possível fazer uma análise mais completa permitindo visualizar e perceber efetivamente a mudança que ocorreu nos valores de determinados indicadores ao longo dos anos, podendo estas alterações vir a conferir ou não uma situação de risco ao município. Desta forma é também possível retirar conclusões mais consistentes e desenhar formas de atuação mais eficazes.

Para tal, tal como mencionado anteriormente, primeiramente foi feita uma seleção de dados, sendo que os dados utilizados nesta análise foram recolhidos da base de dados estatísticos disponibilizada pela PORDATA. Inicialmente foram consideradas diversas variáveis, de forma a conseguir avaliar as demais características dos 308 municípios. Pretendia-se, desta forma, obter uma imagem global de cada município quer a nível demográfico, quer financeiro, económico, social e, também, quanto a tendências evolutivas nos últimos cinco anos.

Posteriormente, procedeu-se ao tratamento dos dados em questão, tendo-se, para isso recorrido à Análise Fatorial exploratória. Esta análise tem como objetivo condensar a informação contida nas variáveis originais com uma perda mínima de informação. Após a utilização deste método as variáveis que foram efetivamente consideradas para o presente estudo foram: Receitas municipais por mil habitantes; Empresas não financeiras por mil habitantes; Taxa de variação anual do número de empresas não financeiras; Ganho médio dos trabalhadores por conta de outrem; Percentagem de empresas criadas no setor da Indústria, Construção e Energia; Taxa média de variação anual da densidade populacional; Índice de envelhecimento; Taxa média de variação anual do Índice de envelhecimento; Percentagem de população em idade ativa; Percentagem de população estrangeira; Número de anos de escolaridade das mulheres que foram mães; Número médio de anos de escolaridade dos empregadores; Número médio de anos de escolaridade dos trabalhadores por conta de outrem; Taxa média de variação anual do número de anos de escolaridade dos trabalhadores por conta de outrem; Alojamentos turísticos por área do município e Alojamentos familiares clássicos por habitante.

Com o objetivo de traçar um perfil destes municípios foi aplicado posteriormente o método *Two Step Cluster*. Este método permitiu agrupar os diversos municípios com base na sua similaridade, através da organização de grupos “homogéneos”. Por fim, foi feita a análise relativa ao risco de cada município, permitindo a identificação das áreas com perspetivas mais positivas e negativas, de acordo com a informação apurada ao longo deste estudo.

CAPÍTULO 2

METODOLOGIA

Para o desenvolvimento desta investigação recorreu-se aos dois tipos de métodos: o método quantitativo e o método qualitativo.

O método qualitativo está presente na escolha das variáveis que foram consideradas para o desenvolvimento deste estudo. Esta escolha foi feita tendo por base os indicadores que considerámos serem promotores do sucesso dos municípios, procurando recolher informação que permitisse ter uma visão global das diversas áreas que permitem avaliar o desempenho de um município.

Quantitativo pois para o desenvolvimento do presente estudo recorreu-se aos seguintes métodos estatísticos: Análise Fatorial Exploratória e *Two Step Cluster*.

O público-alvo deste estudo corresponde aos 308 municípios de Portugal, sendo 278 relativos a Portugal Continental, 19 da Região Autónoma dos Açores e 11 da Região Autónoma da Madeira.

Este estudo irá basear-se numa análise estatística multivariada que é aquela que remete para os métodos estatísticos que analisam simultaneamente diversos indicadores/variáveis relativamente a determinado indivíduo/objeto que se encontra sob investigação.

2.1 Análise Fatorial

A análise fatorial é uma técnica de análise multivariada cujo objetivo é definir, num estudo que contenha diversas variáveis, conjuntos de variáveis altamente correlacionados, denominados de fatores. Estes “fatores” têm como objetivo reduzir as diversas variáveis a um conjunto de menor dimensão com uma perda mínima de informação. Desta forma, sempre que existam variáveis que se encontrem fortemente relacionadas, as mesmas devem ser agrupadas num determinado grupo, de modo a que variáveis que estejam em diferentes grupos apresentem uma fraca correlação.

Na análise fatorial as variáveis y_1, y_2, \dots, y_m , são expressas como combinações lineares de umas outras variáveis F_1, F_2, \dots, F_p ($p < m$) denominadas de fatores. Os fatores são variáveis latentes que vão dar origem a outras variáveis. Assim como ocorre nas variáveis originais, também os fatores variam de indivíduo para indivíduo, porém estes fatores não são mensuráveis nem observáveis. Na análise fatorial, procura-se explicar as covariâncias e/ou correlações existentes entre as variáveis. O modelo pode ser escrito, algebricamente da seguinte forma:

$$Y_i = a_{i1} \times F_1 + a_{i2} \times F_2 + \dots + a_{ip} \times F_p + e_i, \quad i = 1, \dots, m \quad (2.1)$$

Aqui a_{is} é o *factor loading* do fator s para a variável i e e_i corresponde à parte da variável Y_i que não é explicada pelos fatores.

Existem dois tipos de análise fatorial: a Análise Fatorial Confirmatória e a Análise Fatorial Exploratória. A Análise Fatorial Confirmatória parte do pressuposto que já existe uma teoria sobre quais as variáveis que têm maior poder explicativo em cada fator, permitindo confirmar o grau de ajuste dos dados observados à teoria que foi previamente feita. A Análise Fatorial Exploratória vai

analisar o padrão de correlações existentes entre as variáveis, utilizando posteriormente esses padrões de correlações para agrupar as variáveis em fatores.

Para o desenvolvimento deste estudo recorreu-se à Análise Fatorial Exploratória, tendo, primeiramente, sido consideradas as seguintes variáveis: Receitas municipais por mil habitantes; Taxa média de variação anual das receitas; Saldo financeiro por mil habitantes; Empresas não financeiras por mil habitantes; Taxa de variação anual do número de empresas não financeiras; Desempregados registados nos centros de emprego; Taxa de variação anual de desempregados registados nos centros de emprego; Ganho médio dos trabalhadores por conta de outrem; Percentagem de empresas criadas no setor Indústria, Construção e Energia; Densidade populacional; Taxa média de variação anual da densidade populacional; Índice de envelhecimento; Taxa média de variação anual do Índice de envelhecimento; Percentagem de população em idade ativa; Percentagem de população estrangeira; Número de alunos no ensino secundário; Número de anos de escolaridade das mulheres que foram mães; Número médio de anos de escolaridade dos trabalhadores por conta de outrem; Taxa média de variação anual do número de anos de escolaridade dos trabalhadores por contra de outrem; Número médio de anos de escolaridade dos empregadores; Alojamentos turísticos por área do município; Valor médio de avaliação bancária dos alojamentos; Alojamentos familiares clássicos por habitante; Pessoal ao serviço nos hospitais.

No entanto, após realizar a Análise Fatorial concluiu-se que algumas variáveis apresentavam um peso diminuto na tabela das comunalidades, traduzindo-se numa menor contribuição para a explicação da variabilidade total dos dados, pelo que, as variáveis que acabaram efetivamente por ser consideradas foram: Receitas municipais por mil habitantes; Empresas não financeiras por mil habitantes; Taxa de variação anual do número de empresas não financeiras; Ganho médio dos trabalhadores por conta de outrem; Percentagem de empresas criadas no setor da Indústria, Construção e Energia; Taxa média de variação anual da densidade populacional; Índice de envelhecimento; Taxa média de variação anual do Índice de envelhecimento; Percentagem de população em idade ativa; Percentagem de população estrangeira; Número de anos de escolaridade das mulheres que foram mães; Número médio de anos de escolaridade dos empregadores; Número médio de anos de escolaridade dos trabalhadores por conta de outrem; Taxa média de variação anual do número de anos de escolaridade dos trabalhadores; Alojamentos turísticos por área do município e Alojamentos familiares clássicos por habitante.

Após ter sido feita a seleção das variáveis procedeu-se então à Análise Fatorial Exploratória novamente, utilizando o método das componentes principais. Posteriormente, procedeu-se à análise dos resultados obtidos sendo necessário, primeiramente, a validação do pressuposto de normalidade das distribuições das variáveis.

Após aplicação da análise fatorial há ainda que observar a matriz de correlações pois, para que esta metodologia seja útil é importante que exista um grau de correlação entre as variáveis, duas a duas, suficientemente elevado. O objetivo é que existam algumas componentes que consigam explicar grande parte da variabilidade dos dados. Caso tal não se verifique, a aplicação desta análise não faz sentido.

Posto isto, irá ser então analisado o teste relativo à validade da análise fatorial, o teste de esfericidade de Bartlett, em que vão ser consideradas duas hipóteses:

H₀: Considerar matriz de correlações como matriz de identidade (se sig for $\geq 0,05$, não se rejeita H₀, e se assim for não se deve proceder à análise fatorial)

H₁: Matriz de correlações é de similaridade e não de identidade (se sig $< 0,05$, rejeita-se H₀, e a análise fatorial irá contribuir de forma útil para a redução da dimensionalidade)

É necessário ainda avaliar a Medida de KMO, que serve para avaliar a intensidade da análise fatorial e só é aplicável se a hipótese nula tiver sido rejeitada no teste de Bartlett. O teste KMO (*Kaiser-Meyer-Olkin*) compara as correlações simples com as correlações parciais observadas entre as variáveis, (Marôco., 2010b) e (Pestana & Gageiro., 2000) e é calculada da seguinte forma:

$$KMO = \frac{\sum \sum_{j \neq k} r_{jk}^2}{\sum \sum_{j \neq k} r_{jk}^2 + \sum \sum_{j \neq k} q_{jk}^2} \quad (2.2)$$

Onde r_{jk}^2 corresponde ao quadrado dos elementos da matriz de correlação original fora da diagonal e q_{jk}^2 representa o quadrado das correlações parciais entre as variáveis.

Os valores do índice KMO que vão indicar se a Análise Fatorial é apropriada ou não variam consoante o autor, tendo sido considerada a seguinte interpretação:

Tabela 2.1 - Valores KMO

KMO	Grau de ajuste à Análise Fatorial
0,9 a 1	Muito boa
0,8 a 0,9	Boa
0,7 a 0,8	Média
0,6 a 0,7	Razoável
0,6 a 0,6	Má
< que 0,5	Inaceitável

Fonte: adaptado de Pestana & Gageiro (2000)

Posteriormente é necessário escolher o método de extração das componentes, podendo ser utilizado o método das componentes principais, da máxima verosimilhança, dos mínimos quadrados não ponderados, dos mínimos quadrados generalizados, da factorização do eixo principal, da factorização alfa ou o método de factorização de imagem.

O método de extração utilizado para o desenvolvimento deste estudo foi o método das componentes principais. Esta técnica de análise procura combinações lineares das variáveis aleatórias originais, com variância máxima e não correlacionadas entre si. As novas variáveis assim obtidas, denominam-se de componentes principais. As componentes principais por norma são obtidas a partir de uma matriz de covariâncias, porém podem também ser obtidas através de uma matriz de correlações.

É possível mostrar que as componentes principais coincidem com os vetores próprios da matriz de covariâncias. Mais precisamente, o método das componentes principais consiste no cálculo dos valores próprios e vetores próprios da matriz de variâncias-covariâncias (matriz de coeficientes de correlação)

entre variáveis padronizadas. Os vetores próprios são calculados de modo a definirem vetores com tamanho unitário. Isso é conseguido através da multiplicação de cada elemento do vetor próprio normalizado pela raiz quadrada do correspondente valor próprio. O resultado é um fator, vetor constituído por cargas fatoriais. A matriz de variâncias e covariâncias observadas, [S2], é igual ao produto da matriz de cargas fatoriais (*factor loading*) [AR], de dimensões m x p, multiplicada pelo seu transposto, mais uma matriz diagonal de variâncias únicas, [var Ejj].

$$[S2] = [AR][AR]' + [varEjj] \quad (2.3)$$

Seguidamente há que escolher o número de fatores que serão usados na Análise Fatorial. Existem vários critérios que ajudam a determinar o número de fatores a reter. São eles o critério de *Kaiser*, o critério do *Scree-plot* e a percentagem da variância explicada. O critério utilizado neste estudo foi o critério de *Kaiser* e segundo este critério, o número de fatores deve ser igual ao número de valores próprios maiores ou iguais que a média aritmética dos p valores próprios. Quando usada a matriz de correlações esta média corresponde a 1. Devem ainda ser interpretados os fatores cuja percentagem cumulativa de variância seja superior a 50%, ou seja, os fatores devem conseguir explicar mais de 50% da variação dos dados (Marôco., 2010b).

Posteriormente há que recorrer à rotação das matrizes de correlações, de forma a maximizar a variância das cargas fatoriais. Esta rotação vai procurar uma forma corrigida de as variáveis se relacionarem dentro de determinado fator sem que se relacionem com outros. Permite desta forma que cada variável tenha carga fatorial elevada no menor número possível de fatores.

$$s_j^2 = \frac{p \sum_{j=1}^m \left(\frac{a_{jp}^2}{h_j^2} \right)^2 - \left(\frac{\sum_{j=1}^m a_j^2}{h_j^2} \right)^2}{p^2} \quad (2.4)$$

Na expressão acima, s_j^2 corresponde à variância das cargas fatoriais, p é o número de fatores, m corresponde ao número de variáveis originais, a_{jp} à carga da variável j no fator p e h_j^2 corresponde à comunalidade da j'ésima variável.

A quantidade que se deseja maximizar com a rotação é:

$$V = \sum_{j=1}^p s_j^2 \quad (2.5)$$

Existem dois tipos de rotações, a ortogonal e a oblíqua. As rotações oblíquas permitem que os fatores estejam correlacionados entre si. Existem vários métodos de rotação oblíqua tais como: *Oblimin*, *Quartimin*, *Promax*, entre outros. No geral, todos eles têm tendência a apresentar resultados semelhantes (Costello & Osborne., 2005).

As rotações ortogonais assumem que os fatores extraídos são independentes uns dos outros, ou seja, não apresentam correlações entre si. Existem três métodos de rotação ortogonais que podem ser aplicados: o método *Varimax*, *Quartimax*, e o *Equimax*. O mais utilizado é o *Varimax* tendo sido também aquele que foi utilizado no corrente estudo. O método *Varimax* tem como objetivo obter uma

estrutura fatorial na qual uma e apenas uma das variáveis originais esteja fortemente associada com um único fator e pouco associada com os restantes. Desta forma, este método procura maximizar a variância dos quadrados dos *loadings* em cada fator, na matriz dos fatores (S).

Ambos os métodos de rotação tornam os *factor loadings* elevados mais elevados, os baixos mais baixos, fazendo desaparecer os *loadings* com valores intermédios. Cada *factor loading* assume valores no intervalo [-1,1], sendo de realçar que, por norma, são considerados significativos os *loadings* maiores ou iguais a 0,5 (Sharma., 1996). Do ponto de vista interpretativo, *factor loadings* com valores próximos de -1 e 1 indicam que o fator tem a capacidade de afetar fortemente a variável em questão. Estes *loadings* são frequentemente utilizados para propósitos de diagnóstico e análise.

2.2 Análise de Clusters

A análise de *clusters* é uma técnica exploratória de análise multivariada que permite organizar em grupos “homogêneos”, agrupando indivíduos ou variáveis com base na sua similaridade, relativamente a uma ou mais características (C. Fraley & A. Raftery., 1998). Assim sendo, a análise de *clusters* permite maximizar a homogeneidade dentro dos grupos e a heterogeneidade entre os grupos.

Neste estudo, vai ser utilizada a análise de *clusters* para traçar o perfil dos municípios, permitindo o seu agrupamento consoante as semelhanças existentes entre os mesmos. Esta semelhança vai ser mensurada através de uma análise das várias variáveis que foram selecionadas para avaliar o desempenho dos 308 municípios.

O agrupamento de *clusters* pode ser realizado através de dois métodos: o método hierárquico e o método não-hierárquico.

Os métodos hierárquicos são capazes de fornecer mais que um tipo de repartição de dados. Este método vai gerar vários agrupamentos possíveis e não necessita que já se tenha um número inicial de *clusters* como acontece nos métodos não-hierárquicos. São ainda considerados inflexíveis, pois não permitem a troca de elementos entre grupos. O método hierárquico é mais apropriado para amostras de dimensão inferior a 250. Por sua vez, os métodos hierárquicos são classificados em dois tipos: métodos aglomerativos e métodos divisivos. Nos métodos aglomerativos, primeiramente todos os elementos se encontram separados e vão depois sendo agrupados em etapas, um a um, até ser obtido um único *cluster* que contém todos os elementos. Por fim, é escolhido o número ótimo de *clusters* tendo em conta todas as possíveis combinações. Nos métodos divisivos todos os elementos se encontram juntos num único *cluster* e vão depois sendo separados um a um, até que cada elemento seja o seu próprio *cluster*. Por fim, é escolhido o número ótimo de *clusters* tendo em conta todas as possíveis combinações.

A utilização do método não-hierárquico implica a necessidade de definir inicialmente um número de *clusters*. Este método é caracterizado pela sua flexibilidade, pois os elementos podem ser trocados entre grupos durante a execução do algoritmo. Por norma, estes métodos são mais eficientes que os métodos-hierárquicos especialmente se existe um elevado número de observações. O principal método não hierárquico utilizado corresponde ao método *K-means*. Este método é mais apropriado para amostras de dimensão superior a 250 e utiliza como método de agrupamento a distância euclidiana.

Outro método utilizado corresponde ao *Two Step Cluster*. O método *Two Step Cluster* é o mais apropriado para amostras que possuam variáveis categóricas e é ainda recomendado para amostras de grande dimensão.

