

# Divisão Digital na União Europeia

Frederico Cruz-Jesus <sup>1</sup>, Tiago Oliveira <sup>2</sup>, Fernando Bacao <sup>3</sup>

1) Instituto Superior de Estatística e Gestão de Informação – Universidade Nova de Lisboa

[frederico.macruz@gmail.com](mailto:frederico.macruz@gmail.com)

2) Instituto Superior de Estatística e Gestão de Informação – Universidade Nova de Lisboa

[toliveira@isegi.unl.pt](mailto:toliveira@isegi.unl.pt)

2) Instituto Superior de Estatística e Gestão de Informação – Universidade Nova de Lisboa

[bacao@isegi.unl.pt](mailto:bacao@isegi.unl.pt)

## Resumo

O nosso trabalho tem como objectivo fornecer aos decisores políticos uma análise completa e multidimensional às assimetrias no desenvolvimento digital, dentro da União Europeia a 27 (UE-27), através da Análise Factorial e da Análise de Clusters. Foram encontradas 2 dimensões latentes e 5 grupos de países distintos no desenvolvimento digital dos diferentes estados membros da UE-27.

**Palavras-chave:** União Europeia, Divisão Digital, Adopção de Tecnologias de Informação, Sociedade da Informação, Comércio Electrónico e Desenvolvimento Digital.

## 1. Introdução

A atenção dada pelos líderes a nível mundial aos conceitos de sociedade da informação e de divisão digital tem, ao longo dos últimos anos, aumentado significativamente. Na Cimeira Mundial sobre a Sociedade da Informação (WSIS) foi declarado que o desafio global para o novo milénio é a construção de uma sociedade “onde todos possam criar, aceder, utilizar e partilhar a informação e conhecimento, permitindo que indivíduos, comunidades e povos possam alcançar o seu pleno potencial na promoção do seu desenvolvimento sustentável e melhoria da qualidade de vida” (WSIS, 2003, 2005).

Paralelamente, a União Europeia (UE) acaba de lançar a Estratégia “Europa 2020”, que tem como objectivo atingir “um crescimento inteligente, sustentável e inclusivo para a Economia Europeia” (European Commission, 2010b) e “para sair da crise e preparar a economia da UE para os desafios da próxima década” (European Commission, 2010a). Este crescimento económico será realizado, entre outros factores, pelo desenvolvimento de uma economia baseada no conhecimento e na inovação (European Commission, 2010b). Incluídos na Estratégia Europa 2020, como um dos sete pilares da Estratégia, a Agenda Digital para a Europa foi desenvolvida tendo como objectivo definir o papel central que o uso das tecnologias de informação e comunicação (TIC) devem ter, se a Europa quiser concretizar as suas ambições para 2020 (European Commission, 2010a). Logo, assimetrias digitais significativas, dentro da União Europeia a 27 (UE-27), devem ser detectadas e minimizadas de forma a não comprometer os objectivos expressos na Estratégia 2020.

Pelo que conhecemos, até agora, há uma carência de estudos que abordam a temática da divisão digital, sobretudo no contexto da UE-27. Considerando a importância que a Comissão Europeia dá a um desenvolvimento digital homogêneo a todos os seus membros, o primeiro passo a tomar para alcançar este objectivo é avaliar a situação actual dentro da UE. Portanto, esperamos que este estudo possa preencher esta lacuna e lançar alguma luz sobre esta questão, para que políticas eficientes e assertivas possam ser desenvolvidas.

## **2. Desenvolvimento e Divisão Digital**

O termo “divisão digital” (*digital divide*) foi usado pela primeira vez em meados da década de 90 pelo ex-secretário adjunto de Comunicação e Informação dos Estados Unidos, do Departamento de Comércio, Larry Irving Junior (Dragulanescu, 2002). Segundo a Organização para a Cooperação Económica e Desenvolvimento (OCDE), “o termo divisão digital refere-se às disparidades entre indivíduos, famílias, empresas e áreas geográficas em diferentes níveis socioeconómicos, tanto em relação às suas oportunidades de acesso às TIC como ao uso da Internet para uma ampla variedade de actividades” (OECD, 2001).

