

# UE a 20 e a educação ao longo da vida:

## O impacto do local residência sobre a educação através de uma Análise de Clusters

UNIVERSIDADE ABERTA

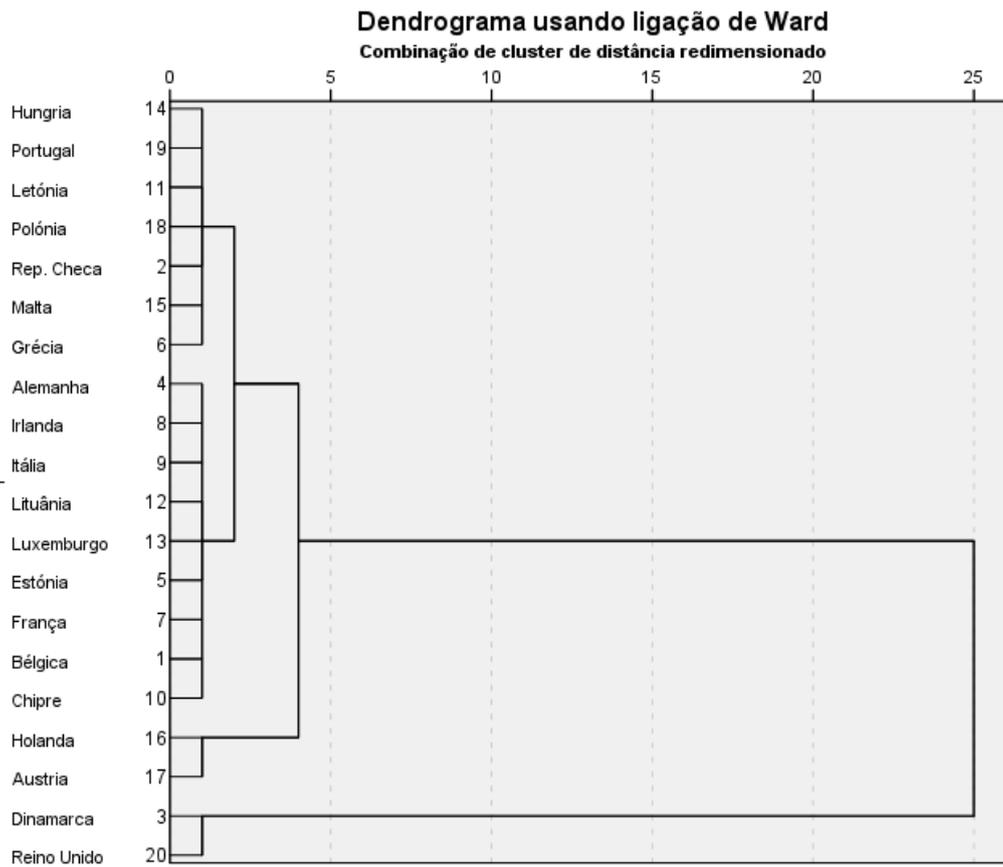
22002 - Análise de dados multivariados e aplicações

Ricardo Neves Pires 2001645

O estudo em análise remete-se para a aprendizagem/educação ao longo da vida nos diferentes estados membros da União europeia. Pretende-se detetar se os diferentes períodos de adesão dos respetivos países tem algum impacto de como os rendimentos dos agentes económicos são aplicados na constante aprendizagem. Seja esta através da aprendizagem de novos conteúdos em sede profissional de requalificações, reforço e atualização de conhecimentos já adquiridos ou simplesmente por uma questão de curiosidade e estímulo intelectual. Pela análise de clusters iremos tentar perceber se existe efetivamente uma diferenciação pelo agrupamento dos países e se isto se deve consoante o momento de adesão do país à União Europeia ou por outro facto, se o momento de adesão nada tem haver com a **percentagem de pessoas em aprendizagem ao longo da vida** e da **percentagem de PIB gasto em educação** mas sim por simples políticas interna dos países. Isto é, países que tenha aderido tardiamente à UE e subsequentemente obtido apoios de fundos de desenvolvimento e pela implementação de directrizes políticas de elevação e nivelção de padrões de educação e aprendizagem. Ou, que não seja de todo o caso. Temos na nossa base de dados 20 países e duas variáveis já supramencionadas a negrito. Relativamente ao método a ser usada, este recai no Hierárquico pelo facto de não existir suspeitas de quantos K cluster existam, o que seria necessário no método não hierárquico e pelo facto de estarmos perante uma base dados relativamente reduzida não levando a problemas de computação, visto que este método tende a ser lento para base de dados grandes. Relativamente às variáveis opta-se por não efetuar qualquer tipo de transformação visto estarmos perante valores relativos (percentagem). Usa-se a medida de distância euclidiana ao quadrado (1). Dois critérios de agregação vão ser usados, o de Ward e o single linkage, por forma a comparar diferenças que possam existir, visto que ambos os critérios tem algumas desvantagens [2].