2.2.1 Método *Two Step Cluster*

O método utilizado para o desenvolvimento deste estudo foi o *Two Step Cluster*. Este método foi proposto por Chiu *et al.* (2001) e tem como objetivo encontrar um número final de *clusters* através de duas fases:

a) 1ª Fase: Pré-cluster

É utilizada uma abordagem sequencial que tem como objetivo agrupar os indivíduos dando origem a muitos pequenos grupos, grupos estes que vão ser utilizados na próxima fase para que sejam agregados, reduzindo o seu número.

b) 2ª Fase: Clusterização

Os pré-*clusters* definidos no passo anterior vão ser reagrupados dando origem aos *clusters* finais segundo um número ideal de agrupamentos, através da utilização do método hierárquico aglomerativo. Existem duas medidas de distância que podem ser utilizadas neste método, são elas: a distância euclidiana e a distância baseada no logaritmo da verosimilhança, sendo que esta última permite a utilização quer de variáveis contínuas, quer de variáveis categóricas. A medida de distância euclidiana corresponde à distância da "linha reta" entre dois *clusters* e só pode ser utilizada quando todas as variáveis são contínuas. A medida de distância utilizada neste estudo foi a baseada no logaritmo da verosimilhança. Esta medida vai ajustar distribuições de probabilidades às variáveis assumindo que as variáveis contínuas seguem uma distribuição normal e que as categóricas seguem uma distribuição multinomial. Considera ainda que todas as variáveis são independentes. A distância baseada no logaritmo da verosimilhança correspondente à distância entre dois *clusters* i e s é definida como:

$$d(i, s) = \xi_i + \xi_s - \xi_{\langle i, s \rangle} \quad (2.6)$$

$$\xi_i = -n_i \left(\sum_{j=1}^p \frac{1}{2} \log(\hat{\sigma}_{ij}^2 + \hat{\sigma}_j^2) - \sum_{j=1}^q \sum_{l=1}^{m_j} \hat{\pi}_{ijl} \log(\hat{\pi}_{ijl}) \right) \quad (2.7)$$

$$\xi_s = -n_s \left(\sum_{j=1}^p \frac{1}{2} \log(\hat{\sigma}_{sj}^2 + \hat{\sigma}_j^2) - \sum_{j=1}^q \sum_{l=1}^{m_j} \hat{\pi}_{sjl} \log(\hat{\pi}_{sjl}) \right) \quad (2.8)$$

$$\xi_{\langle i, s \rangle} = -n_{\langle i, s \rangle} \left(\sum_{j=1}^p \frac{1}{2} \log(\hat{\sigma}_{\langle i, s \rangle j}^2 + \hat{\sigma}_j^2) - \sum_{j=1}^q \sum_{l=1}^{m_j} \hat{\pi}_{\langle i, s \rangle jl} \log(\hat{\pi}_{\langle i, s \rangle jl}) \right) \quad (2.9)$$

ξ_v pode ser interpretado como uma espécie de dispersão (variância) dentro do cluster v ($v = i, s, \langle i, s \rangle$) e constituído por duas partes a primeira $-n_v \sum_{j=1}^p \frac{1}{2} \log(\hat{\sigma}_{\langle i, s \rangle j}^2 + \hat{\sigma}_j^2)$ mede a dispersão das variáveis contínuas x_j dentro do cluster v . A segunda parte $-n_v \sum_{j=1}^q \sum_{l=1}^{m_j} \hat{\pi}_{\langle i, s \rangle jl} \log(\hat{\pi}_{\langle i, s \rangle jl})$ permite medir a dispersão das variáveis categóricas.

À semelhança do que acontece na utilização da técnica hierárquica aglomerativa, os clusters com a menor distância $d(i,s)$ são agrupados em cada passo. A função do logaritmo da verosimilhança para o passo com k clusters é calculada da seguinte forma:

$$l_k = \sum_{v=1}^k \xi_v \quad (2.10)$$

A função l_k pode ser interpretada como a dispersão dentro dos clusters.

Relativamente ao número de clusters, é possível especificar um número ou então este número pode ser determinado automaticamente, através da utilização do critério BIC (critério *Bayesiano de Schwarz*) ou AIC (critério de informação de *Akaike*). O método que acabou por ser utilizado neste estudo foi o AIC, uma vez que o desempenho do BIC penaliza grandes diferenças entre clusters, como por exemplo, relativamente ao tamanho dos mesmos.

O critério de informação de *Akaike*, é uma técnica que permite mensurar a qualidade de um modelo estatístico (Akaike., 1974). Esta técnica estima a quantidade de informação que é perdida quando se considera um modelo com menos parâmetros face a um inicial, sendo que quanto menor for a informação que determinado modelo perde, maior é a qualidade do modelo ou seja, um bom modelo é aquele que possui um menor valor de AIC.

$$AIC_k = -2l_k + 2r_k \quad (2.11)$$

r_k – corresponde ao número de parâmetros independentes.

CAPÍTULO 3

APRESENTAÇÃO, ANÁLISE E DISCUSSÃO DOS RESULTADOS

O *software* de tratamento estatístico de dados onde foram implementadas as metodologias acima descritas foi o SPSS.

3.1 Resultados obtidos com a Análise Fatorial

3.1.1 Teste de KMO e Bartlett

Tabela 3.1 – Resultado obtido no Teste de KMO e Bartlett em SPSS

Teste de KMO e Bartlett		
Medida Kaiser-Meyer-Olkin de adequação de amostragem.		,760
Teste de esfericidade de Bartlett	Aprox. Qui-quadrado	2480,309
	gl	120
	Sig.	,000

Tal como foi mencionado acima, no teste de esfericidade de Bartlett, as hipóteses a ser consideradas são as seguintes:

H0: A matriz de correlação é uma matriz identidade, não há correlação suficiente entre as variáveis (Análise Fatorial não traz valor acrescentado)

H1: A matriz de correlação é uma matriz de similaridade, existe correlação suficiente entre as variáveis (Análise Fatorial é adequada)

O valor obtido no teste de esfericidade foi de $X^2 = 2480,309$ com 120 graus de liberdade, como $\text{sig}=0,00$ que é inferior a 0,1, 0,05 e 0,01, rejeita-se H0 e como tal é possível assumir que a Análise Fatorial é adequada e que as variáveis estão correlacionadas entre elas, positiva ou negativamente. De notar também que a medida de KMO é de 0,76 o que revela uma capacidade média de adequação de amostragem à Análise Fatorial. Como tal, é útil prosseguir com a Análise Fatorial.

3.1.2 Comunalidades

A comunalidade (*communality*), h^2 , corresponde à soma dos quadrados das cargas fatoriais das variáveis sobre cada fator e indica a eficiência dos mesmos na explicação da variabilidade total. Ou seja, as comunalidades (h^2_j) indicam quanto da variabilidade total está a ser explicada pelo conjunto de fatores. Field (2005) defende que a percentagem de variabilidade explicada de cada variável quando agrupada em determinado fator deve ser superior a 0,7, para um número de variáveis inferior a 30, ou acima de 0,6, se o tamanho da amostra for superior a 250.

Na tabela 3.2 encontra-se a tabela das comunalidades que foi obtida através da implementação desta análise em SPSS.

Tabela 3.2 – Tabela das Comunalidades obtida em SPSS

Comunalidades		
	Inicial	Extração
Receitas municipais p/mil habitantes	1,000	,717
Empresas não financeiras p/ mil habitantes	1,000	,635
Taxa de variação anual das empresas não financeiras	1,000	,667
Ganho médio dos trabalhadores por conta de outrem	1,000	,609
% de empresas criadas no setor Indústria, Construção e Energia	1,000	,686
Taxa média de variação anual da Densidade Populacional	1,000	,751
Índice de Envelhecimento	1,000	,879
Taxa média de variação anual do Índice de Envelhecimento	1,000	,713
População em idade ativa (%)	1,000	,869
População Estrangeira (%)	1,000	,774
Nr. De anos de escolaridade das mulheres que foram mães	1,000	,817
Nr. Médio de anos de escolaridade dos empregadores	1,000	,508
Nr. Médio de anos de escolaridade dos trabalhadores por conta de outrem	1,000	,799
Tx. Média de variação anual do nr. de anos de escolaridade dos trabalhadores por contra de outrem	1,000	,375
Alojamentos Turísticos por área do município	1,000	,557
Alojamentos familiares clássicos por habitante	1,000	,828
Método de Extração: análise de Componente Principal.		

Ao observar a tabela 3.2 é possível concluir que a variável com maior peso corresponde ao Índice de envelhecimento. Sendo que, também a percentagem de população em idade ativa, o número de anos de escolaridade das mulheres que foram mães e o número de alojamentos familiares clássicos por habitante apresentam um peso considerável.

A variável com menor peso corresponde à taxa média de variação anual do número de anos de escolaridade dos trabalhadores por conta de outrem. Também o número médio de anos de escolaridade dos empregadores apresenta um peso pouco significativo.

3.1.3 Variância total explicada

Tabela 3.3 – Tabela da Variância total explicada obtida em SPSS

Variância total explicada									
Componente	Autovalores iniciais			Somadas de extração de carregamentos ao quadrado			Somadas de rotação de carregamentos ao quadrado		
	Total	% de variância	% cumulativa	Total	% de variância	% cumulativa	Total	% de variância	% cumulativa
1	4,705	29,404	29,404	4,705	29,404	29,404	4,078	25,489	25,489
2	2,880	18,002	47,406	2,880	18,002	47,406	2,188	13,674	39,164
3	1,385	8,653	56,059	1,385	8,653	56,059	1,892	11,825	50,988
4	1,234	7,712	63,772	1,234	7,712	63,772	1,636	10,223	61,211
5	,981	6,134	69,906	,981	6,134	69,906	1,391	8,695	69,906
6	,886	5,540	75,446						
7	,824	5,147	80,593						
8	,601	3,753	84,347						
9	,565	3,529	87,876						
10	,484	3,028	90,904						
11	,401	2,505	93,409						
12	,308	1,926	95,335						
13	,282	1,763	97,097						
14	,236	1,476	98,574						
15	,150	,939	99,513						
16	,078	,487	100,000						

Método de Extração: análise de Componente Principal.

A tabela 3.3 indica a percentagem total da variância que é explicada pelos fatores obtidos. Os fatores devem apresentar variância acumulada de pelo menos 60%.

De notar que, o número de variáveis latentes não observáveis foi 16 e que os fatores/componentes extraídos apresentam uma percentagem cumulativa de variância total explicada de 69,906%. Ou seja, com 5 fatores é possível explicar 69,906% da variação dos dados. Foi definido um número fixo de 5 fatores, dado que, o quinto fator tem um valor próprio muito próximo de 1.

3.1.4 Matriz de componentes

Tabela 3.4 – Tabela da Matriz de componentes obtida em SPSS

Matriz de componente ^a					
	Componente				
	1	2	3	4	5
Receitas municipais p/mil habitantes	,798	,153	,104	-,213	-,006
Empresas não financeiras p/mil habitantes	,416	,369	,366	,405	,166
Taxa de variação anual das empresas não financeiras	,504	,252	,495	,322	,011
Ganho médio dos trabalhadores por conta de outrem	-,499	,523	-,231	-,171	,060
% de empresas criadas no setor Ind, Const e Energia	,124	-,532	-,122	,469	-,392
Taxa média de variação anual da Densidade Populacional	-,774	,217	,161	,075	-,272
Índice de Envelhecimento	,851	-,049	-,378	,076	,062
Taxa media de variação anual do Índice de Envelhecimento	-,500	-,499	,053	,348	,299
População em idade ativa (%)	-,810	-,275	,365	,033	,053
População Estrangeira (%)	,061	,622	,456	,128	-,398
Nr. De anos de escolaridade das mulheres que foram mães	-,218	,320	-,490	,651	-,058
Nr. Médio de anos de escolaridade dos empregadores	-,111	,672	-,096	-,179	,054
Nr. Médio de anos de escolaridade dos trabalhadores por conta de outrem	-,263	,755	-,366	,160	,011

Tx. Média de variação anual do nr. de anos de escolaridade dos trabalhadores por contra de outrem	,446	-,203	-,019	,149	,336
Alojamentos Turísticos por área do município	-,138	,397	,170	,135	,577
Alojamentos familiares clássicos por habitante	,892	,124	-,029	-,002	-,126

Método de Extração: análise de Componente Principal.

a. 5 componentes extraídos.

A matriz de componentes revela quais as variáveis que apresentam maior poder explicativo em cada fator. No entanto, esta é a solução não rotacionada. Após a rotação, através da utilização do método *Varimax*, foi obtida a seguinte matriz de componentes:

Tabela 3.5 – Tabela da Matriz de componentes rotacionada obtida em SPSS

Matriz de componente rotativa^a					
	Componente				
	1	2	3	4	5
Receitas municipais p/mil habitantes	,740	,038	,277	-,300	,030
Empresas não financeiras p/mil habitantes	,216	,084	,751	,095	-,085
Taxa de variação anual das empresas não financeira	,253	-,065	,768	-,086	,027
Ganho médio dos trabalhadores por conta de outrem	-,249	,592	-,218	,307	,236
% de empresas criadas no setor Ind, Const e Energia	,010	-,802	-,028	,204	-,030
Taxa média de variação anual da Densidade Populacional	-,710	,139	-,052	,203	,430
Índice de Envelhecimento	,874	-,192	,047	,096	-,260
Taxa média de variação anual do Índice de Envelhecimento	-,634	-,268	-,095	,086	-,472
População em idade ativa (%)	-,914	-,042	-,091	-,153	-,006

População Estrangeira (%)	-,003	,175	,588	,064	,628
Nr. De anos de escolaridade das mulheres que foram mães	-,075	-,011	,043	,899	-,019
Nr. Médio de anos de escolaridade dos empregadores	,077	,619	,052	,203	,274
Nr. Médio de anos de escolaridade dos trabalhadores por conta de outrem	-,005	,555	,019	,657	,243
Tx. Média de variação anual do nr. de anos de escolaridade dos trabalhadores por contra de outrem	,322	-,133	,171	-,076	-,468
Alojamentos Turísticos por área do município	-,185	,539	,344	,110	-,319
Alojamentos familiares clássicos por habitante	,839	-,146	,304	-,095	,039

Método de Rotação: Varimax com Normalização de Kaiser.

A matriz de componentes rotacionada indica, após realizar a rotação dos fatores, que variáveis têm maior poder explicativo em cada fator. Sendo que cada linha representa uma variável e cada uma das colunas corresponde a um dos fatores. Os valores dispostos na matriz representam as cargas fatoriais, isto é, a correlação de cada variável com cada fator. Identificamos as variáveis que explicam cada fator observando os valores das cargas.

Através da análise desta tabela foram identificados 5 fatores a que demos as seguintes designações:

Fator 1 - Continuidade Demográfica;

Fator 2 - Escolaridade e Bons Salários;

Fator 3 - Empresas e População Estrangeira;

Fator 4 - População Jovem Escolarizada;

Fator 5 - População Estrangeira Pouco Escolarizada.

Relativamente ao Fator Continuidade Demográfica as variáveis com maior carga fatorial positiva são: Receitas municipais por mil habitantes; Índice de envelhecimento e Alojamentos familiares clássicos por habitante. Quanto às variáveis com maior carga fatorial negativa, são elas: Percentagem de população em idade ativa; Taxa média de variação anual da densidade populacional e Taxa média de variação anual de Índice de envelhecimento. Resumidamente municípios que têm valores elevados positivos neste fator são municípios que, quando comparados com os restantes, têm um elevado valor de receitas municipais por mil habitantes, elevado Índice de envelhecimento, mais alojamentos familiares clássicos por habitante, menor percentagem de população em idade ativa e uma taxa média de variação anual de densidade populacional e de Índice de envelhecimento mais baixa também, ou

seja, municípios onde a densidade populacional tem vindo a baixar e onde o Índice de envelhecimento ou tem diminuído ou não tem vindo a aumentar tanto como nos restantes municípios do país.

Quanto ao Fator Escolaridade e Bons Salários as variáveis com maior carga fatorial positiva são: Número médio de anos de escolaridade dos trabalhadores por conta de outrem; Ganho médio dos trabalhadores por conta de outrem; Número médio de anos de escolaridade dos empregadores, Alojamentos turísticos por área do município. Relativamente às variáveis com maior carga fatorial negativa destaca-se essencialmente a percentagem de empresas criadas no setor Indústria, Construção e Energia. Conclui-se assim que municípios que têm valores positivos elevados neste fator são municípios em que os trabalhadores e empregadores são mais escolarizados, os salários são mais elevados e existe um elevado número de alojamentos turísticos por área de município e em contrapartida têm uma baixa percentagem de empresas criadas no setor Indústria, Construção e Energia.

No Fator Empresas e População Estrangeira as variáveis que apresentam maior carga fatorial positiva são: Taxa de variação anual das empresas não financeiras; Percentagem de população estrangeira e Empresas não financeiras por mil habitantes. Isto permite concluir que municípios que apresentam valores positivos elevados neste fator são municípios com uma percentagem mais elevada de população estrangeira, e onde existe um maior número de empresas não financeiras.

No que diz respeito ao Fator População Jovem Escolarizada as variáveis com maior carga fatorial positiva são: Número de anos de escolaridade das mulheres que foram mães e Número médio de anos de escolaridade dos trabalhadores por conta de outrem. É possível desta forma concluir que municípios que têm valores positivos neste fator são municípios onde as mulheres que foram mães são mais escolarizadas e onde o número médio de anos de escolaridade dos trabalhadores por conta de outrem é mais elevado também.

Relativamente ao Fator População Estrangeira Pouco Escolarizada, as variáveis com maior carga fatorial positiva são: Percentagem de população estrangeira e Taxa média de variação anual da densidade populacional. Quanto às variáveis com maior carga fatorial negativa, são elas: Taxa média de variação anual do número médio de anos de escolaridade dos trabalhadores por conta de outrem e Taxa média de variação anual do Índice de envelhecimento. Ou seja, municípios que têm valores positivos elevados neste fator são municípios que apresentam uma mais elevada percentagem de população estrangeira, uma variação positiva relativamente à densidade populacional, ou seja, municípios onde os valores de densidade populacional têm vindo a aumentar ao longo dos anos ou, pelo menos, a diminuir menos que na generalidade dos municípios. São ainda municípios que apresentam uma variação negativa relativamente ao Índice de envelhecimento e ao número médio de anos de escolaridade dos trabalhadores por conta de outrem, ou seja, municípios onde o Índice de envelhecimento e o número de anos de escolaridades dos trabalhadores por conta de outrem ou têm diminuído ou não têm vindo a aumentar tanto como nos restantes municípios do país.