Em primeiro lugar, a exclusão digital foi entendida em termos binários, ou seja, a diferença entre em “ter” ou “não ter” acesso às Tecnologias de Informação. Hoje, no entanto, essa diferença é considerada insuficiente, uma vez que outros factores têm de ser considerados. Com efeito, hoje em dia, a divisão digital é percebida como um fenómeno complexo multidimensional (Bertot, 2003; Okazaki, 2006; Warschauer, 2002).

Existem dois tipos distintos de divisão digital. O primeiro está localizado a um nível internacional, ou seja, entre diferentes países. O segundo refere-se a uma dimensão intranacional, ou seja, dentro de um país. Nestes tipos de divisões digitais, as assimetrias podem verificar-se entre diferentes regiões, ou grupos de indivíduos, quando características de diferentes naturezas são verificadas como, por exemplo, socioeconómicas, de género, etnia ou mesmo idade (Ono & Zavodny, 2007; Unesco, 2003). Assim, de acordo com esta definição, a divisão digital pode representar uma ameaça real para todas as estratégias de inclusão digital, tanto a nível nacional como internacional, espalhadas por todo o mundo, incluindo a agenda digital para a Europa (Cuervo & Menéndez, 2006; OECD, 2009).

O desenvolvimento e utilização das TIC tem apresentado um crescimento exponencial nas últimas décadas. Estas tecnologias estão a desempenhar um papel decisivo na melhoria da maioria dos aspectos da nossa sociedade, incluindo o sector dos negócios, comunicação, política e economia. Assim sendo, novos tipos de interacções, ou serviços avançados, estão a tornar-se cada vez mais comuns, como o comércio electrónico, governo electrónico, saúde electrónica, ensino electrónico, banca electrónica ou finanças electrónicas, entre outros (Çilan, Bolat, & Coskun, 2009; European Commission, 2006; Selwyn & Facer, 2007; María Rosalía Vicente & Gil-de-Bernabé, 2010; Maria Rosalia Vicente & Lopez, 2010). Acções e tecnologias como navegar na Internet, aceder ao YouTube, receber e enviar e-mails, aceder a wikis ou consultar bibliotecas on-line, estão a ser amplamente utilizadas no nosso dia-a-dia, melhorando a forma como as pessoas interagem umas com as outras e têm acesso a informações privilegiadas face àqueles que não têm acesso a estas tecnologias (Brooks, Donovan, & Rumble, 2005). Deste modo, é consensual que as TIC afectam, positivamente, a economia e o bem-estar em variadas dimensões distintas (Çilan, et al., 2009; The World Bank, 2006). As TIC criam vantagens competitivas nas empresas, melhoram os sistemas nacionais de saúde (Bakker, 2002) através da saúde electrónica, melhoram os sistemas de educação através do ensino electrónico (Cukusic, Alfirovic, Granic, & Garaca, 2010), que traz vantagens na redução das distâncias geográficas entre os alunos e os estabelecimentos de ensino, criando também novos sectores de actividade e negócios que diminuem as taxas de desemprego (Castells, 2007; Castells & Himanen, 2007). Assim, para estes benefícios serem atingidos, alguns obstáculos devem ser ultrapassados,

nomeadamente as desigualdades dentro e entre países, no que se refere ao acesso a estas tecnologias, ou seja, as assimetrias digitais devem ser minimizadas.

### **3. Quantificação da Divisão Digital**

#### **Enquadramento**

Com o objectivo de medir a adopção das TIC, em cada um dos Estados-Membros da UE, para uma posterior comparação múltipla, alguns constrangimentos tiveram de ser ultrapassados. Devido à sua inquestionável importância na melhoria da economia e condições de vida, o problema de como medir esta adopção, tem adquirido uma importância crescente (Cuervo & Menéndez, 2006; OECD, 2009; The World Bank, 2006; Vehovar, Sicherl, Husing, & Dolnicar, 2006).

Em primeiro lugar, não há uma definição clara, única e padronizada de desenvolvimento digital, sociedade da informação ou divisão digital, uma vez que, as considerações sobre este assunto diferem em diferentes países, regiões geográficas, e modelos de sociedade da informação adoptados. Na verdade, existem vários modelos na sociedade de informação, como os da Finlândia, Singapura ou Estados Unidos (Silicon Valley). Cada um destes dá especial ênfase a características e objectivos distintos, que devem estar devidamente adequados às respectivas realidades nacionais (Castells & Himanen, 2007).