$$d_{ij} = \sum_{v=1}^n (X_{iv} - X_{jv})^2 \quad (1)$$

em que  $v$  são as variáveis e,  $i$  e  $j$  são os valores dos sujeitos. O dendograma do critério de Ward é representado aqui pela fácil leitura que este nos trás. De referir que pelo critério de single linkage os resultados são idênticos, dando indícios de que o critério extrai a estrutura existente nos dados e não uma estrutura alheia.



Observam-se pelo dendograma quatro clusters. Contudo, testa-se perante o critério do  $R^2$  onde é possível verificar a variabilidade retida em cada uma das soluções dos Clusters e se de facto os quatro cluster retem grande parte da variância total. Para os quatro clusters obtém-se,

$$SQC = 952.049 + 5.828 = 957.877$$

$$SQT = 972.478 + 18.152 = 990.63$$

em que o quociente destes dois valores remete para o  $R^2 = 96,6\%$ , que explica a variância total. Testou-se a análise para três clusters em que se obteve  $R^2 = 93.18\%$ , inferior ao de quatro clusters, sendo que retemos estes últimos. De notar que seria pouco expectável pela permissa inicial de adesão à UE, encontrar países como a Lituânia, Estónia e Chipre agrupados aos países fundadores. Sendo que a mesma fica descartada pela análise de clusters. Inclusive,

seria de esperar e pelo o que foi afirmado acima, que países com níveis de rendimentos e PIB elevados tivessem agrupados. Como o caso dos dois clusters onde se encontram Holanda e Austaria por um lado e Dinamarca e Reino Unido por outro. Por último, o cluster onde estão os restante sete países, incluído Portugal, seria expetável que estes pudessem estar agrupados. Por forma a validar ainda mais a nossa análise, procedeu-se à transformação das variáveis, tendo chegado às mesmas conclusões. Após o uso de dois critérios de agregação, critério do  $R^2$  em número de cluster diferentes e tendo efetuado uma análise das variáveis em separado, na tentativa de suportar a nossa permissa, que tal não foi possível. Existe um grau de confiança sobre os resultados, em que o período de adesão à UE não é um fator de diferenciação. Não obstante, o número e tipo de variáveis, e aumento da base de dados possa conduzir a resultados diferentes. No caso concreto é bem possível que as políticas internas e delineamento histórico de cada país, independentemente do período de adesão reflitam o agrupamento e semelhanças no que toca à percentagem do PIB gasto em educação e percentagem de pessoas em aprendizagem ao longo da vida. No essencial os países que tenham aderido à UE tardiamente estão a par e passo dos países considerados os líderes Europeus. Per si a adesão não é a causa da nivelação existente entre estados membros mas talvez sim, as políticas internas e as capacidades de cada estado membro incutir as directrizes europeias atempadamente e eficientemente ao abrigo da UE.

## BIBLIOGRAFIA

[1] Marôco, J., *Análise Estatística com o SPSS STATISTICS v.27*, 8<sup>a</sup> edição, Report Number, 2021

[2] Reis, E., *Estatística Multivariada Aplicada*, 2<sup>a</sup> Edição, Edições Sílabo, 2001.