São considerados municípios mais competitivos, ou seja, que apresentam um melhor desempenho não conferindo uma situação de risco ao país, aqueles que apresentam valores mais elevados de receitas municipais por mil habitantes, que apresentam um maior número de empresas não financeiras por mil habitantes e onde esse mesmo número tem vindo a aumentar mais. São também mais competitivos os municípios onde o ganho médio dos trabalhadores por conta de outrem é mais elevado. Outro fator de relevo consiste em apresentar uma percentagem elevada na criação de empresas no setor da Indústria,

Construção e Energia. Também uma população escolarizada é essencial e consequentemente o aumento relativamente ao nível desta escolarização ao longo dos últimos anos. Municípios que apresentam um valor mais elevado quanto ao número de alojamentos familiares clássicos por habitante e de alojamentos turísticos por área do município conferem, por sua vez, maior competitividade ao município também. O aumento da densidade populacional é também um fator que pode apresentar capacidade de influenciar positivamente o desempenho de determinado município conferindo vantagens competitivas. Por outro lado, conferem uma situação de risco os municípios que apresentam um Índice de envelhecimento mais elevado e onde o mesmo tem vindo a aumentar mais ao longo dos anos. Também uma baixa percentagem de população em idade ativa afeta o desempenho do município. Relativamente à percentagem de população estrangeira, valores elevados neste indicador podem constituir uma vantagem competitiva aos municípios podendo apresentar um impacto positivo em termos económicos. Para além disso a imigração tem sido também responsável pelo aumento de efetivos em idade jovem e ativa em Portugal, contribuindo ainda para o aumento da taxa de natalidade (Gomes., 2017).

3.2 Normalização das variáveis

De forma a facilitar a interpretação dos dados que foram obtidos procedeu-se à normalização das variáveis, uma vez que, a maioria das variáveis utilizadas neste estudo encontram-se expressas em diferentes escalas com dispersões bastante distintas. Este processo de *standardização* tem como objetivo reduzir as variáveis à mesma escala, tornando-as independentes das unidades de medida em que foram originalmente expressas, facilitando desta forma a sua interpretação. Para efetuar a normalização procedeu-se à utilização da técnica *Z-scores*:

$$x_{i,j} = \frac{x_{i,j} - \bar{x}_j}{s_j} \quad (3.1)$$

Onde, $x_{i,j}$ representa a i -ésima observação da j -ésima variável, \bar{x}_j corresponde à média da j -ésima variável e s_j indica o desvio padrão da j -ésima variável.

Na tabela 3.6 estão listados os Municípios com os 5 valores standardizados mais elevados e com os 5 valores standardizados mais baixos em cada um dos fatores:

Tabela 3.6 - Valores mais positivos e negativos nos fatores

Escolaridade e bons salários		Empresas e população estrangeira		População jovem escolarizada	
Municípios	Valor standardizado	Municípios	Valor standardizado	Municípios	Valor standardizado
Porto	5,35	Albufeira	4,42	Oeiras	2,62
Funchal	5,26	Lagos	3,97	Coimbra	2,29
Lisboa	5,16	Aljezur	3,68	Vimioso	2,23
Oeiras	2,77	Alfândega da Fé	3,59	Lisboa	2,18
Alcochete	2,58	Vila do Bispo	3,50	Monção	2,03
Lousada	-1,87	Constância	-1,62	Corvo	-2,39
Paredes de Coura	-1,94	Vila Velha de Ródão	-1,65	Ribeira Grande	-2,82
Felgueiras	-2,44	Vila Nova da Barquinha	-1,71	Monforte	-3,27
Vila Franca do Campo	-2,66	Nelas	-1,74	Porto Moniz	-3,53
Pampilhosa da Serra	-3,73	Alcochete	-1,82	Mourão	-4,92

Continuidade Demográfica		População estrangeira pouco escolarizada	
Municípios	Valor standardizado	Municípios	Valor standardizado
Alcoutim	4,27	Vila do Bispo	3,80
Vila Velha de Ródão	3,43	Loulé	3,10
Idanha-a-nova	3,02	Lagos	3,03
Pampilhosa da Serra	2,84	Albufeira	2,94
Sabugal	2,66	Aljezur	2,79
Santa Cruz	-1,90	Manteigas	-2,25
Paços de Ferreira	-1,95	Vinhais	-2,28
Vizela	-1,98	Tarouca	-2,53
Lousada	-2,01	Porto	-3,20
Funchal	-2,26	Funchal	-5,69

3.3 Resultados obtidos do *Two Step Cluster*

As variáveis utilizadas para aplicação deste método correspondem aos *factor loadings* obtidos previamente através da implementação da Análise Fatorial e posteriormente analisados e interpretados. São eles: Fator 1- Continuidade Demográfica; Fator 2 - Escolaridade e Bons Salários; Fator 3 - Empresas e População Estrangeira; Fator 4 - População Jovem Escolarizada; Fator 5 - População Estrangeira Pouco Escolarizada.

Após implementação deste método de análise estatística em SPSS foi obtido o seguinte *output*:

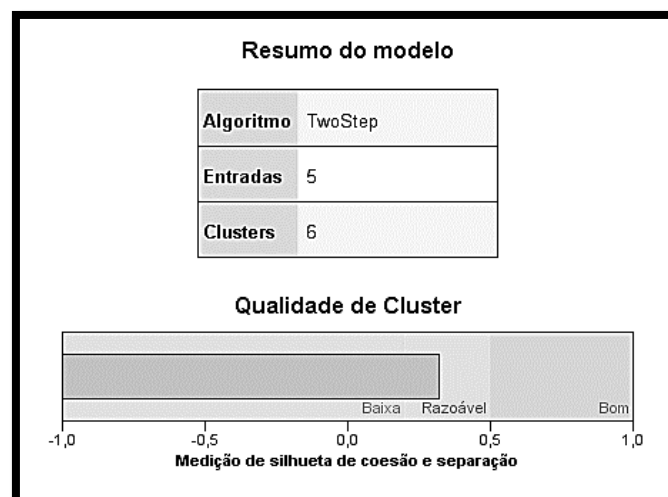


Figura 3.1 - Resumo do modelo e Qualidade do cluster obtido em SPSS

Na figura 3.1 a “Qualidade de Cluster” é avaliada através do coeficiente de silhueta, proposto por Rousseeuw (1987). Trata-se de uma medida que reflete a coesão e a separação das repartições de *clusters*, permitindo avaliar a qualidade do agrupamento que foi obtido pelo método da Análise de *Clusters* em questão, indicando o quão semelhantes são as observações que estão contidas em

determinado *cluster* em comparação com as observações contidas nos restantes *clusters*. A coesão permite avaliar a homogeneidade existente entre as observações pertencentes a determinado *cluster*. A separação, por sua vez, quantifica o grau de separação existente entre os *clusters*. Para obter o coeficiente de silhueta global (CS) calcula-se a média de todos os valores de silhueta, sendo que o valor de silhueta para cada observação *i* é obtido da seguinte forma:

$$S(i) = \frac{b_i - a_i}{\max(b_i, a_i)} \quad (3.2)$$

Onde, a_i representa a distância média entre a observação *i* e todas as observações do mesmo *cluster*, e b_i corresponde à distância média entre a observação *i* e todas as observações do *cluster* vizinho mais próximo.

O valor $s(i)$ varia entre -1 e 1 e, se o valor de silhueta estiver próximo ou for igual a 1 então os dados foram bem agrupados e atribuídos ao *cluster* apropriado, $a(i) < b(i)$. Caso o valor seja aproximadamente 0, significa que a observação também poderia ser afeta ao *cluster* mais próximo, isso ocorre quando $a(i) = b(i)$, indicando que a observação se encontra num ponto intermédio entre dois *clusters*. Se o valor for próximo ou igual a -1 significa que a observação foi mal agrupada, pois $a(i) > b(i)$, ou seja, a observação *i*, em média, está mais distante das observações do seu próprio *cluster*, estando mais próximo das observações do *cluster* vizinho mais próximo. Concluindo, quanto mais próximo de 1, melhor a qualidade do *cluster* (Souza., 2007).

$$CS = \frac{1}{n} \sum_{i=1}^n S(i) \quad (3.3)$$

em que n é o número de observações. Na tabela 3.7 estão representados os valores de referência empiricamente aceites para interpretar o coeficiente de silhueta.

Tabela 3.7 - Interpretação dos valores do coeficiente de silhueta

SC	Interpretação
0,71 - 1	<i>Clusters</i> possuem uma estrutura bastante robusta
0,51 – 0,7	<i>Clusters</i> possuem uma estrutura razoável
0,26 – 0,5	<i>Clusters</i> podem apresentar uma estrutura fraca e artificial
≤ 0,25	Nenhuma estrutura foi descoberta

Fonte: Adaptada de Kaufman & Rousseeuw (1990)

O valor de silhueta (CS) obtido foi de 0,3 e o número de *clusters* obtido foi 6, o que significa que as variáveis utilizadas, neste caso os *factor loadings*, conseguem separar os municípios em 6 *clusters*, mas sem uma capacidade de discriminação de dados suficiente para ser considerada boa. Isto significa que, para alguns municípios, uma pequena alteração nos valores das respetivas variáveis poderia originar uma mudança para outro *cluster*. No entanto, dado haver bastante coerência geográfica, a caracterização de cada *cluster* irá dar-nos, adicionalmente, uma visão do perfil social e económico das regiões de predominância de cada *cluster*.

Tabela 3.8 - Tamanho do maior e menor cluster obtido em SPSS

Tamanho do menor Cluster	3 (1%)
Tamanho do maior Cluster	87 (28,2%)

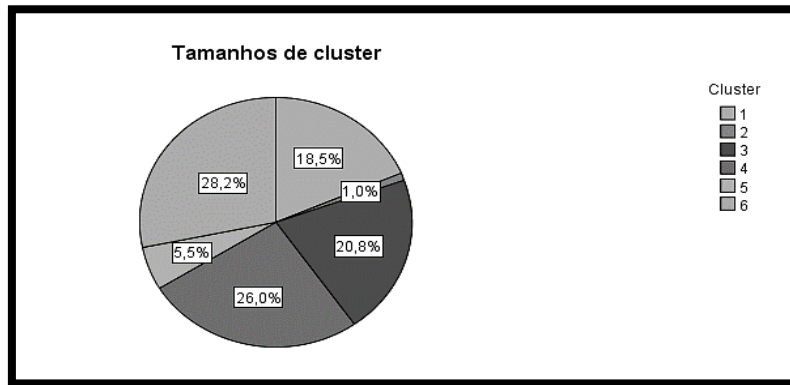
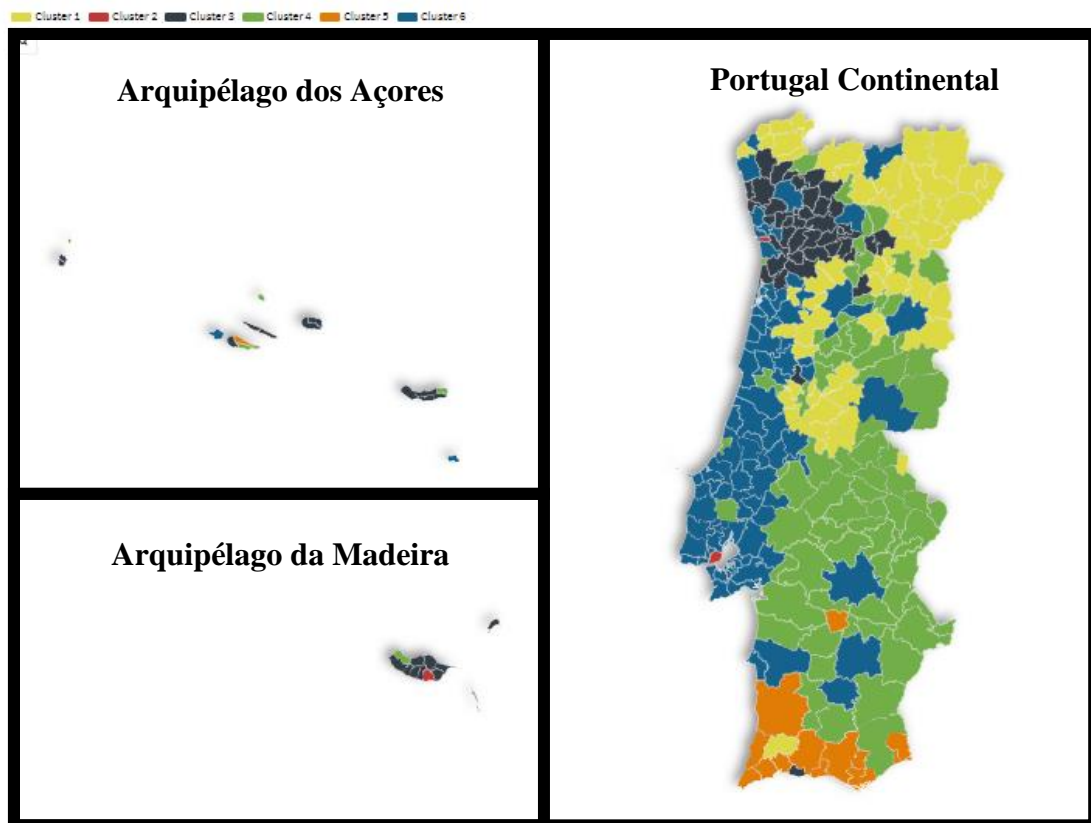


Figura 3.2 - Tamanhos de cluster obtido em SPSS

Distribuição territorial dos clusters
(Método Two Step clustering)



Fonte: Plordata

Figura 3.3 - Distribuição geográfica dos municípios elaborado no Flourish

(Nota: o mapa interativo está disponível em <https://public.flourish.studio/visualisation/4199571/>)

Através da análise da tabela 3.8 e das figuras 3.2 e 3.3, conclui-se que ao implementar o método *Two Step Cluster* em SPSS foram obtidos 6 *clusters* de dimensões bastante distintas. O *cluster* 1 contém 57 municípios que se encontram, na sua grande maioria, a norte de Portugal na zona de Trás os Montes e Alto Douro, contendo ainda alguns municípios da região da Beira Alta. O *cluster* 2 é o *cluster* de menor dimensão, sendo constituído somente por 3 municípios – Lisboa, Porto e Funchal – podendo-se desta forma concluir que este *cluster* é constituído pelas “grandes” metrópoles. O *cluster* 3, por sua vez, é constituído por 64 municípios pertencendo a sua grande maioria à região do Minho e Douro Litoral e contendo ainda a maioria dos municípios pertencentes à Região Autónoma da Madeira e ao Arquipélago dos Açores. Relativamente ao *cluster* 4, fazem parte deste *cluster* 80 municípios, municípios estes que se situam essencialmente na região da Beira Baixa, Alto Alentejo e Baixo Alentejo. Inclui, no entanto, dois municípios da orla atlântica (Nazaré e Espinho) e outros dois municípios perto do litoral (Alenquer e Soure). O *cluster* 5 é constituído por 17 municípios situados essencialmente na zona do Algarve, contendo ainda alguns municípios pertencentes à Costa Vicentina. Por fim, o *cluster* 6 corresponde ao *cluster* de maior dimensão, contendo 87 municípios, que se situam na sua grande maioria na zona litoral norte e centro de Portugal, mas que inclui também alguns municípios que são capitais de distrito, fora da zona litoral, como Viseu, Guarda, Castelo Branco, Évora e Beja.

3.3.1 Descrição dos Clusters

Esta descrição vai ser feita com recurso a uma análise de *boxplots* ou caixa com bigodes. As caixas com bigodes vão fornecer uma informação visual relativamente às seguintes medidas amostrais: localização, dispersão, assimetria, comprimento da cauda da distribuição e *outliers*. Relativamente à localização, este gráfico fornece informação quanto à localização central que corresponde à mediana amostral ou segundo quartil, acabando por fornecer ainda informação relativamente a localizações representativas, são elas o mínimo, o primeiro quartil, o terceiro quartil e o máximo. A dispersão, por sua vez, é representada pela amplitude do gráfico, que corresponde à diferença entre o máximo e o mínimo. Quanto maior for a amplitude, maior é a dispersão dos dados. Por fim, a assimetria é analisada consoante a posição relativa dos quartis. Uma distribuição simétrica iria apresentar a mediana no centro do retângulo sendo esta coerente com a média, caso a mediana seja próxima do primeiro quartil então os dados são positivamente assimétricos, caso a mediana seja próxima do terceiro quartil então os dados são negativamente assimétricos.

Outlier corresponde a um valor que é significativamente distinto quando comparado com o restante conjunto de observações, ou valor discrepante em relação aos restantes dados que foram observados (Barnett *et al.*, 1994). Na figura 3.4 encontra-se um exemplo deste gráfico.

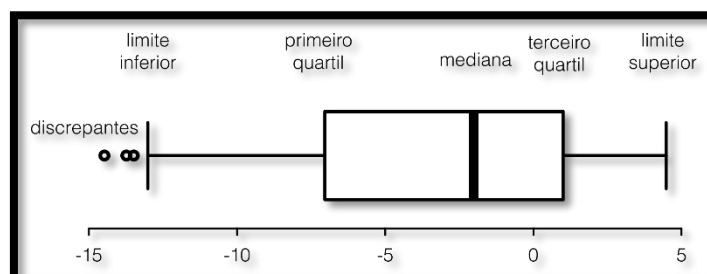


Figura 3.4 - Exemplo de caixa com bigodes/boxplot

Fonte: Wikipédia

O primeiro quartil, corresponde ao número que deixa 25% das observações à esquerda e 75% a direita. O terceiro quartil, corresponde ao valor que deixa 75% dos dados à esquerda e 25% à sua direita. A mediana ou o segundo quartil, por sua vez, corresponde ao valor que deixa 50% dos dados à sua esquerda e 50% à sua direita.

3.3.2 Cluster 1

Cluster caracterizado pela conciliação entre a continuidade demográfica e bons níveis de escolarização da população jovem, mas ainda com população ativa pouco escolarizada e baixos salários.

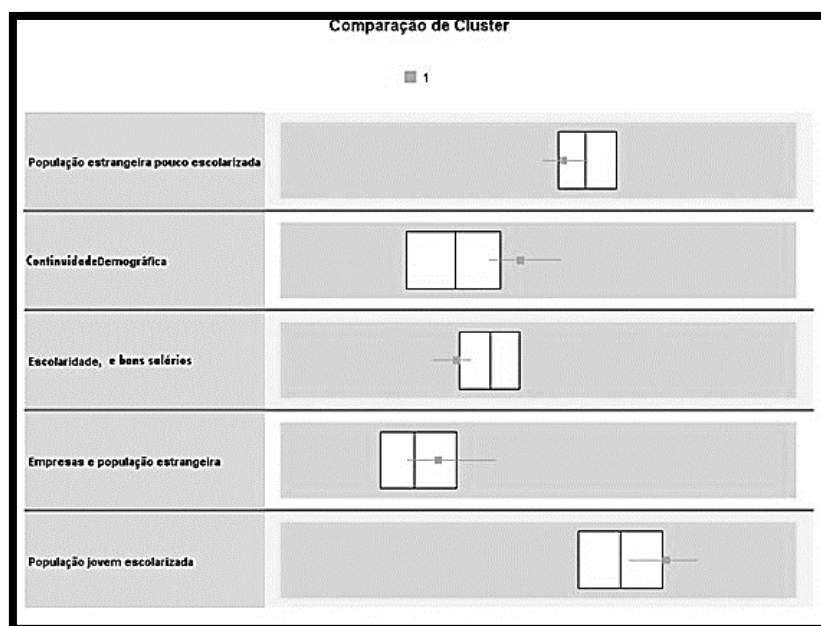


Figura 3.5 - Cluster 1 obtido em SPSS

Valores centrados no percentil 25 da distribuição nacional relativamente aos fatores população estrangeira pouco escolarizada e escolaridade e bons salários. Valores entre a mediana e o percentil 75 no fator empresas e população estrangeira, acima do percentil 75 no fator continuidade demográfica, e centrados no percentil 75 no fator população jovem escolarizada.