A segunda restrição está relacionada com a falta de dados harmonizados disponíveis, para esta temática, quando se consideram vários países. Assim, existe normalmente uma relação limitada entre a profundidade e a largura da análise. Ou seja, quanto mais indicadores tentamos usar, menos são os países que podem ser incluídos na nossa análise (Cuervo & Menéndez, 2006). Apesar de tudo, este problema foi atenuado pelo facto da nossa análise incidir sobre países membros da União Europeia, uma vez que o Eurostat, entre outras organizações, dispõe de um leque variado de oferta de dados harmonizados e específicos sobre esta temática para cada um dos 27 países. Assim, todos os dados utilizados neste estudo foram recolhidos a partir da categoria “Estatísticas da Sociedade da Informação” no site do Eurostat e dizem respeito ao ano de 2009.

De acordo com as recomendações da OCDE, as variáveis que devem ser usadas para medir a divisão digital variam com o objectivo da pesquisa. Por exemplo, se pretendemos medir as disparidades no acesso às TIC a nível nacional, devemos desagregar os indicadores a dimensões como grupos etários, género, idade, classe social, escolaridade, local geográfico, etc. Para medir a divisão digital entre diferentes países a nível internacional, os indicadores devem estar agregados à realidade nacional, excluindo assimetrias internas. Como temos como objectivo medir a divisão digital dentro da União Europeia, vamos seguir esta metodologia.

Estudos recentes concluíram que a divisão digital internacional é principalmente uma consequência de desigualdades económicas entre diferentes países. Na verdade, termos como “rico em informação” e “pobre em informação” já foram utilizados para classificar países no âmbito do desenvolvimento digital. Além do desenvolvimento económico, países com menor nível de escolaridade na sua população, tendem igualmente a apresentar níveis de adopção de TIC mais baixas (M. D. Chinn & Fairlie, 2007; Menzie D. Chinn & Fairlie, 2010; Hargittai, 1999; Kiiski & Pohjola, 2002; Pohjola, 2003).

#### **Dados**

Para medir os diferentes níveis das sociedades de informação dentro da UE-27, utilizaram-se 13 variáveis, de acordo com as recomendações da OCDE, da Comissão Europeia e que já foram utilizados em estudos anteriores. As variáveis são as seguintes (ver Tabela 1).

Código	Variável
HsInt	% de habitações com ligação à Internet
HsBro	% de habitações com ligação de Banda Larga à Internet
IntPop	% da população que utiliza regularmente a Internet
IntSrc	% da população que utiliza a Internet para pesquisar produtos ou informações comerciais
ebank	% da população que utiliza serviços de e-banking
elearn	% da população que utiliza serviços de e-learning
email	% da população que utiliza correio electrónico (email)
ehealth	% da população que utiliza a Internet para pesquisar informações sobre saúde
egovInf	% da população que utiliza a Internet para obter informações de entidades públicas
egovInt	% da população que utiliza a Internet para interagir com entidades públicas
ecom	% de empresas que vendem online
esafe	% de empresas que vendem online oferecendo capacidade de ligações seguras
egovsup	% de serviços públicos disponíveis online

Tabela 1 – Códigos e descrição das variáveis usadas

Tivemos em conta, considerando as limitações na disponibilidade de dados para todos os países da UE, a inclusão de indicadores já utilizados bem como indicadores novos que considerámos relevantes para o estudo. A percentagem de domicílios com acesso a conexões de banda larga (HsBro) e a percentagem de serviços públicos disponíveis online (egovsup) foram usadas por Cuervo e Menéndez (2006), entre outros indicadores, para medir a divisão digital dentro da UE-15. Variáveis semelhantes à percentagem de domicílios com acesso à Internet (HsInt) e com acesso a conexões de banda larga (HsBro), à percentagem da população que utiliza regularmente a Internet (IntPop), à percentagem da população que utiliza regularmente a Internet para encontrar informações comerciais (IntSrc), à percentagem da população que utiliza a Internet para obter informações e interagir com autoridades públicas (egovInf e egovInt, respectivamente) e à percentagem da população que utiliza regularmente serviços de e-learning (eLearn); foram usados por Çılan, Bolat e Coskun (2009), para analisar a divisão digital entre países membros e países candidatos à UE, antes do alargamento de 2004. Além disso, a UE através da Agenda Digital para a Europa (2010a) dá especial ênfase, ao papel de “serviços avançados” como a saúde electrónica, educação electrónica, banca electrónica e governo electrónico. Os serviços de e-banking e e-health são considerados como “alguns dos serviços online mais inovadores e avançados” (2010a); Os serviços de governo electrónico também são descritos na Agenda Digital, uma vez que “apesar de existir um elevado nível na disponibilidade de serviços públicos electrónicos na Europa, diferenças significativas no acesso a estes, entre Estados-Membros, ainda se verificam” (2010a); Por isso incluímos estas variáveis para que possam medir a adopção destes serviços, sendo estas particularmente relevantes para nossa análise uma vez que, segundo a literatura consultada, existe uma lacuna na quantificação destas variáveis.