O município de Vimioso apresenta valores em linha com o centro da distribuição de cada um dos fatores e pode, por isso, considerar-se o município de referência do *cluster* 1.

Tabela 3.9 - Município mais característico do Cluster 1

Município	Continuidade demográfica	Escolaridade e salários	Empresas e população estrangeira	População jovem escolarizada	População estrangeira pouco escolarizada
Vimioso	2,346	-1,272	0,517	2,226	0,064

Fator continuidade demográfica

Ao *standardizar* os dados e posteriormente ao recorrer à observação dos seus valores, conclui-se que os dados relativos à variável taxa média de variação anual da densidade populacional são na sua grande maioria negativos nos municípios pertencentes a este *cluster*, ou seja, são municípios onde a densidade populacional tem vindo a diminuir mais ao longo dos anos. Destaca-se essencialmente o municí-

pio Almeida cujo valor *standardizado* para essa variável é de -2,07 o que significa que a taxa média de variação anual da densidade populacional neste município se encontra 2,07 desvios padrões abaixo da média.

Outra variável que se revela bastante consistente possibilitando uma melhor interpretação deste *cluster* corresponde ao Índice de envelhecimento. A grande maioria dos municípios pertencentes a este *cluster* apresentam valores acima da média relativamente ao Índice de envelhecimento, excetuando os seguintes municípios: Valença cujo valor *standardizado* para essa variável é de -0,056 e Bragança com -0,068. Ainda que nestes municípios o valor *standardizado* seja negativo, é de notar que o mesmo não apresenta um valor muito afastado da média.

O mesmo acontece com a percentagem de população em idade ativa: os municípios pertencentes a este *cluster* apresentam, maioritariamente, valores negativos em comparação com os restantes municípios nesta variável. Destacando-se os municípios: Vimioso cujo valor *standardizado* para essa variável é de -2,593, Vinhais com -2,848, Sabugal com -2,026 e Pampilhosa da Serra com -2,354.

Quanto à variável alojamentos familiares clássicos por habitante os valores *standardizados* dos municípios pertencentes a este *cluster*, nesta variável, são na sua grande maioria elevados. Destacam-se os municípios: Montalegre cujo valor *standardizado* para essa variável é de 2,123 e Vimioso com 2,094. Existem, no entanto, municípios pertencentes a este *cluster* que apresentam valores *standardizados* negativos nesta variável, destacando-se o município Anadia cujo valor *standardizado* para essa variável é de -0,729, o que significa que o número de alojamentos familiares clássicos, por habitante, neste município se encontra 0,729 desvios padrões abaixo da média. Anadia será, assim, um município de algum modo atípico no *cluster* 1 quanto a esta variável. De notar que, geograficamente, Anadia se situa na fronteira entre o *cluster* 1 e o *cluster* 6.

Os valores apresentados nos municípios pertencentes a este *cluster* nestas variáveis explicam o motivo de o fator continuidade demográfica se encontrar acima do percentil 75 neste *cluster*, uma vez que, recorde-se, este fator é caracterizado por valores elevados relativamente às variáveis: Índice de envelhecimento e Alojamentos familiares clássicos por habitante e valores baixos relativamente às variáveis: Percentagem de população em idade ativa e Taxa média de variação anual da densidade populacional.

Fator escolaridade e bons salários

Relativamente à variável ganho médio dos trabalhadores por conta de outrem, os valores *standardizados* para esta variável nos municípios pertencentes ao *cluster* 1 são na sua grande maioria negativos.

Quanto à percentagem de empresas criadas no setor Indústria, Construção e Energia, a mesma é bastante elevada nos municípios pertencentes a este *cluster*, destacando-se essencialmente os municípios: Monção cujo valor *standardizado* para essa variável é de 2,421, o que significa que a percentagem de empresas criadas no setor Indústria, Construção e Energia neste município se encontra 2,421 desvios padrões acima da média. Ainda de destacar Paredes de Coura que apresenta um valor *standardizado* de 3,490, Anadia com 2,188, Sabugal com 2,286, Pampilhosa da Serra com 5,440 e Vila de Rei com 2,662. Existem, porém, municípios pertencentes a este *cluster* que apresentam valores *standardizados* negativos nesta variável, destacando-se o município Marvão cujo valor *standardizado* para essa variável é de -0,759, o que significa que a percentagem de empresas criadas no setor Indústria, Construção e Energia neste município se encontra 0,759 desvios padrões abaixo da média. No *cluster* 1, Marvão é

um município atípico nesta variável e também territorialmente se encontra fora da zona de predominância do cluster. Todos os municípios que circundam Marvão pertencem ao *cluster 4*.

Relativamente aos alojamentos turísticos por área do município, os valores *standardizados* nesta variável para os municípios pertencentes a este *cluster* são negativos em todos os municípios.

O facto de o ganho médio dos trabalhadores por conta de outrem apresentar na maior parte dos municípios pertencentes a este *cluster* valores *standardizados* negativos, a percentagem de empresas criadas no setor Indústria, Construção e Energia ser elevada e o facto de os valores *standardizados* na variável alojamentos turísticos por área de município serem negativos pode explicar o facto de este *cluster* apresentar valores centrados no percentil 25 relativamente ao fator escolaridade e bons salários, uma vez que, este fator é caracterizado por valores baixos na variável percentagem de empresas criadas no setor Indústria, Construção e Energia e valores elevados relativamente à variável ganho médio dos trabalhadores por conta de outrem e alojamentos turísticos por área de município.

Fator empresas e população estrangeira

Ainda que os valores *standardizados* da variável percentagem de população estrangeira nos municípios pertencentes a este *cluster* sejam maioritariamente negativos, existem valores bastante elevados em alguns municípios relativamente às variáveis empresas não financeiras por mil habitantes e à taxa de variação das mesmas, o que pode explicar o facto de este *cluster* apresentar valores entre a mediana e o percentil 75 no fator empresas e população estrangeira. Nesta sequência, destacam-se os municípios: Murça cujo valor *standardizado* para essa variável é de 3,088, Valpaços com 3,531 Alfândega da Fé com 3,453, Macedo de Cavaleiros com 2,557, Mogadouro com 2,662, Vinhais com 3,412, Penedono com 3,684, Pinhel com 2,149 e Vila Nova de Foz Côa com 2,223 por apresentarem valores *standardizados* elevados na variável empresas não financeiras por mil habitantes.

Relativamente à taxa de variação das mesmas destacam-se os seguintes municípios: Alfândega da Fé cujo valor *standardizado* para essa variável é de 4,686, Freixo de Espada à Cinta com 6,227, Torre do Moncorvo com 3,358, Vila Flor com 2,259 e Vinhais 2,098. Existem, no entanto, municípios pertencentes a este *cluster* que apresentam valores *standardizados* negativos nestas variáveis, destacando-se os municípios: Pampilhosa da Serra cujo valor *standardizado* para a variável empresas não financeiras por mil habitantes é de -1,752 e o município Manteigas cujo valor *standardizado* para a variável relativa à variação das empresas não financeiras por mil habitantes é de -1,072.

Fator população jovem escolarizada

Este fator encontra-se centrado no percentil 75 neste *cluster*, principalmente fruto do facto de o número de anos de escolaridade das mulheres que foram mães apresentarem, nos municípios pertencentes a este cluster, valores maioritariamente positivos. Neste sentido destaca-se o município Vimioso cujo valor *standardizado* para essa variável é de 2,812, o que significa que o número de anos de escolaridade das mulheres que foram mães neste município se encontra 2,812 desvios padrões acima da média. Existem, no entanto, municípios pertencentes a este *cluster* que contrariam esta tendência apresentando valores *standardizados* negativos nesta variável. Neste sentido destacam-se os municípios: Freixo de Espada à Cinta cujo valor *standardizado* é de -0,718, Macedo de Cavaleiros com -0,757, Sernancelhe com -1,103, Vouzela com -0,842 e Castanheira de Pêra com -0,877.

A variável número médio de anos de escolaridade dos trabalhadores por conta de outrem tanto apresenta valores positivos como negativos nos municípios pertencentes a este *cluster*. No entanto, como a

carga fatorial mais significativa diz respeito à variável número médio de anos de escolaridade das mulheres que foram mães, estes valores acabam por pouco influenciar o facto de este *cluster* apresentar valores entre a mediana e o percentil 75 no fator população jovem escolarizada.

Fator população estrangeira pouco escolarizada

O facto de este *cluster* apresentar valores centrados no percentil 25 relativamente ao fator população estrangeira pouco escolarizada deve-se a serem maioritariamente negativos os valores *standardizados* das variáveis percentagem de população estrangeira e taxa média de variação anual da densidade populacional, uma vez que, este fator é caracterizado por percentagens elevadas de população estrangeira bem como variações positivas relativamente à densidade populacional.

Resumindo, conclui-se que este *cluster* inclui municípios que apresentam um elevado número de alojamentos familiares clássicos por habitante, que apresentam um elevado índice de envelhecimento, uma baixa percentagem de população em idade ativa e ainda municípios com valores relativamente elevados quanto ao nível de escolarização das mulheres que foram mães. Outro aspeto caracterizante destes municípios prende-se ainda com o facto de o ganho médio dos trabalhadores por conta de outrem ser baixo, assim como o número de alojamentos turísticos por área de município, e onde a percentagem de empresas no setor da Indústria, Construção e Energia é elevada.

3.3.3 Cluster 2

Cluster caracterizado pela conciliação entre bons níveis de escolarização da população, bons salários e onde existe uma elevada aposta no turismo.

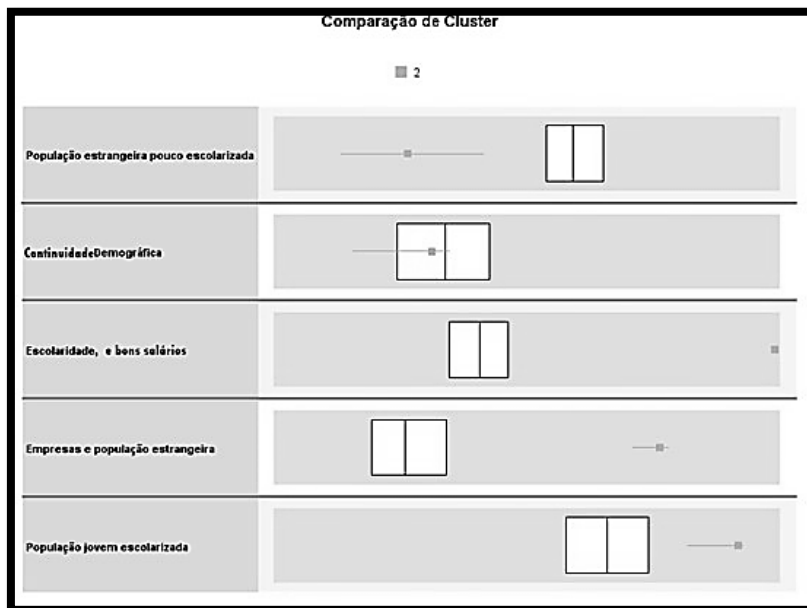


Figura 3.6 - Cluster 2 obtido em SPSS

Valores acima do percentil 75 nos fatores: Escolaridade e bons salários; Empresas e população estrangeira e População jovem escolarizada. Valores abaixo do percentil 25 no fator população estrangeira pouco escolarizada e valores situados entre o percentil 25 e a mediana no fator continuidade demográfica.

Este *cluster* é constituído somente por 3 municípios: Lisboa, Porto e Funchal e apesar de alguns valores não estarem em linha com o centro da distribuição de cada um dos fatores, é possível considerar qualquer um destes municípios como representante do *cluster* 2.

Tabela 3.10 - Municípios característicos do Cluster 2

Município	Continuidade demográfica	Escolaridade e salários	Empresas e população estrangeira	População jovem escolarizada	População estrangeira pouco escolarizada
Lisboa	-0,227	5,351	2,259	1,980	-3,197
Porto	0,254	5,161	2,929	2,183	-0,303
Funchal	-2,258	5,255	3,165	0,490	-5,692

Fator continuidade demográfica

Nas variáveis taxa média de variação anual do Índice de envelhecimento e percentagem de população em idade ativa os valores *standardizados* nos municípios Lisboa e Porto são negativos e no Funchal o valor *standardizado* nestas variáveis é positivo, estes valores contraditórios entre municípios explica o porquê de neste *cluster* o fator continuidade demográfica se encontrar entre o percentil 25 e a mediana mas próximo da mediana.

Fator escolaridade e bons salários

Quanto à variável ganho médio dos trabalhadores por conta de outrem, nos 3 municípios pertencentes a este *cluster* os valores *standardizados* são bastante elevados, destacando-se os municípios: Lisboa e Porto cujo valor *standardizado* nessa variável é de 3,712 e 2,323, respetivamente.

Relativamente à variável alojamentos turísticos por área do município os valores *standardizados* para os 3 municípios pertencentes a este *cluster* são bastante elevados apresentando os seguintes valores: Porto cujo valor *standardizado* para essa variável é de 8,598, Lisboa com 7,188 e Funchal com 12,044.

Ao observar estes valores percebemos o porquê de o fator escolaridade e bons salários se encontrar acima do percentil 75 neste *cluster*, uma vez que, nestes 3 municípios os valores *standardizados* relativos às variáveis ganho médio dos trabalhadores por conta de outrem e alojamentos turísticos por área de município são elevados, o que é uma característica desse fator. Para além disso, os valores *standardizados* para estes 3 municípios relativamente à variável percentagem de empresas criadas no setor Indústria, Construção e Energia estão abaixo da média outra característica do fator escolaridade e bons salários.

De realçar também que o número médio de anos de escolaridade dos trabalhadores por conta de outrem apresenta igualmente valores *standardizados* elevados nos 3 municípios, o que também contribui para que o fator escolaridade e bons salários se encontre acima do percentil 75 neste *cluster*, dado que é outra das características deste fator. Destacam-se principalmente os municípios: Lisboa e Porto cujo valor *standardizado* nessa variável é de 3,811 e 3,317, respetivamente.

Fator empresas e população estrangeira

Ao observar os valores *standardizados* da variável empresas não financeiras por mil habitantes nos municípios pertencentes a este *cluster* podemos concluir que os mesmos são elevados, o que pode

explicar o facto de o fator empresas e população estrangeira se encontrar acima do percentil 75 no *cluster* 2. Destaca-se principalmente o município Lisboa cujo valor *standardizado* para essa variável é de 2,904, o que significa que o número de empresas não financeiras por mil habitantes neste município se encontra 2,904 desvios padrões acima da média.

Fator população jovem escolarizada

Ao observar os valores *standardizados* destes 3 municípios nas variáveis: Número médio de anos de escolaridade dos trabalhadores por conta de outrem; Número médio de anos de escolaridade dos empregadores e o Número de anos de escolaridade das mulheres que foram mães, verificamos que estes valores são positivos, podendo concluir-se portanto que são municípios bastante escolarizados, por comparação com os restantes.

Ao observar estes valores percebemos o porquê de o fator população jovem escolarizada se encontrar acima do percentil 75 neste *cluster*, uma vez que, nestes 3 municípios os valores *standardizados* relativos ao número de anos de escolaridade das mulheres que foram mães e ao número médio de anos de escolaridade dos trabalhadores por conta de outrem são elevados o que é uma característica deste fator.

Fator população estrangeira pouco escolarizada

Ainda que nestes três municípios os valores *standardizados* sejam positivos relativamente à variável percentagem de população estrangeira, os valores relativos à taxa de variação do número médio de anos de escolaridade dos trabalhadores por conta de outrem são bastante elevados, o que representa um aumento ao longo dos anos relativamente à escolaridade destes trabalhadores quando comparados com os restantes municípios. Este acontecimento pode explicar o motivo de o fator população estrangeira pouco escolarizada se encontrar abaixo do percentil 25 neste *cluster*. Relembre-se que este fator é caracterizado por uma elevada percentagem de população estrangeira porém, a taxa de variação do número médio de anos de escolaridade dos trabalhadores por conta de outrem é baixa.

Concluindo, os municípios que pertencem a este *cluster* são municípios onde existe um elevado número de alojamentos turísticos por área de município, onde o ganho médio dos trabalhadores por conta de outrem é elevado e onde a população é bastante escolarizada. Estes municípios são os centros estratégicos das respetivas regiões e, por isso, é neles que os indicadores de natureza económica são mais elevados.

3.3.4 Cluster 3

Relativamente a este *cluster* para cada fator há algum dos outros que contraria o seu potencial efeito positivo no dinamismo e criação de riqueza. Em termos demográficos, os municípios deste *cluster* são menos afetados pelos efeitos da perda de população, têm uma percentagem mais elevada da população em idade ativa mas apresentam uma maior tendência de envelhecimento. Adicionalmente, têm valores acima da mediana no fator empresas e população estrangeira. No entanto, os valores são baixos nos fatores que retratam a escolaridade e os salários e, nesse sentido, pode ser considerado um *cluster* de risco.