Consideramos também que ter extraído todas estas variáveis do Eurostat dá a garantia que o resultado da análise não é comprometido pela qualidade e adequabilidade dos dados utilizados. Os dados são apresentados na Tabela 2.

Pais	HsInt	HsBro	IntPop	IntSrc	ebank	elearn	email	ehealth	egovInf	egovInt	ecom	esafe	egovsup
Austria	70	58	67	54	35	34	63	35	35	39	10	4	100
Belgica	67	63	70	59	46	35	68	32	27	31	16	6	70
Bulgaria	30	26	40	17	2	14	34	9	8	10	3	1	40
Chipre	53	47	45	39	15	25	38	16	21	22	7	3	50
Republica Checa	54	49	54	50	18	27	55	20	23	24	15	7	60
Dinamarca	83	76	82	74	66	58	81	46	65	67	19	10	84
Estónia	63	62	67	54	62	31	62	32	43	44	11	3	90
Finlandia	78	74	79	73	72	69	75	56	45	53	15	7	89
França	63	57	65	60	42	55	60	37	36	39	12	5	80
Alemanha	79	65	71	69	41	40	70	47	35	37	18	6	74
Grecia	38	33	38	33	5	26	31	14	11	12	6	3	45
Hungria	55	51	57	48	16	32	55	35	23	25	6	2	63
Irlanda	67	54	60	54	30	40	56	23	23	28	21	9	83
Italia	53	39	42	33	16	35	39	20	15	17	4	2	70
Letonia	58	50	61	50	42	38	54	28	22	23	4	2	65
Lituania	60	50	55	44	32	28	47	28	18	19	18	7	60
Luxemburgo	87	71	83	75	54	66	81	54	44	54	9	3	68
Malta	64	63	55	48	32	39	51	30	23	24	12	6	100
Países Baixos	90	77	86	79	73	34	85	50	50	55	22	8	79
Polonia	59	51	52	29	21	33	45	22	16	18	5	2	53
Portugal	48	46	42	40	17	40	40	28	18	21	16	3	100
Romenia	38	24	31	12	2	18	28	15	6	6	3	1	45
Eslovaquia	62	42	66	50	26	21	61	30	26	31	6	2	55
Eslovenia	64	56	58	49	24	38	53	32	31	32	11	7	95
Espanha	54	51	54	47	24	38	52	32	29	30	10	4	80
Suecia	86	79	86	77	71	45	83	36	48	57	21	9	95
Reino Unido	77	69	76	64	45	45	74	34	30	35	16	12	100

Tabela 2 – Dados usados para a análise da divisão digital (Porcentagem)

Pela análise da Tabela 2 evidenciam-se disparidades, dentro da UE-27, na adopção das TIC ao longo dos diferentes países. A Bulgária tem 30% de domicílios ligados à Internet e uma adopção no uso de serviços de e-banking situado por volta dos 2%. Estes valores contrastam com, por exemplo, os dos Países Baixos que apresentam valores de 90% e 73% para as mesmas variáveis, respectivamente. Outro exemplo das disparidades presentes nos dados verifica-se na disponibilidade de serviços públicos online. Existem quatro países (Áustria, Malta, Portugal e Reino Unido) com uma disponibilidade situada nos 100% e três (Bulgária, Grécia e Roménia) com menos de 50%. Estas distribuições altamente assimétricas não afectam a nossa análise, uma vez que os métodos multivariados usados não têm qualquer tipo de pressupostos sobre as distribuições das variáveis usadas. No entanto, estas são bastante elucidativas sobre a divisão que existe entre os países. Além disso, a dimensionalidade dos dados, com 13 dimensões, torna impossível abordar o tema da divisão digital com estatísticas univariadas sem que com isso tenhamos resultados incompletos.

## 4. Metodologia

### Análise Factorial

A análise factorial (AF) é um método que usa a estrutura de correlações entre as variáveis de forma a encontrar factores, ou dimensões, latentes dentro dos dados analisados. Para se utilizar uma AF com sucesso algumas suposições têm de ser assumidas e confirmadas. Como referido, o uso desta técnica depende da estrutura de correlação entre os dados de entrada. Logo, é necessário confirmar se essa correlação é suficientemente forte, caso contrário, a AF pode originar conclusões erróneas. A nossa análise envolveu várias etapas. Primeiro foi necessário analisar a estrutura de correlação dos dados, pela matriz de correlações. Em segundo lugar, a adequação dos dados foi testada pelo método de Kaiser-Mayer-Olkin (KMO). Terceiro, foi

necessário escolher o método de extracção a ser utilizado. Quarto, o número de factores a ser extraído foi definido e, finalmente, deu-se a interpretação dos factores com base nos *loadings*, ou seja, a correlação entre as variáveis originais e os factores latentes obtidos. A matriz de correlação (ver Tabela 3) mostra que cada variável tem, no mínimo, um coeficiente de correlação absoluto de 0,67 com uma outra variável. Este facto garante que as variáveis usadas dizem respeito ao mesmo fenómeno, a medição do desenvolvimento digital. Pela análise da mesma, verifica-se que um dos mais altos níveis de correlação (0,98) existe entre a percentagem da população que utiliza a Internet para obter informações das autoridades públicas (egovInf) e percentagem da população que usa regularmente a Internet para interagir com autoridades públicas (egovInt), o que indica que a pesquisa serviços públicos online origina a interação, também online, da população com as autoridades públicas. Por outro lado, a percentagem de serviços públicos disponíveis online (egovsup) apresenta um coeficiente de correlação de 0,51 e 0,53, com a percentagem da população que utiliza regularmente a Internet (IntPop) e a percentagem da população que utiliza correio electrónico (e-mail), respectivamente. Estes níveis de correlações indicam que, aparentemente, em alguns países a adopção das TIC pela população têm uma correlação linear moderada com os esforços dos países para incentivar o uso destas tecnologias.

	HsInt	HsBro	IntPop	IntSrc	ebank	elearn	email	ehealth	egovInf	egovInt	ecom	esafe	egovsup
HsInt	1	0,94	0,94	0,93	0,88	0,69	0,94	0,86	0,86	0,89	0,67	0,66	0,58
HsBro		1	0,92	0,93	0,91	0,73	0,92	0,84	0,88	0,89	0,72	0,7	0,67
IntPop			1	<b>0,95</b>	0,92	0,66	<b>0,99</b>	0,85	0,89	0,92	0,62	0,62	0,51
IntSrc				1	0,89	0,73	<b>0,95</b>	0,87	0,89	0,91	0,72	0,67	0,58
ebank					1	0,68	0,9	0,82	0,89	0,91	0,64	0,58	0,56
elearn						1	0,66	0,79	0,69	0,72	0,42	0,45	0,56
email							1	0,86	0,89	0,92	0,65	0,63	0,53
ehealth								1	0,83	0,85	0,52	0,43	0,53
egovInf									1	<b>0,98</b>	0,6	0,56	0,56
egovInt										1	0,6	0,56	0,56
ecom											1	0,85	0,58
esafe												1	0,58
egovsup													1

Tabela 3 – Matriz de Correlações

Para confirmar a adequação dos dados para (AF), foi usada a medida de KMO. Esta apresenta o valor de 0,85 que expressa a boa adequabilidade dos dados à AF (S. Sharma, 1996).