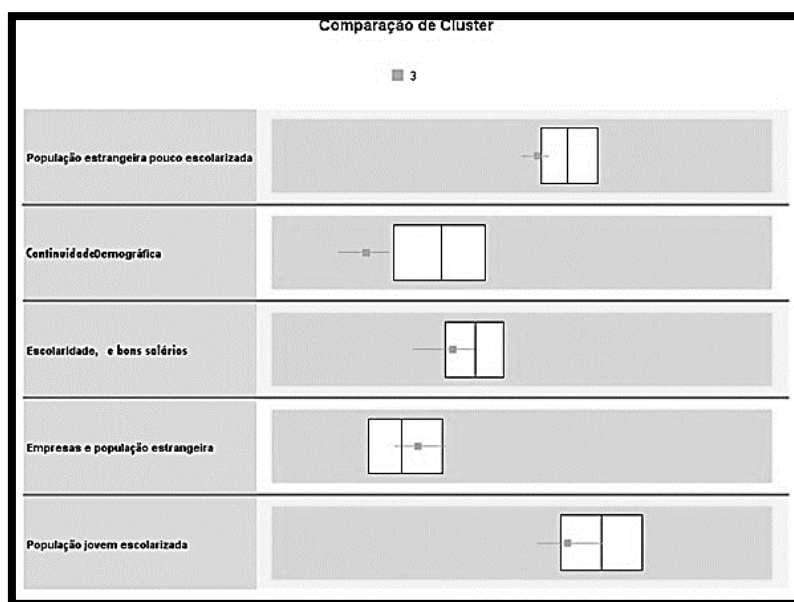


Figura 3.7 - Cluster 3 obtido em SPSS

Valores centrados no percentil 25 no fator população estrangeira pouco escolarizada. Valores abaixo do percentil 25 no fator continuidade demográfica. Valores situados entre o percentil 25 e a mediana nos fatores: Escolaridade e bons salários e População jovem escolarizada ainda que muito próximos do percentil 25. Valores situados entre a mediana e o percentil 75 no fator empresas e população estrangeira.

O município de Vizela apresenta valores em linha com o centro da distribuição de cada um dos fatores e pode, por isso, considerar-se o município de referência do *cluster* 3.

Tabela 3.11 - Município mais característico do Cluster 3

Município	Continuidade demográfica	Escolaridade e salários	Empresas e população estrangeira	População jovem escolarizada	População estrangeira pouco escolarizada
Vizela	-1,979	-1,452	0,398	-0,844	-0,822

Fator continuidade demográfica

Na variável receitas municipais por mil habitantes, ao observar os valores *standardizados* nos municípios que pertencem a este *cluster* constata-se que a grande maioria apresenta valores negativos. Existem, no entanto, municípios pertencentes a este *cluster* que apresentam valores *standardizados* positivos nesta variável. Neste sentido destacam-se os municípios: Calheta cujo valor *standardizado* é de 0,863 e o município Lajes das Flores cujo valor *standardizado* é de 1,948.

Ao observar os valores *standardizados* relativamente à variável Índice de envelhecimento nos municípios pertencentes a este *cluster* podemos concluir que a maioria dos municípios apresenta valores negativos, o que revela uma população mais jovem. Existem, no entanto, municípios pertencentes a este *cluster* que apresentam valores *standardizados* positivos nesta variável, neste sentido destaca-se principalmente o município Tabuaço cujo valor *standardizado* para essa variável é de 0,744.

Os valores *standardizados* na variável taxa média de variação anual do Índice de envelhecimento são elevados na grande maioria dos municípios pertencentes a este *cluster*. Neste sentido destacam-se os municípios Mesão Frio cujo valor *standardizado* é de 2,033, Marco de Canaveses com 2,231, Paços de Ferreira com 2,166, Santo Tirso com 2,005, Tabuaço com 2,155 e Machico com 2,280. São exceções a esta tendência os municípios Madalena cujo valor *standardizado* para essa variável é de -1,320 e Calheta com -0,092.

O mesmo sucede na variável percentagem de população em idade ativa. Destacam-se por apresentarem valores *standardizados* elevados os seguintes municípios: Vizela cujo valor *standardizado* é de 2,038, Lousada com 2,038 Porto Santo com 2,344 e Santa Cruz com 2,179. É exceção a esta tendência o município da Calheta cujo valor *standardizado* para essa variável é de -0,123, o que significa que a percentagem de população em idade ativa neste município se encontra 0,123 desvios padrões abaixo da média.

Também os valores *standardizados* na variável alojamentos familiares clássicos, por habitante, nos municípios pertencentes a este *cluster*, são na sua grande maioria negativos. Existem, no entanto, municípios pertencentes a este *cluster* que apresentam valores *standardizados* positivos nessa variável e, neste sentido destacam-se principalmente os municípios de Tarouca e Porto Santo, cujo valor *standardizado* para esta variável são 0,945 e 0,779, respetivamente.

Isto explica o porquê de o fator continuidade demográfica se encontrar abaixo do percentil 25 neste *cluster*, uma vez que uma das características deste fator é apresentar valores elevados nas receitas municipais por mil habitantes, no Índice de envelhecimento e nos alojamentos familiares clássicos por habitante e valores baixos relativamente à percentagem de população em idade ativa e à taxa de variação anual do Índice de envelhecimento.

Fator escolaridade e bons salários e fator população jovem escolarizada

Ao observar os valores *standardizados* dos municípios pertencentes a este *cluster* na variável número médio de anos de escolaridade dos trabalhadores por conta de outrem conclui-se que a maioria dos municípios apresenta valores negativos, destacando-se principalmente os municípios: Lousada cujo valor *standardizado* para essa variável é de -2,281, Marco de Canaveses com -2,054, Cinfães com -2,606, São João da Pesqueira com -2,412 e Vila Franca do Campo com -2,426. Destacam-se por serem exceções a esta tendência os municípios de Ponta Delgada cujo valor *standardizado* para essa variável é de 0,875, Gondomar com 0,723 e Lajes das Flores com 0,696.

Relativamente aos valores *standardizados* dos municípios pertencentes a este *cluster* na variável número médio de anos de escolaridade dos empregadores, os mesmos são também na sua grande maioria negativos, destacando-se os seguintes municípios: Calheta, cujo valor *standardizado* para essa variável é de -2,331, Lajes das Flores com -3,415, Ponta Delgada com -2,069, Povoação com -2,417, Santa Cruz das Flores com -2,397, Velas com -2,752 e Vila Franca do Campo com -3,112. São exceções a esta tendência os municípios de São João da Madeira cujo valor *standardizado* para essa variável é de 0,835 e Angra do Heroísmo com 0,698.

O facto de os valores *standardizados* dos municípios nestas variáveis serem na sua grande maioria negativos explica o porquê de este *cluster* apresentar valores situados entre o percentil 25 e a mediana no fator escolaridade e bons salários, uma vez que, este fator se caracteriza por valores elevados quer no número médio de anos de escolaridade dos trabalhadores por conta de outrem quer no número mé-

dio de anos de escolaridade dos empregadores. Os valores negativos relativamente à variável escolaridade dos trabalhadores por conta de outrem também podem explicar o facto de este *cluster* apresentar valores situados entre o percentil 25 e a mediana no fator população jovem escolarizada, uma vez que, valores elevados desta variável é um aspeto que caracteriza esse fator.

Fator empresas e população estrangeira

Os valores *standardizados* relativos à percentagem de população estrangeira são maioritariamente negativos nos municípios pertencentes a este *cluster*. Quanto às empresas não financeiras por mil habitantes e à taxa de variação anual das mesmas os valores são irregulares apresentando quer valores positivos quer valores negativos. Importante realçar no entanto que, neste *cluster* a mediana é de -0,2, sendo o percentil 75 muito próximo de 0. Logo, ainda que os valores sejam na sua grande maioria negativos, é importante observar que metade dos municípios pertencentes a este *cluster* apresenta valores acima de -0,2 na variável taxa de variação anual das empresas não financeiras por mil habitantes, que é a variável que apresenta maior poder explicativo neste fator. Tal justifica o porquê de este fator apresentar valores situados entre a mediana e o percentil 75 neste *cluster*.

Fator população estrangeira pouco escolarizada

Este *cluster* apresenta valores centrados no percentil 25 no fator população estrangeira pouco escolarizada devido ao facto de os valores *standardizados* relativos à variável percentagem de população estrangeira serem maioritariamente negativos, nos municípios pertencentes a este *cluster*. Também o facto de a taxa média de variação anual do Índice de envelhecimento apresentar valores *standardizados* positivos na maioria dos municípios pertencentes a este *cluster* contribui para este resultado. Neste sentido destacam-se os municípios de Mesão Frio cujo valor *standardizado* para essa variável é de 2,033, Marco de Canaveses com 2,231, Paços de Ferreira com 2,166, Santo Tirso com 2,005, Tabuaço com 2,155 e Machico com 2,280. São exceções a esta tendência os municípios: Madalena cujo valor *standardizado* corresponde a -1,320 e Calheta -0,092.

Resumindo no *Cluster 3*, encontram-se os municípios que apresentam receitas negativas, uma população em fase de acentuado envelhecimento, embora ainda com elevada percentagem em idade ativa, e pouco escolarizada.

3.3.5 Cluster 4

Cluster caracterizado pela conciliação entre a continuidade demográfica e a melhoria dos níveis de escolarização da população ativa, mas com população jovem pouco escolarizada, baixa percentagem de população em idade ativa e fraca atratividade de população estrangeira.

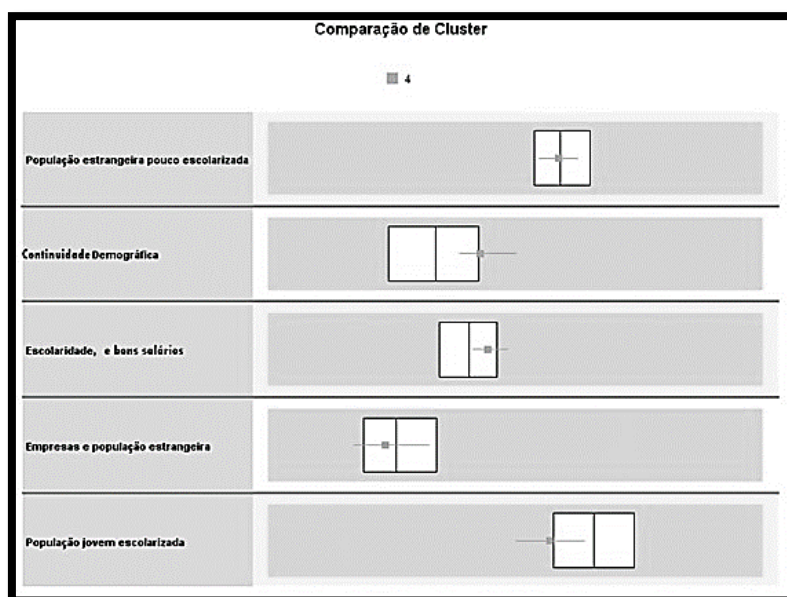


Figura 3.8 - Cluster 4 obtido em SPSS

Valores próximos à mediana relativamente ao fator população estrangeira pouco escolarizada e valores situados entre o percentil 25 e a mediana ainda que próximo da mediana no fator empresas e população estrangeira o que leva portanto a que estes fatores não sejam representativos deste *cluster*. Valores centrados no percentil 75 no fator continuidade demográfica. Valores situados entre a mediana e o percentil 75 no fator escolaridade e bons salários. Valores centrados no percentil 25 no fator população jovem escolarizada.

É o *cluster* dos municípios cujos indicadores só se alteram pelo impacto direto do avanço do tempo, sendo o município de Mora aquele que apresenta valores em linha com o centro da distribuição de cada um dos fatores e pode, por isso, considerar-se o município de referência do *cluster* 4.

Tabela 3.12 - Município mais característico do cluster 4

Município	Continuidade demográfica	Escolaridade e salários	Empresas e população estrangeira	População jovem escolarizada	População estrangeira pouco escolarizada
Mora	1,606	0,229	-0,730	-0,439	-0,656

Fator continuidade demográfica

Na variável receitas municipais por mil habitantes os valores *standardizados* são bastante elevados em alguns dos municípios pertencentes a este *cluster*, destacando-se os municípios: Corvo cujo valor *standardizado* para essa variável é de 5,482, Alcoutim com 5,134, Vila Velha de Ródão com 2,753, Crato com 2,235, Mourão com 2,249, Mértola 2,119 e Porto Moniz com 2,018.

Também no Índice de envelhecimento a grande maioria dos municípios pertencentes a este *cluster* apresentam valores *standardizados* elevados, destacando-se os municípios: Penamacor cujo valor *standardizado* para essa variável é de 3,846, Vila Velha de Ródão que apresenta o valor 4,233 e Alcoutim com 4,182. São exceções a esta tendência os municípios: Alenquer cujo valor *standardizado* para essa variável é de -0,924, Corvo com -0,867 e Nordeste que apresenta o valor -0,879.

O mesmo não sucede na variável taxa de variação anual do Índice de envelhecimento, dado que os valores *standardizados* nos municípios pertencentes a este *cluster* nesta variável são maioritariamente negativos. Ou seja, ainda que seja uma população envelhecida, este Índice de envelhecimento tem vindo a diminuir ou a aumentar menos nestes municípios que nos outros. Municípios que se destacam por apresentarem valores negativos nessa variável: Vila Velha de Ródão cujo valor *standardizado* para essa variável é de -3,382, Alter do Chão com -2,395, Gavião com -2,611, Viana do Alentejo com -2,151 e Barrancos com -2,225. Ainda relativo a esta variável acabam por se destacar, por serem exceções a esta tendência, os seguintes municípios: Sabrosa cujo valor *standardizado* é de 0,959, Espinho com 0,896 e Armamar com 0,927.

Relativamente à variável percentagem de população em idade ativa a maioria dos municípios pertencentes a este *cluster* apresentam valores negativos destacando-se os municípios: Idanha-a-Nova cujo valor *standardizado* para essa variável é de -3,130, Penamacor -2,213 e Alcoutim -3,693. Existem, no entanto, municípios pertencentes a este *cluster* que apresentam valores *standardizados* positivos nesta variável. Neste sentido destacam-se principalmente os municípios do Corvo cujo valor *standardizado* para esta variável é de 0,911, Nordeste que apresenta um valor *standardizado* de 1,122 e Santa Cruz da Graciosa com 0,911.

Também a taxa média de variação anual da densidade populacional apresenta na maioria dos municípios pertencentes a este *cluster* valores negativos, destacando-se os municípios: Idanha-a-Nova cujo valor *standardizado* para esta variável é de -2,070, Gavião com -2,948, Nisa com -2,520, Alenquer com -4,789, Mora com -2,210 e Alcoutim com -3,388. Ou seja, são municípios onde a densidade populacional tem vindo a diminuir ao longo dos anos. Existem, no entanto, municípios pertencentes a este *cluster* que apresentam valores *standardizados* positivos nesta variável, neste sentido destaca-se principalmente o município do Corvo cujo valor *standardizado* para esta variável é de 1,688.

De realçar que os municípios dos Açores que se encontram dentro deste *cluster* estão presentes em muitas das exceções, apresentando valores contraditórios relativamente ao que seria de esperar dentro deste cluster, principalmente, o município do Corvo. No entanto o facto de apresentar um valor bastante elevado em comparação com os restantes municípios na variável receitas municipais por mil habitantes faz com que este município pertença a este *cluster* dado que, essa é uma das principais características dos municípios pertencentes a este cluster.

Os dados acima explicam o facto de o fator continuidade demográfica apresentar valores centrados no percentil 75 neste *cluster*. Recorde-se que, este fator é caracterizado por valores elevados nas variáveis receitas municipais por mil habitantes e no Índice de envelhecimento e valores baixos relativamente às variáveis percentagem de população em idade ativa, taxa de variação de densidade populacional e taxa de variação anual do Índice de envelhecimento.

Fator escolaridade e bons salários

Ainda que a grande maioria dos valores *standardizados* dos municípios pertencentes a este *cluster* 4 para a variável número médio de anos de escolaridade dos trabalhadores por conta de outrem e dos empregadores sejam negativos, assim como os valores relativos à variável alojamentos turísticos por área de município e ganho médio por conta de outrem, o fator escolaridade e bons salários apresenta valores situados entre a mediana e o percentil 75 neste *cluster*. Tal, deve-se ao facto de os municípios pertencentes a este *cluster* apresentarem valores *standardizados* negativos na variável percentagem de empresas criadas no setor Indústria, Construção e Energia, o que é uma característica do fator escola-

ridade e bons salários e é aquela que apresenta um maior valor absoluto de carga fatorial dentro deste fator na matriz de componentes. Neste sentido, destacam-se os municípios: Monforte -2,411 e Corvo -2,411. Existem, porém, municípios pertencentes a este *cluster* que apresentam valores *standardizados* positivos nesta variável, destacando-se principalmente os municípios: Penalva do Castelo cujo valor *standardizado* para essa variável é de 0,971, Fornos de Algodres com 1,112, Arganil com 0,768, Oliveira do Hospital com 0,700, Crato com 0,841, Nisa com 0,771, Aljustrel com 0,856 e Santa Cruz da Graciosa com 1,166.

Fator empresas e população estrangeira

Relativamente ao fator empresas e população estrangeira, o mesmo apresenta valores situados entre o percentil 25 e a mediana, ainda que próximo da mediana, neste *cluster*, principalmente devido ao facto de os municípios apresentarem valores *standardizados* negativos na variável percentagem de população estrangeira. Ferreira do Alentejo, no entanto destaca-se por ser uma exceção a esta tendência apresentando um valor *standardizado* para essa variável de 1,058.

Relativamente às restantes variáveis que caracterizam este fator, ou seja, a variável empresas não financeiras por mil habitantes e a sua respetiva variação apresentam valores *standardizados* contraditórios sendo eles positivos em alguns municípios e negativos noutros.

Fator população jovem escolarizada

Quanto ao fator população jovem escolarizada, o mesmo apresenta valores centrados no percentil 25 neste *cluster*. Tal deve-se ao facto de os valores *standardizados* dos municípios pertencentes a este *cluster* na variável número médio de anos de escolaridade dos trabalhadores por conta de outrem serem na sua grande maioria negativos tal como já tinha mencionado anteriormente, o mesmo sucedendo com o número de anos de escolaridade das mulheres que foram mães. Destacam-se por apresentarem valores negativos na variável número de anos de escolaridade das mulheres que foram mães os seguintes municípios: Penamacor cujo valor *standardizado* nessa variável é de -2,783, Avis com -2,597, Monforte com -3,216, Sousel com -2,088, Mourão com -5,922, Moura com -2,584, Vidigueira com -2,606 e Porto Moniz -3,161. Existem, no entanto, municípios pertencentes a este *cluster* que apresentam valores positivos nesta variável, destacando-se os municípios: Soure cujo valor *standardizado* nessa variável é de 1,163, Covilhã com 1,464, Golegã com 0,861, Castelo de Vide com 1,273, Portalegre com 1,166 e Vila Viçosa com 0,791.

Relativamente ao número médio de anos de escolaridade dos trabalhadores por conta de outrem, tal como mencionado anteriormente, a grande maioria dos municípios pertencentes a este *cluster* apresentam valores negativos nesta variável, destacando-se o município Santa Cruz da Graciosa cujo valor *standardizado* nessa variável é de -2,056. Existem, no entanto, municípios que contrariam essa tendência apresentando valores positivos nessa variável, destacando-se os municípios: Vila Velha de Ródão cujo valor *standardizado* nessa variável é de 1,506, Nelas com 0,737, Covilhã com 0,980, Fundão com 0,755, Abrantes com 1,032, Campo Maior com 0,807, Elvas com 0,727, Portalegre com 0,870.

Fator população estrangeira pouco escolarizada

O fator população estrangeira pouco escolarizada apresenta valores situados entre o percentil 25 e a mediana ainda que próximo da mediana neste *cluster*, uma vez que, ainda que as variáveis percentagem de população estrangeira e taxa média de variação anual da densidade populacional, apresentem na maioria dos municípios pertencentes a este *cluster* um valor *standardizado* negativo, o mesmo ten-

de a ser próximo de 0. Também os valores relativos à taxa média de variação anual do Índice de envelhecimento são na sua grande maioria negativos nos municípios pertencentes a este cluster.

Resumindo, municípios que pertençam a este *cluster* são municípios onde a taxa média de variação anual da densidade populacional apresenta valores negativos, ou seja, a densidade populacional tem vindo a diminuir mais nesses municípios e onde o Índice de envelhecimento é elevado quando comparado com os restantes municípios. Apresenta ainda valores relativos à percentagem de população em idade ativa baixos e receitas elevadas. Também o número de anos de escolaridade das mulheres que foram mães apresentam valores muito baixos quando comparados com os restantes municípios.

3.3.6 Cluster 5

Cluster caracterizado por uma aposta na mão-de-obra estrangeira e/ou pouco investimento na melhoria da escolarização.

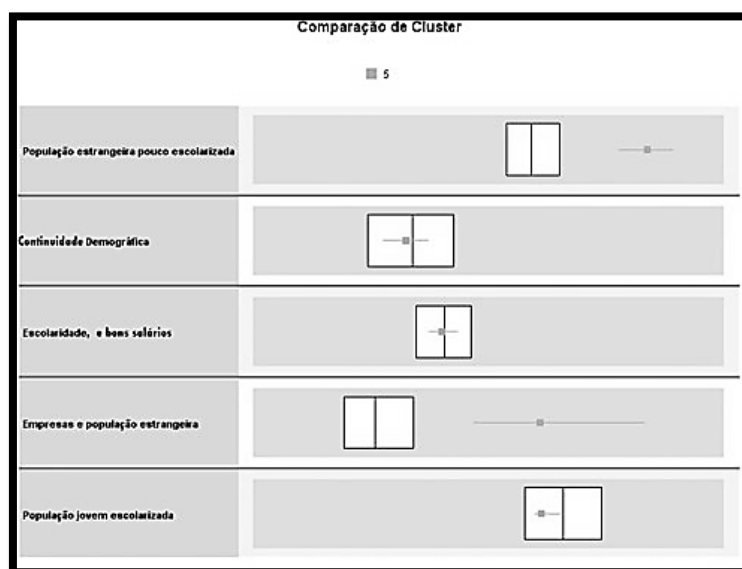


Figura 3.9 - Cluster 5 obtido em SPSS

Valores acima do percentil 75 no fator população estrangeira pouco escolarizada e no fator empresas e população estrangeira. Valores próximos da mediana no fator escolaridade e bons salários e no fator continuidade demográfica, ainda que neste último os valores estejam um pouco abaixo da mediana. Valores situados entre o percentil 25 e a mediana no fator população jovem escolarizada.

O município de Albufeira apresenta valores em linha com o centro da distribuição de cada um dos fatores e pode, por isso, considerar-se o município de referência do *cluster* 5.

Tabela 3.13 - Município mais característico do Cluster 5

Município	Continuidade demográfica	Escolaridade e salários	Empresas e população estrangeira	População jovem escolarizada	População estrangeira pouco escolarizada
Albufeira	-0,906	0,462	4,419	-0,398	2,943

Fator continuidade demográfica

Olhando para as variáveis que mais caracterizam este fator, constata-se que para o Índice de envelhecimento, os valores *standardizados* nos municípios pertencentes a este *cluster* são baixos, sendo praticamente todos negativos, ainda que não muito afastados da média. São exceções a esta tendência os seguintes municípios: Odemira cujo valor *standardizado* para essa variável é de 0,156, Albufeira com 0,162 e Aljezur com 0,258.

Tal pode explicar o facto de este *cluster* apresentar valores próximos da mediana no fator continuidade demográfica ainda que um pouco abaixo da mediana. Isso juntamente com o facto de a taxa média de variação anual de densidade populacional apresentar valores *standardizados* positivos exceto no município Odemira. No entanto, as restantes características representantes deste fator estão em concordância com o que seria esperado dentro do mesmo, ou seja, os valores correspondentes ao número de alojamentos familiares clássicos por habitante, são na sua maioria positivos, o mesmo sucede com as receitas municipais por mil habitantes, e por sua vez os valores relativos à taxa média de variação anual do Índice de envelhecimento apresentam na grande maioria dos municípios pertencentes a este *cluster* valores negativos. Não é possível no entanto retirar padrões muito conclusivos relativamente à variável percentagem de população em idade ativa. Estes dados permitem perceber o porquê de que ainda que este fator se encontre abaixo da mediana neste *cluster*, este não esteja muito afastado da mesma.

Fator Escolaridade e bons salários

De uma forma geral, o ganho médio dos trabalhadores por conta de outrem, o número médio dos trabalhadores e empregadores, número de alojamentos turísticos por área de município e a percentagem de empresas criadas no setor Indústria, Construção e Energia não se encontram muito afastados da média o que justifica o facto de o fator escolaridade e bons salários não ser fortemente característico deste *cluster*, apresentando valores próximos da mediana.

Fator empresas e população estrangeira

Os valores *standardizados* relativamente à variável empresas não financeiras por mil habitantes são na sua grande maioria positivos nos municípios pertencentes a este *cluster*, destacando-se principalmente os municípios: Albufeira cujo valor *standardizado* é de 2,192 e Aljezur com 2,073. São exceções a esta tendência os municípios: Alvito cujo valor *standardizado* é de -0,564, Castro Marim com -0,044 e Olhão com -0,080.

O mesmo sucede com a variável taxa de variação do número de empresas não financeiras por mil habitantes. Os valores *standardizados* nos municípios pertencentes a este *cluster* são positivos, destacando-se os municípios: Aljezur cujo valor *standardizado* para essa variável é de 2,852, Lagoa com 2,824, Lagos com 2,690 e Tavira com 3,227. São portanto, municípios onde o número de empresas não financeiras tem vindo a aumentar mais ao longo dos anos quando comparado com os restantes municípios. É exceção a esta tendência o município Vila Real de Santo António em que esse valor é de -0,139.

Para além disso, os valores *standardizados* relativos à variável percentagem de população estrangeira são todos positivos nos municípios pertencentes a este *cluster* apresentando valores bastante elevados em alguns dos municípios. Neste sentido destacam-se principalmente os municípios: Odemira cujo valor *standardizado* é de 4,970, Albufeira com 6,189, Aljezur com 4,763, Lagoa com 3,474, Lagos com 5,407, Loulé com 4,257, Portimão com 2,738, Silves com 2,508, Tavira com 4,049, Vila do Bispo

com 6,212, Vila Real de Santo António com 2,232. Os valores elevados nestas variáveis explicam o facto de o fator empresas e população estrangeira apresentar valores acima do percentil 75 neste *cluster*, dado que, este fator caracteriza-se por valores elevados nessas mesmas variáveis.

Fator população jovem escolarizada

Relativamente aos valores *standardizados* relativos ao número de anos de escolaridade das mulheres que foram mães nos municípios pertencentes a este *cluster* são maioritariamente negativos. O que pode explicar o facto de o fator população jovem escolarizada apresentar valores situados entre o percentil 25 e a mediana neste *cluster* ainda que a outra variável caracterizante deste fator, o número médio de anos de escolaridade dos trabalhadores por conta de outrem, apresente valores *standardizados* positivos nos municípios pertencentes a este cluster.

Fator população estrangeira pouco escolarizada

É este o *cluster* que enquadra os municípios com valores mais elevados no fator população estrangeira pouco escolarizada. Ou seja, para a maioria dos municípios deste *cluster*, os valores da percentagem de população estrangeira são elevados e os da densidade populacional tem vindo a aumentar ao longo dos anos. São ainda municípios que apresentam uma variação negativa relativamente ao Índice de envelhecimento e ao número médio de anos de escolaridade dos trabalhadores por conta de outrem, ou seja, municípios onde o Índice de envelhecimento e o número de anos de escolaridades dos trabalhadores por conta de outrem ou têm diminuído ou não têm vindo a aumentar tanto como nos restantes municípios do país.

Concluindo, no *cluster* 5 encontram-se municípios com uma elevada percentagem de população estrangeira com um nível de escolaridade que não acompanha as melhorias verificadas a nível nacional. Estes municípios são ainda caracterizados por uma população jovem, onde existe um elevado número de empresas não financeiras e onde esse número apresentou um elevado crescimento desde 2013.

3.3.7 Cluster 6

Cluster caracterizado por uma aposta na melhoria da escolarização mas pouco dinamismo na criação de empresas.

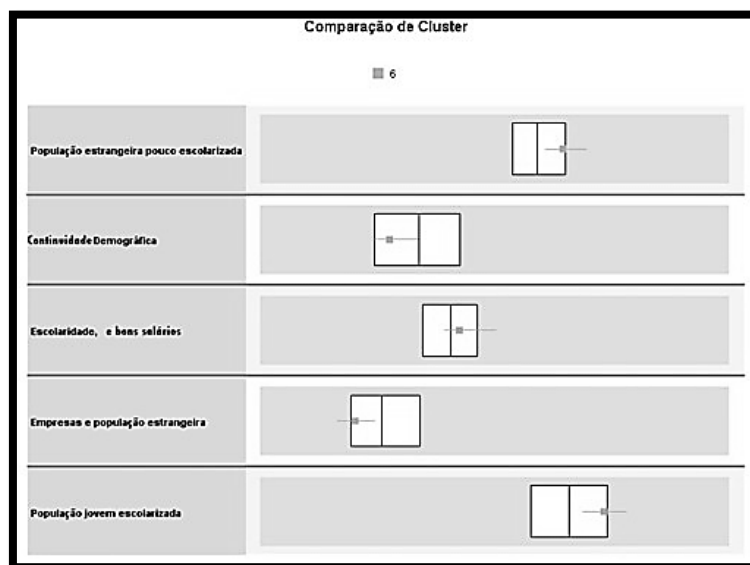


Figura 3.10 - Cluster 6 obtido em SPSS

Valores centrados no percentil 75 no fator população estrangeira pouco escolarizada e população jovem escolarizada. Valores centrados no percentil 25 no fator empresas e população estrangeira. Valores situados entre o percentil 25 e a mediana no fator continuidade demográfica. Valores situados entre a mediana e o percentil 75 no fator escolaridade e bons salários ainda que próximo da mediana.

O município de Condeixa-a-Nova apresenta valores em linha com o centro da distribuição de cada um dos factores e pode, por isso, considerar-se o município de referência do *cluster* 6.

Tabela 3.14 - Município mais característico do Cluster 6

Município	Continuidade demográfica	Escolaridade e salários	Empresas e população estrangeira	População jovem escolarizada	População estrangeira pouco escolarizada
Condeixa-a-Nova	-0,656	0,294	-0,710	1,044	0,576

Fator continuidade demográfica

Ao observar os valores *standardizados* das receitas municipais por mil habitantes nos municípios pertencentes a este *cluster* verificamos que esses valores são maioritariamente negativos. Tal não se verifica somente nos seguintes municípios: Vila Nova de Cerveira cujo valor *standardizado* para essa variável é de 0,536, Guarda com 0,079, Óbidos com 0,182, Constância com 0,710 e Castro Verde com 0,410.

O mesmo sucede com os dados relativos ao Índice de envelhecimento: a grande maioria dos municípios pertencentes a este *cluster* apresenta valores *standardizados* negativos nesta variável (no entanto não se encontram muito afastados da média, ou seja, não são valores muito discrepantes).

Os valores *standardizados* relativos à variável taxa média de variação anual densidade populacional são elevados na grande maioria dos municípios pertencentes a este *cluster*, o que indica que a densidade populacional nestes municípios tem vindo a aumentar mais ao longo dos anos (ou a diminuir menos) quando comparado com os restantes municípios, destacando-se os municípios: Entroncamento cujo valor *standardizado* para essa variável é de 2,106, Amadora com 2,020, Arruda dos Vinhos com 3,040, Mafra com 2,519, Odivelas com 2,802, Alcochete com 2,841 e Montijo com 2,720.

Quanto aos alojamentos familiares clássicos, por habitante, os valores *standardizados* nos municípios pertencentes a este cluster relativamente a esta variável são na sua grande maioria negativos.

As observações descritas acima explicam o porquê de o fator continuidade demográfica apresentar valores situados entre o percentil 25 e a mediana neste *cluster*, dado que uma das características deste fator prende-se com um valor elevado de receitas municipais por mil habitantes, de Índice de envelhecimento, de alojamentos familiares clássicos por habitante e uma taxa de variação anual de densidade baixa, ou seja, municípios onde a densidade populacional tem vindo a diminuir ao longo do tempo.

Fator escolaridade e bons salários

O valor relativo ao número médio de anos de escolaridade dos trabalhadores por conta de outrem, número médio de anos de escolaridade dos empregadores e ganho médio dos trabalhadores por conta de outrem são na maioria dos municípios pertencentes a este *cluster* positivos o que pode explicar o

porquê de o fator escolaridade e bons salários se encontrar ligeiramente acima da mediana. Ainda assim, os valores apresentados por este fator encontram-se próximos da mediana. Isto pode dever-se ao facto de o número de alojamentos turísticos por área de município, outra característica deste fator, apresentar valores *standardizados* negativos na maioria dos municípios pertencentes a este *cluster*.

Quanto à variável número médio de anos de escolaridade dos empregadores destacam-se principalmente os municípios: Oeiras cujo valor *standardizado* para essa variável é de 2,649, Cascais com 1,930 e Constância com 1,852. Existem, no entanto, municípios que apresentam valores negativos nesta variável destacando-se principalmente os municípios: Vila Nova de Poiares cujo valor *standardizado* é de -1,039, Pombal com -0,810, Porto de Mós com -0,713, Salvaterra de Magos com -0,869, Cadaval com -0,766 e Vila do Porto -0,734.

Relativamente à variável número médio de anos de escolaridade dos trabalhadores destacam-se principalmente os municípios: Vila Real cujo valor *standardizado* para essa variável é de 2,112, Coimbra com 2,573, Entroncamento com 2,131, Amadora com 2,285, Cascais com 2,306, Oeiras com 3,397 e Almada com 2,009. Existem, no entanto, municípios que apresentam valores negativos nesta variável destacando-se principalmente os municípios: Sobral de Monte Agraço cujo valor *standardizado* é de -0,799, Cadaval com -0,807, Salvaterra de Magos com -0,701.

No que diz respeito à variável ganho médio dos trabalhadores por conta de outrem, os valores *standardizados* relativos aos municípios pertencentes a este *cluster* são na sua grande maioria elevados quando comparados com os restantes municípios. Destacam-se os seguintes municípios: Vila do Porto cujo valor *standardizado* para essa variável é de 5,442, Castro Verde com 4,953, Sines com 4,422, Palmela com 2,121, Alcochete com 6,801, Oeiras com 4,721 e Amadora com 2,400.

Fator empresas e população estrangeira

Relativamente aos valores *standardizados* das variáveis empresas não financeiras por mil habitantes e respetiva taxa de variação nos municípios pertencentes a este *cluster* eles são baixos. Quanto à variável empresas não financeiras por mil habitantes existem, no entanto, municípios que não seguem esta tendência, destacando-se nesse sentido o município: Óbidos cujo valor *standardizado* é de 1,016.

Estes valores podem explicar o motivo de o fator empresas e população estrangeira apresentar valores centrados no percentil 25 neste *cluster*, uma vez que, este fator é caracterizado por valores elevados relativamente a essas variáveis.

Fator população jovem escolarizada

Quanto ao número de anos de escolaridade das mulheres que foram mães, os valores *standardizados* desta variável nos municípios pertencentes a este *cluster* são na sua grande maioria elevados, o mesmo sucedendo com o número médio de anos de escolaridade dos trabalhadores por conta de outrem. Isto pode explicar o porquê de o fator população jovem escolarizada apresentar valores centrados no percentil 75 neste *cluster*, uma vez que este fator se caracteriza por valores elevados nestas variáveis.

Quanto à variável anos de escolaridade das mulheres que foram mães destaca-se, por apresentar um valor especialmente elevado, o município Coimbra cujo valor *standardizado* para essa variável é de 2,198. Existem, no entanto, municípios que apresentam valores *standardizados* negativos nesta variável.

vel. Neste sentido destacam-se os municípios: Murtosa cujo valor *standardizado* para essa variável é de -0,821, Bombarral com -0,719, Montijo com -0,883 e Vila do Porto com -0,957.

Relativamente à variável número médio de anos de escolaridade dos trabalhadores por conta de outrem, a grande maioria dos municípios pertencentes a este *cluster* apresentam valores *standardizados* positivos, tal como foi mencionado anteriormente.

Fator população estrangeira pouco escolarizada

A taxa média de variação do número médio de anos de escolaridade dos trabalhadores por conta de outrem apresenta valores *standardizados* abaixo da média na maioria dos municípios pertencentes a este *cluster*, destacando-se principalmente o município Oeiras cujo valor *standardizado* é de -2,592. De realçar que ainda que este município apresente um valor elevado relativamente à variável número médio de anos de escolaridade dos trabalhadores por conta de outrem o mesmo foi o que menos aumentou ao longo dos anos por comparação com os restantes municípios. Existem, no entanto, municípios que apresentam valores *standardizados* positivos nesta variável. Neste sentido destacam-se os municípios: Óbidos cujo valor *standardizado* para essa variável é de 0,740 e Castro Verde com 1,558. Estes valores juntamente com o facto da variação anual de densidade populacional apresentar valores *standardizados* elevados na maioria dos municípios pertencentes a este *cluster* contribuem para explicar o facto de neste *cluster* o fator população estrangeira pouco escolarizada apresentar valores centrados no percentil 75. Ainda que a percentagem de população estrangeira apresente maioritariamente valores negativos, os mesmos não apresentam um valor muito afastado da média pelo que acabam por não influenciar este resultado. Destacam-se por apresentarem valores elevados na variável taxa média de variação anual da densidade populacional os municípios: Entroncamento cujo valor *standardizado* para essa variável é de 2,106, Amadora com 2,020, Arruda dos Vinhos com 3,040, Mafra com 2,519, Odivelas com 2,802, Alcochete com 2,841 e Montijo com 2,720.

Resumindo, neste *cluster* estão municípios onde a população é escolarizada, as receitas municipais por mil habitantes são baixas, onde existe um baixo número de empresas por habitante e, são ainda municípios onde a densidade populacional tem vindo a aumentar ou, pelo menos a diminuir menos que nos demais municípios e onde os salários são muitos elevados.

De realçar que o facto de existir um baixo número de empresas por habitante, por si só não indica uma situação de risco, pois podem existir grandes empresas que acabam por empregar um grande número de habitantes do município em questão. Posto isto, recorreu-se à observação dos indicadores: Número médio de pessoas ao serviço nas empresas não financeiras e a percentagem de pessoal ao serviço das quatro maiores empresas não financeiras em alguns dos municípios pertencentes ao *cluster* 6. Constatou-se que, de facto essa situação se verifica. Considerando o município Vila Nova de Cerveira, verificou-se que este é o 5º município com o maior número médio de pessoas ao serviço nas empresas não financeiras, sendo esse valor de 4,9 (quando a média nacional é 3,2) e é ainda o 6º a apresentar uma maior percentagem de pessoal ao serviço das quatro maiores empresas não financeiras, correspondendo esse valor a 39,3%. O mesmo sucede com outros municípios pertencentes a este *cluster* como é o caso de Matosinhos e Mangualde que apresentam no indicador número médio de pessoas ao serviço nas empresas não financeiras o valor de 4,8 e 4,2, respetivamente, e apresentam ainda valores bastante elevados no que diz respeito à percentagem de pessoal ao serviço das quatro maiores empresas não financeiras, sendo esse valor de 35,1% em Matosinhos e de 38% em Mangualde.

CAPÍTULO 4

CONCLUSÃO

Após o desenvolvimento deste estudo foi possível concluir que as técnicas estatísticas que foram utilizadas demonstraram ser capazes de fazer uma separação razoavelmente eficaz dos municípios portugueses, de acordo com as suas características, permitindo desta forma traçar um perfil dos mesmos e fazer uma análise relativamente ao seu desenvolvimento. Foi possível ainda concluir que os municípios que se localizam no litoral de Portugal são mais desenvolvidos que aqueles que se encontram no interior do país, tal como seria de esperar. De notar que os municípios que apresentam os salários mais elevados se encontram no *cluster* 6, que se localizam maioritariamente no litoral de Portugal, zona que, tal como foi mencionado anteriormente é mais desenvolvida e também mais escolarizada. Estas características também estão presentes nos municípios pertencentes ao *cluster* 2.

Desta forma, para que seja possível colmatar esta diferença existente entre o litoral de Portugal e o interior, deveria existir uma maior aposta no interior, quer através da criação de emprego, quer através da melhoria dos serviços públicos a nível de saúde, educação e transportes. De realçar, no entanto, que já é de notar uma ligeira diferença relativamente a essa realidade, pois, nos municípios localizados no interior do país pertencentes ao *cluster* 1 e 4, é de notar que ainda que o Índice de envelhecimento seja tendencialmente elevado e a percentagem de população em idade ativa seja baixa, a variação relativamente ao Índice de envelhecimento apresentou tendencialmente uma variação negativa ao longo dos últimos anos, o que significa que o Índice de envelhecimento nestes municípios tem vindo a diminuir mais que nos restantes. O que demonstra uma tendência de mudança relativamente à realidade mencionada acima, tendência esta que, tudo indica, se irá manter, podendo significar que futuramente estas zonas mais rurais venham a ser efetivamente mais desenvolvidas em diversos setores. Tal pode dever-se a diversos fatores, podendo o facto de ultimamente existir uma maior aposta no turismo rural ser um deles. Segundo o grupo de consultoria, Mercal, “O turismo rural em Portugal registou um substancial crescimento em 2017 e tem sido uma atividade que tem registado um elevado crescimento e com excelentes perspetivas futuras” tendência esta que se manteve. Porém, em 2020 dada a situação pandémica em que vivemos nos dias que correm, diversos setores foram seriamente afetados, tendo sido o turismo um daqueles que mais sofreu as suas consequências. Ainda assim, dentro do grande setor que é o turismo, o turismo rural conseguiu destacar-se pela positiva quando comparado com os restantes, dado que esta modalidade de turismo se caracteriza principalmente por um distanciamento da vida cidadina, apelando a um contacto com a natureza, longe das grandes confusões. Posto isto, dada a necessidade de distanciamento social, as pessoas acabaram por optar por estes estabelecimentos ao invés de outros, tal como foi mencionado no artigo do jornal digital, dinheiro vivo, que menciona que “As dormidas em alojamentos de turismo rural e de habitação caíram 22,7% sendo que os restantes estabelecimentos apresentaram descidas acima dos 50%”.

Para além disso, existem ainda diversas medidas que foram implementadas pelo governo e incentivos do Estado com o propósito de mudar esta realidade e tornar a zona interior do país mais rejuvenescida e desenvolvida relacionadas com o turismo rural. São exemplos dessas medidas a SI2E², Linha de

² O SI2E pretende estimular o surgimento de iniciativas empresariais e a criação de emprego em territórios de baixa densidade e por essa via promover o desenvolvimento e a coesão económica e social do país. Não se aplica exclusivamente aos territórios de baixa densidade, o SI2E favorece através de majorações específicas os investimentos nelas realizados e sobretudo cria condições para uma maior dinâmica empresarial ao ajustar tipologias de projetos às condições reais das micro e pequenas empresas do interior. Fonte: Portal dos incentivos.

Apoio à Valorização Turística do Interior³ e o Programa de Investimento em Territórios de Baixa Densidade⁴.

Também o facto de existir uma maior aposta no setor da agricultura ao longo dos últimos anos veio a influenciar esta mudança. Segundo o artigo, “É na agricultura que há maior crescimento de empresas”, publicado pelo Diário de Notícias, Portugal tem apresentado um maior número de criação de empresas, principalmente no setor da agricultura e pescas. Segundo o Banco de Portugal, entre 2013 e 2017, o número de empresas nesse setor cresceu mais de 31%. Concluindo que, o ritmo de criação de empresas na agricultura e pescas é superior em mais de quatro vezes ao das restantes atividades económicas. A aposta que se tem verificado ao longo dos últimos anos neste setor, também pode ter sido influenciada por determinadas medidas que foram implementadas pelo Estado, medidas estas que têm como objetivo incentivar a investimentos neste ramo, como por exemplo o PDR⁵.

Outro fator importante que permite corroborar a teoria de que está a existir uma mudança nos municípios mais rurais, é o facto de que os valores relativos à variação do número médio de anos de escolaridade dos trabalhadores sofreram uma maior variação, em termos positivos, exatamente em municípios pertencentes aos *clusters* 1, 3 e 4, *clusters* estes que são constituídos maioritariamente por municípios que se localizam no interior do país.

Esta mudança pode ainda ser justificada com o facto de a percentagem de população estrangeira residente em Portugal ter vindo a aumentar consideravelmente nos últimos anos. Segundo um artigo publicado pela República Portuguesa “A população estrangeira residente em Portugal aumentou em 2019 pelo quarto ano consecutivo, totalizando 590 348 cidadãos, o valor mais elevado registado pelo Serviço de Estrangeiros e Fronteiras desde o seu surgimento, em 1976”. Esta tendência teve como consequência uma valorização do mercado imobiliário pois, muitos dos investimentos que são feitos por estrangeiros neste ramo, são bastante elevados, fator que pode também estar relacionado com a existência do Visto Gold. O Visto Gold ou Golden Visa é uma autorização de residência concedida a cidadãos nacionais de Estados Terceiros atribuída mediante a realização de investimento, empresarial ou imobiliário, em território português. Este visto pode ser concedido em caso de aquisição de bens imóveis de valor igual ou superior a 500 mil euros. Como consequência, deu-se um aumento quase exponencial dos preços das casas nos últimos anos, resultado consequentemente de uma maior procura pelas mesmas, especialmente na zona de Lisboa. Segundo um artigo publicado em julho pelo Diário de Notícias, o preço das casas em Lisboa aumentou mais de 55% nos últimos três anos. Esta realidade tem levado a população mais jovem a abdicar da vida citadina lisboeta e a deslocar-se para zonas onde o preço das casas é mais acessível, interior do país inclusive. Observando os dados relativos ao valor médio de avaliação bancária dos alojamentos disponibilizados pela PORDATA, presentes na tabela abaixo, é possível observar que estes valores são mais elevados nos municípios pertencentes ao *cluster*

³ A Linha de Apoio à Valorização Turística do Interior tem como objetivo apoiar o investimento em projetos e iniciativas que, por meio de atividades com relevância ou interesse para o turismo, contribuam para o desenvolvimento do interior e para a coesão económica e social do território. Fonte: Turismo de Portugal.

⁴ Lançado pelo Governo, através da Turismo Fundos, o programa tem como objetivo apoiar o investimento na aquisição de imóveis, nas regiões abrangidas pelo Programa Nacional para a Coesão Territorial, para desenvolvimento de atividades turísticas nos territórios de baixa densidade. Fonte: Consulstaff

⁵ O PDR 2020, Programa de Desenvolvimento Rural de Portugal - Continente (2014 - 2020) foi aprovado na sequência da decisão da Comissão Europeia - Decisão C (2014) 9896 final de 12 de dezembro de 2014 e visa essencialmente apoiar as atividades deste setor e na produção de bens transacionáveis dirigidas a agentes diretamente envolvidos na criação de valor a partir de atividades agroflorestais assente numa gestão eficiente dos recursos. Fonte: Portal dos Incentivos

2, caracterizado pelas grandes metrópoles, e ainda nos municípios pertencentes ao distrito de Lisboa, especialmente nos arredores da cidade. Estes valores são ainda elevados em alguns dos municípios pertencentes ao *cluster* 5, caracterizado especialmente por apresentar uma elevada percentagem de população estrangeira.

Tabela 4.1 - Os 10 Municípios onde o valor médio de avaliação bancária dos alojamentos é mais elevado

Municípios	Distritos	População Es-trangeira (%)	Alojamentos turísticos por área de município	Valor médio de avaliação bancária dos alojamentos
Porto	Porto	5,40%	685,37%	1804
Cascais	Lisboa	11,90%	84,54%	2039
Lisboa	Lisboa	15,50%	576,00%	2300
Loures	Lisboa	7,90%	2,40%	1691
Odivelas	Lisboa	9,00%	0,00%	1799
Oeiras	Lisboa	5,60%	23,91%	2129
Loulé	Faro	21,70%	13,35%	1765
Albufeira	Faro	30,10%	148,23%	1790
Lagos	Faro	26,70%	46,95%	1704
Tavira	Faro	20,80%	6,26%	1648

Fonte: PORDATA (elaboração própria)

De realçar, que este aumento que se tem verificado relativamente à percentagem de população estrangeira, poderá também estar relacionado com o aumento do número de refugiados que tem ocorrido ao longo dos últimos anos em diversas partes do globo, Portugal inclusive. Sendo esta crise já considerada pela ONU como a crise humanitária mais intensa do século. Na tabela 4.2 é possível observar os números relativos a esta realidade acima mencionada:

Tabela 4.2 - Refugiados em Portugal, entre 2013 e 2018

Ano	Refugiados	
	N	Taxa de Variação anual
2013	598	
2014	699	+16,9%
2015	699	0
2016	1194	+70,8%
2017	1623	+35,9%
2018	2136	+31,9%

Fonte - UNHCR – Global Trends 2018, sistematização e cálculos por Oliveira, C. (2020)

Relativamente ao *cluster 3*, este *cluster* é constituído essencialmente por municípios a norte de Portugal, na zona Minho e Douro litoral, e ainda pela maioria dos municípios dos arquipélagos dos Açores e da Madeira. Nestes municípios o principal fator de relevo e que os diferencia daqueles que pertencem aos *clusters 1* e *4*, prende-se essencialmente com o facto de a percentagem de população em idade ativa apresentar valores bastante elevados e de a variação anual do Índice de envelhecimento ter apresentado uma variação mais positiva que nos demais municípios, o que representa um aumento relativamente ao Índice de envelhecimento ao longo dos últimos anos neste municípios, contrariando a tendência de diminuição deste Índice que se tem vindo a notar nos municípios *1* e *4*. Prevendo-se desta forma um envelhecimento da população nesses municípios, o que pode conferir uma situação de risco. Focando nas regiões autónomas, este facto já foi considerado pela Direção Regional de Estatística da Madeira⁶, que, num estudo que desenvolveu, aponta para o envelhecimento da população madeirense.

Concluiu-se também que as grandes metrópoles, correspondendo aos municípios do *cluster 2*, sendo consequentemente os mais desenvolvidos, têm uma característica essencial em comum que se prende principalmente com o turismo. Conclusões retiradas através da observação do indicador relativo ao número de alojamentos turísticos por área de município, que é bastante elevado nos três municípios pertencentes ao *cluster 2*. De realçar que, tal como foi mencionado pela página Turismo de Portugal “O setor do turismo é a maior atividade económica exportadora do país, tendo sido, em 2019, responsável por 52,3% das exportações de serviços e por 19,7% das exportações totais. As receitas turísticas registaram um contributo de 8,7% para o PIB nacional”. Também os níveis de escolarização e salários são elevados nestes municípios, tal como seria de esperar.

Também como seria de esperar os municípios localizados no Algarve e também na Costa Vicentina, presentes no *cluster 5*, apresentam uma elevada percentagem de população estrangeira, pois são locais bastante atrativos a nível turístico e que permitem aos imigrantes desempenhar funções consideradas relegadas pelos autóctones, permitindo ainda desenvolver negócios sustentáveis relacionados com um ambiente mais rural que possuem um papel importante na economia nacional, regional e local através da criação de emprego, promoção da sustentabilidade, permitindo gerar de forma indireta receitas e crescimento de diversos setores de atividade.

Por fim, concluiu-se que o principal fator de risco, para os 308 municípios envolvidos neste estudo, prende-se essencialmente com o envelhecimento da população, que acaba por influenciar negativamente o desenvolvimento destes municípios e consequentemente de Portugal. Também uma maior aposta no setor das tecnologias seria necessário dado que é um fator bastante diferenciador nos dias que correm e que caracterizam as grandes economias dos dias de hoje.

⁶ Relatório de Atividades 2016. Direção Regional de Estatística da Madeira. Julho 2017.

BIBLIOGRAFIA

Alberto, D., & Ferreira, J. J. Competitividade Regional: Conceito, Instrumentos e Modelos de Avaliação. Tomar, 2008. Disponível em: <https://core.ac.uk/download/pdf/62717753.pdf>.

Consultado a 6 de setembro de 2020.

Amaral, E. (2011). Apontamentos das aulas teóricas de Correlação e Análise Fatorial. Universidade Federal de Minas Gerais, Brasil. Disponível em: <http://www.ernestoamaral.com/docs/dcp046-111/Aula28.pdf>.

Consultado a 7 de abril de 2020.

Bacher, J., Wenzig, K., & Vogler, M. (2004). SPSS Two Step Cluster - a first evaluation. (2., corr. ed.) (Arbeitsund Diskussionspapiere / Universität Erlangen-Nürnberg, Sozialwissenschaftliches Institut, Lehrstuhl für Soziologie, 2004-2).

Bação, F., Painho, M., & Catarino, S. Novas Metodologias na Análise de Clusters Aplicada a Estudos Geográficos. Instituto Superior de Estatística e Gestão de Informação da Universidade Nova de Lisboa, Portugal.

Barnett, V., Lewis, T. (1994). Outliers in statistical data. New York: John Wiley & Sons.

Barroso, R. É na Agricultura que há maior crescimento de empresas. Diário de Notícias, 06 de fevereiro de 2019. Disponível em: <https://www.dn.pt/edicao-do-dia/06-fev-2019/e-na-agricultura-que-ha-maior-crescimento-de-empresas-10539894.html>.

Consultado a 29 de outubro de 2020.

Basílio, E. (2018). Os Determinantes do Desequilíbrio Financeiro dos Municípios Portugueses. (Tese de mestrado). Instituto Politécnico de Bragança, Portugal.

Bryman, A. (1984). The debate about quantitative and qualitative research: a question of method or Epistemology?. The British Journal of Sociology, 35(1), 75-92.

Bulteel, K., Wilderjans, T.F., Tuerlinckx, F. *et al.* Hull, C. (2013). An alternative to AIC and BIC in the context of mixtures of factor analyzers. Behav Res 45, 782–791.

Caldas, P. Rating Municipal Português: O Indicador de Sustentabilidade. Lisboa, 7 de maio de 2019. Disponível

em: http://www.ordemeconomistas.pt/xportalv3/file/XEOCM_Documento/58973792/file/Apresentacao%20do%20RMP%20de%20maio%202019%20FIM1.pdf.

Consultado a 23 de março de 2020.

Carvalho, F. (2013). Análise Fatorial. (Tese de mestrado). Faculdade de Ciências e Tecnologia da Universidade de Coimbra, Portugal.

Castro, V. F. (2009). Imigração, mercado de trabalho e desenvolvimento em contextos regionais de baixas densidades. Cadernos de Geografia n°28/29 – 2009/10, Faculdade de Letras da Universidade de Coimbra, pp. 61-71.

Consulstaff. Programa de Investimento em Territórios de Baixa Densidade. Disponível em: <https://consulstaff.pt/programa-investimento-territorios-baixa-densidade-2-fase-candidaturas/>. Consultado a 30 de outubro de 2020.

Cordeiro, M. (2008). Partição de grupos e validação da análise de agrupamento para equipamentos de fiscalização eletrónica de trânsito. Universidade Federal do Paraná, Brasil.

Costa, G. (2017). Uso da Técnica Two Step Cluster para Segmentação de Funcionários de uma Empresa no Rio de Janeiro segundo Clima Organizacional: Um Estudo de Caso. Revista da Estatística UFOP, vol VI. p. 2237-8111.

Costa, M. (2016). Os Determinantes da Boa Governança Municipal. (Tese de mestrado). Instituto Superior Técnico da Universidade de Lisboa, Portugal.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7), p. 1-9.

Creswell, J. (2009). *Research design: Qualitative, quantitative and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Damáσιο, B. (2012). *Uso da Análise Fatorial em Psicologia*. Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.

Direção Regional de Estatística da Madeira (2016). Relatório de Atividades 2016. Disponível em: <https://estatistica.madeira.gov.pt/>. Consultado a 2 de novembro de 2020.

EDTI (2019). BoxPlot: Saiba tudo sobre o Diagrama de Caixa e como interpretar esse gráfico. Disponível em: <https://www.escolaedti.com.br/o-que-e-um-box-plot>. Consultado a 29 de setembro de 2020.

Emad, A. M., Naugler, C., & Behrouz F. H. (2015). Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology. *Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics*, 32. p. 577-602.

Fagerberg, J., Knell, M., & Srholec, M. (2004). The competitiveness of nations : economic growth in the ECE regions, UNECE Spring Seminar Competitiveness and Economic Growth in the ECE Regions, Geneva.

Fernandes, A. (2017). *Análise de Conglomerados: comparação de técnicas e uma aplicação a dados de fluxo migratório em Portugal*. Universidade de Aveiro, Portugal.

Fernandes, M. J., Camões, P. J., & Jorge, S. (2018). *Anuário Financeiro dos Municípios Portugueses 2018*. Lisboa: Ordem dos Contabilistas Certificados.

Ferreira, A. (2012). *Identificação de Perfis de Hábitos de Consumo através da utilização de cartões bancários*. (Tese de mestrado). Instituto Superior Técnico, Portugal.

Fraley, C., & Raftery, A. E. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis. *Computer Journal*, 4. p. 578–588.

Gaspar, P. (2017). População residente na Madeira diminui e envelhece. *Jornal Económico*. Disponível em: <https://jornaleconomico.sapo.pt/noticias/populacao-residente-na-madeira-cai-06-250343>. Consultado a 1 de novembro de 2020.

Geraldes, J. (2016). *Economia Aplicada e Previsão*. (Tese de mestrado). Instituto Superior de Economia e Gestão, Portugal.

Gomes, N. (2017). *A Mulher Estrangeira na População Residente em Portugal*. Observatório das Migrações, ACM, I.P., V1. p. 8.

Gonçalves, R. F., Matos, F. A., & Manso, P. J. (2012). *Os Municípios e a Qualidade de Vida em Portugal: Construção de um Índice de Concelhio de Desenvolvimento Económico e Social*. Universidade da Beira Interior, Portugal. Disponível em: <http://wordpress.ubi.pt/dge/wp-content/uploads/sites/12/2019/02/qualidade-vida-concelhio-2012-pmanso.pdf>. Consultado a 14 de abril de 2020.

Huang, Z. (1998). Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Mining and Knowledge Discovery*, 2. p. 283–304.

International Institute for Management Development (IMD). *IMD World Competitiveness Ranking 2020*. Lausanne, Suíça. Disponível em: <https://worldcompetitiveness.imd.org/rankings/WCY>. Consultado a 7 de setembro de 2020.

International Institute for Management Development (IMD). *IMD World Competitiveness Ranking 2020: showing strength of small economies*. Lausanne, Suíça. Disponível em: <https://www.imd.org/news/updates/IMD-2020-World-Competitiveness-Ranking-revealed/>. Consultado a 7 de setembro de 2020.

Jick, T. (1979). Mixing qualitative and quantitative methods. *Triangulation in action*. *Administrative Science Quarterly*, 24(4). p. 602-611.

Lima, L., Maroldi, A., & Silva, D. Outlier(s) nos cálculos bibliométricos: Primeiras aproximações. Outlier(s) | in bibliometric calculations: preliminary approach. *Liinc em Revista*, v. 9, n. 1, 27 maio 2013. Disponível em: <https://doi.org/10.18617/liinc.v9i1.522>. Consultado a 20 de agosto de 2020.

Lopes, T. P. A competitividade da Economia. *Mind the Gap!*. Porto, 24 de julho 2020. Disponível em: <https://www.pbs.up.pt/pt/artigos-e-eventos/artigos/a-competitividade-da-economia-portuguesa-mind-the-gap/>. Consultado a 5 de setembro de 2020.

Maciel, A., Vinhas, L., & Câmara, G. (2015). *Algoritmos de Clustering para separação de culturas agrícolas e tipos de uso e cobertura da Terra utilizando dados de sensoriamento remoto*. Instituto Nacional de Pesquisas Espaciais – INPE, Brasil.

Marôco, J. (2010b). Análise Estatística com o PASW Statistics. Pêro Pinheiro: Report Number.

Mercal (2018). Turismo rural em Portugal: situação, conceito e apoios. Disponível em: <https://mercal.pt/turismo-rural-em-portugal/>
Consultado a 29 de outubro de 2020.

Ninho, C. (2011). Aplicação de Mineração de Dados nas Transações de Compras em Empresa do Segmento de Petróleo e Gás. (Tese de Mestrado). Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa em Engenharia, Brasil.

Nunes, D. Turismo rural com as menores perdas em julho. Dinheiro Vivo, 15 de setembro de 2020. Disponível em: <https://www.dinheirovivo.pt/empresas/turismo-rural-com-as-menores-perdas-em-julho-12895878.html>.
Consultado a 19 de outubro de 2020.

Oliveira, B. Análise Fatorial: Uma importante Técnica Multivariada. Brasil, 25 de setembro de 2019. Disponível em: <https://operdata.com.br/blog/analise-fatorial/>.
Consultado em: 7 de abril de 2020.

Oliveira, B. Análise de Componentes Principais. Brasil, 11 de outubro de 2019. Disponível em: <https://operdata.com.br/blog/analise-de-componentes-principais/>.
Consultado em: 7 de abril de 2020.

Oliveira, C. (2020). Observatório das Migrações – Relatório Estatístico do Asilo 2020. Disponível em: <https://www.om.acm.gov.pt/documents/58428/440932/Relatorio-Asilo-2020-OM.pdf/9f22513a-46ff-4f4b-a292-93347463d71c>.
Consultado a 3 de novembro de 2020.

Oliveira, M. (2010). Monitorizando a Evolução de Clusters. (Tese de mestrado). Faculdade de Economia da Universidade do Porto, Portugal.

PDR (2020). Disponível em: <http://www.pdr-2020.pt/>
Consultado a 29 de outubro de 2020.

Pestana, M., & Gageiro, J. (2000). Análise de Dados para Ciências Sociais - A complementaridade do SPSS. Lisboa: Edições Sílabo.

Pinto, I. Preços das casas aumentaram mais de 55% no últimos 3 anos. Diário de Notícias, 10 de julho de 2020. Disponível em: <https://www.dn.pt/edicao-do-dia/10-jul-2020/precos-das-casas-em-lisboa-aumentaram-mais-de-55-nos-ultimos-tres-anos-12406593.html>.
Consultado a 4 de novembro de 2020.

Pires, S. *et al.* Qualidade da governação local em Portugal. Fundação Francisco Manuel dos Santos, Lisboa, novembro de 2018. Disponível em: <https://www.ffms.pt/FileDownload/14b94278-1366-4df0-8740-aedcb3c8876f/qualidade-da-governacao-local-em-portugal>.
Consultado em: 10 de abril de 2020.

PORDATA. Municípios. Disponível em: <https://www.pordata.pt/Municipios>.
Consultado a: 25 de fevereiro de 2020.

Portal Action. Estatísticas descritivas: 2.3 – Quartis. Disponível em: <http://www.portalaction.com.br/estatistica-basica/23-quartis>.
Consultado em: 27 de setembro de 2020.

Portal dos Incentivos. PDR 2020 – Programa de Desenvolvimento Rural de Portugal. Disponível em: <https://www.portaldosincentivos.pt/index.php/pdr-2020>.
Consultado a 29 de outubro de 2020.

Portal dos Incentivos. SI2E – Sistema de Incentivos ao Empreendedorismo e ao Emprego. Disponível em: <https://www.portaldosincentivos.pt/index.php/si2e>.
Consultado a 29 de outubro de 2020.

Portal Educação. Quartil e Percentil. Disponível em: <https://siteantigo.portaleducacao.com.br/conteudo/artigos/administracao/quartil-e-percentil/30565>.
Consultado em: 29 de setembro de 2020.

Portela, D. (2012). Contributo das Técnicas de Análise Fatorial para o Estudo do Programa “Ocupação Científica de Jovens nas Férias”. (Tese de mestrado). Universidade Aberta, Portugal.

Portela, D. (2012). Contributo das Técnicas de Análise Fatorial para o Estudo do Programa “Ocupação Científica de Jovens nas Férias”. (Tese de mestrado). Universidade Aberta, Portugal.

Porto Business School. Portugal alcança o 37º lugar no ranking da competitividade. Porto, 15 de junho 2020. Disponível em: <https://www.pbs.up.pt/pt/artigos-e-eventos/artigos/portugal-sobe-dois-lugares-e-alcanca-o-37%C2%BA-lugar-no-ranking-mundial-de-competitividade/>.
Consultado a 5 de setembro de 2020.

República Portuguesa (2020). População estrangeira residente em Portugal aumenta pelo quarto ano consecutivo. Disponível em: <https://www.portugal.gov.pt/pt/gc22/comunicacao/noticia?i=populacao-estrangeira-residente-em-portugal-aumenta-pelo-quarto-ano-consecutivo>.
Consultado a 4 de novembro de 2020.

SEF/GEFP (2018), Relatório de Emigração, Fronteiras e Asilo de 2018, Barcarena, Serviço de Estrangeiros e Fronteiras.

SEF. Vistos Gold. Disponível em: <https://www.sef.pt/pt/pages/conteudo-detalle.aspx?nID=62>.
Consultado a 5 de novembro de 2020.

Sharma, S. (1996). Applied Multivariate Techniques, New York: Wiley.

Silva, R. (2007). Estudo do desempenho do algoritmo Agrupamento em Duas Etapas através de comparações realizadas sob metodologia de planeamento de experimentos. (Tese de mestrado). Universidade Federal de Santa Catarina, Brasil.

Soares, J. O., Lourenço Marquês, M.M. & Monteiro, C. A. multivariate methodology to uncover regional disparities: A contribution to improve European Union and governmental decisions. *European Journal of Operational Research* 145, p. 121–135.

Souza, E. F. (2007). *Comparação e escolha de agrupamentos: uma proposta utilizando a entropia*. (Tese de mestrado). Instituto de Matemática e Estatística, Universidade de São Paulo, Brasil.

SPHINX Brasil (2014). *Análise: Da Univariada à Multivariada*. Disponível em: <https://www.sphinxbrasil.com/blog/analise-da-univariada-a-multivariada>.

Consultado em: 10 de março de 2020.

Taveira, F. (2014). *Análise de risco com base em indicadores: Remoção indiferenciada de resíduos pelos cantoneiros de limpeza na Câmara Municipal de Lisboa*. (Tese de mestrado). Instituto Politécnico de Setúbal, Portugal.

Teixeira, D. A. *et al.* (2009). *O Relatório de Gestão e a Avaliação e Comparação do Desempenho dos Municípios Portugueses: O Caso dos Municípios no Distrito de Setúbal*. Instituto Politécnico de Setúbal, Portugal. Disponível em:

<https://comum.rcaap.pt/bitstream/10400.26/5415/1/ART%20JORNADAS%20LUSO%20HISP%202012%20ABT%20CM%20PP%20NT.pdf>.

Consultado a 3 de maio de 2020.

Theodoris, S. & Koutrombas, K. (1999). *Pattern recognition*. Academic Press, New York.

Turismo de Portugal (2020). Disponível em:

http://www.turismodeportugal.pt/pt/Turismo_Portugal/visao_geral/Paginas/default.aspx.

Consultado a 29 de outubro de 2020.

UNHRC. Disponível em: <https://www.unhcr.org/>.

Consultado a 3 de novembro de 2020.

ANEXOS

O **Cluster 1** contém 57 municípios:

Tabela 1 - Municípios pertencentes ao Cluster 1

Arcos de Valdevez	Valpaços	Torre de Moncorvo	Sernancelhe	Trancoso	Alvaiázere
Caminha	Vila Pouca de Aguiar	Vila Flor	Tondela	Vila Nova de Foz Côa	Ansião
Melgaço	Alfândega da Fé	Vimioso	Vouzela	Pampilhosa da Serra	Castanheira de Pêra
Monção	Bragança	Vinhais	Aguiar da Beira	Penacova	Pedrógão Grande
Paredes de Coura	Carrazeda de Ansiães	Anadia	Almeida	Penela	Ferreira do Zêzere
Ponte da Barca	Freixo de Espada à Cinta	Sever do Vouga	Gouveia	Oleiros	Mação
Valença	Macedo de Cavaleiros	Castro Daire	Manteigas	Proença-a-Nova	Sardoal
Boticas	Miranda do Douro	Mortágua	Pinhel	Sertã	Marvão
Montalegre		Penedono	Sabugal	Vila de Rei	Monchique
Murça		São Pedro do Sul	Mirandela		
			Mogadouro		

O **Cluster 2** é o cluster de menor dimensão apresenta somente 3 municípios: Lisboa, Porto e Funchal.

O **Cluster 3** contém 64 municípios:

Tabela 2 - Municípios pertencentes ao Cluster 3

Barcelos	Amares	Paços de Ferreira	São João da Madeira	Ponte de Lima	Santa Maria da Feira
Cabeceiras de Basto	Vila Verde	Paredes	Cinfães	Marco de Canaveses	Vila Franca do Campo
Celorico de Basto	Mesão Frio	Penafiel	Lamego	Calheta	Calheta
Esposende	Mondim de Basto	Póvoa de Varzim	Resende	Lagoa	Lajes das Flores
Fafe	Peso da Régua	Santo Tirso	São João da Pesqueira	Lajes das Flores	Lobos
Guimarães	Santa Marta de Penaguião	Trofa	Sátão	Madalena	Machico
Póvoa de Lanhoso	Amarante	Arouca	Tabuaço	Madalena	Ponta do Sol
Vieira do Minho	Baião	Castelo de Paiva	Tarouca	Ponta Delgada	Porto Santo
Vila Nova de Famalicão	Felgueiras	Oliveira de Azeméis	Miranda do Corvo	Povoação	Ribeira Brava
	Gondomar	Ovar	Angra do Heroísmo	Praia da Vitória	Santa Cruz
	Lousada			Ribeira Grande	Santana
				Santa Cruz das Flores	São Vicente
				Velas	

O **Cluster 4** contém 80 municípios:

Tabela 3 - Municípios pertencentes ao Cluster 4

Armamar	Espinho	Sabrosa	Terras de	Alijó	Cuba
Moimenta da	Mêda	Ribeira de	Bouro	Estremoz	Ferreira do
Beira	Seia	Pena	Fronteira	Montemor-o-	Alentejo
Nelas	Arganil	Nazaré	Gavião	Novo	Mértola
Penalva do	Góis	Abrantes	Monforte	Mora	Moura
Castelo	Oliveira do	Alpiarça	Nisa	Mourão	Ourique
Santa Comba	Hospital	Chamusca	Ponte de Sor	Portel	Serpa
Dão	Soure	Coruche	Portalegre	Redondo	Vidigueira
Vila Nova de	Tábua	Golegã	Sousel	Reguengos de	Alcoutim
Paiva	Belmonte	Alter do	Alenquer	Monsaraz	Corvo
Celorico da	Covilhã	Chão	Alcácer do	Vendas Novas	Lajes do Pico
Beira	Fundão	Arronches	Sal	Viana do	Nordeste
Figueira de	Idanha-a-	Avís	Grândola	Alentejo	Santa Cruz da
Castelo Rodrigo	nova	Campo	Alandroal	Vila Viçosa	Graciosa
Fornos de	Penamacor	Maior	Arraiolos	Aljustrel	Porto Moniz
Algodres	Vila Velha	Castelo de	Borba	Almodôvar	
Figueiró dos	de Ródão	Vide			
Vinhos		Crato			
Barrancos		Elvas			

O **Cluster 5** contém 17 municípios:

Tabela 4 - Municípios pertencentes ao Cluster 5

Alvito	Lagoa	Silves
Odemira	Lagos	Tavira
Albufeira	Loulé	Vila do Bispo
Aljezur	Olhão	Vila Real de Santo
Castro	Portimão	António
Marim	São Brás de	São Roque do Pico
Faro	Alportel	

O **Cluster 6** contém 87 municípios:

Tabela 5 - Municípios pertencentes ao Cluster 6

Viana do	Oliveira do	Murtosa		Mealhada	Montemor-o-
Castelo	Bairro	Vila Nova de	Almeirim	Azambuja	velho
Vila Nova de	Vagos	Poiães	Benavente	Cadaval	Mira
Cerveira	Vale de	Castelo	Cartaxo	Cascais	Barreiro
Braga	Cambra	Branco	Constância	Loures	Moita
Chaves	Carregal do	Alcobaça	Entrocament	Lourinhã	Montijo
Vila Real	Sal	Batalha	o	Mafra	Palmela
Maia	Mangualde	Bombarral	Ourém	Odivelas	Santiago do
Matosinhos	Oliveira de	Caldas da	Rio Maior	Oeiras	Cacém
Valongo	Frades	Rainha	Salvaterra de	Sintra	Seixal
Vila do	Viseu	Leiria	Magos	Sobral de	Sesimbra
Conde	Guarda	Marinha	Santarém	Monte	Setúbal

Vila Nova de Gaia Águeda Albergaria-a-Velha Aveiro Estarreja Ílhavo	Cantanhede Coimbra Condeixa-a-Nova Figueira da Foz Lousã	Grande Óbidos Peniche Pombal Porto de Mós Alcanena	Tomar Torres Novas Vila Nova da Barquinha Amadora Arruda dos Vinhos	Agraço Torres Vedras Vila Franca de Xira Alcochete Almada	Sines Évora Beja Castro Verde Horta Vila do Porto
--	--	--	---	---	--