Como método de extracção dos factores foi aplicada a AF em componentes principais. A AF em componentes principais tem como origem uma abordagem empírica, onde os factores extraídos são assumidos como combinações lineares das variáveis usadas. Considerando que o nosso objectivo é reduzir a complexidade e dimensionalidade da divisão digital, tornou-se necessário decidir quantos factores deveríamos extrair na nossa análise. Não existe uma regra óptima para fazê-lo. Existem, no entanto, três grandes critérios para alcançar o número de factores; os métodos de Pearson, Kaiser e “Scree Plot” foram tidos em consideração (Peres-Neto, et al., 2005). Todos eles foram unânimes: o número ideal de factores a ser extraído é dois. Tal como apresentado na Tabela 4, a variância retida nestes dois factores corresponde a 87% do total presente nas 13 variáveis originais. Isto significa que embora tenhamos alterado o número de indicadores na análise de 13 para dois, apenas 13% da informação foi perdida. O passo seguinte relacionou-se com a rotação dos factores de forma a atingir uma melhor divisão das variáveis originais nas duas dimensões extraídas. Embora existam vários tipos de rotação, incluindo rotações ortogonais e oblíquas, as ortogonais tendem a ser as mais utilizadas (S. Sharma, 1996). Em particular, foi aplicada a rotação Varimax. No entanto, as rotações Varimax e Quartimax

levaram a resultados semelhantes. Finalmente, para medir a fiabilidade da escala de cada factor, a medida de Alfa de Cronbach foi também calculada. Esta mede a consistência interna de cada factor dentro de si mesmo. Nunnally (1978) sugere que um valor acima de 0,7 é considerado adequado. Os valores calculados são de 0,99 e 0,86 para o factor 1 e factor 2, respectivamente, o que indica que os factores extraídos têm uma elevada fiabilidade.

<b>Análise Factorial</b>	<b>Factor 1</b>	<b>Factor 2</b>
ehealth	<b>0,91</b>	0,22
egovint	<b>0,90</b>	0,35
IntPop	<b>0,88</b>	0,39
egovinf	<b>0,88</b>	0,35
email	<b>0,87</b>	0,41
ebank	<b>0,85</b>	0,40
IntSrc	<b>0,85</b>	0,48
HsInt	<b>0,84</b>	0,47
HsBro	<b>0,81</b>	0,54
elearn	<b>0,75</b>	0,25
esafe	0,28	<b>0,90</b>
ecom	0,33	<b>0,88</b>
egovsup	0,38	<b>0,66</b>
Variância Explicada (%)	59%	28%
Variância Explicada Acumulada (%)	59%	87%
Alpha de Cronbach	0,99	0,86

Tabela 4 – Resultados da Análise Factorial (Rotação Varimax)

O factor 1 é caracterizado por uma alta densidade na utilização das TIC pela população, juntamente com a disponibilidade da sua infra-estrutura. Logo, o factor 1 expressa a “*adopção e difusão das TIC na população*”.

No factor 2 temos expressa a capacidade das empresas para fornecer conexões seguras, a percentagem de empresas que vendem online e a disponibilidade de serviços públicos online. Logo, o factor 2 está relacionado com o “*comércio, segurança e governo electrónicos*”.

Os “scores” de cada factor para cada país foram calculados para quantificar o nível de cada país nas duas dimensões extraídas (ver Figura 1). O gráfico mostra que os países escandinavos - Suécia, Países Baixos, Dinamarca e Finlândia - são os que têm melhor posição em ambos os factores. O Luxemburgo é o país com maiores níveis na *adopção e difusão das TIC na população* mas, também simultaneamente, um dos países com menores níveis no *comércio, segurança e governo electrónicos*. O Reino Unido é o país com o nível mais elevado no factor 2, “*comércio, segurança e governo electrónicos*”, o que confirma a teoria de que é o país da Europa com níveis mais elevados na disponibilidade de serviços de governo electrónico (Unwin & de Bastion, 2009). Os países do leste europeu, como a Roménia e a Bulgária, são os que se encontram num estado menos desenvolvido em relação ao nível de desenvolvimento digital.

Finalmente, notamos que alguns países apresentam coordenadas muito próximas, indicando perfis digitais semelhantes, como é o caso da Roménia e da Bulgária; Lituânia, República Checa e Portugal; Itália e Chipre ou da França e Estónia. O facto de que estes pares de países não estão necessariamente próximos do ponto geográfico, pode evidenciar a presença de outros factores, como por exemplo, de natureza cultural ou política.

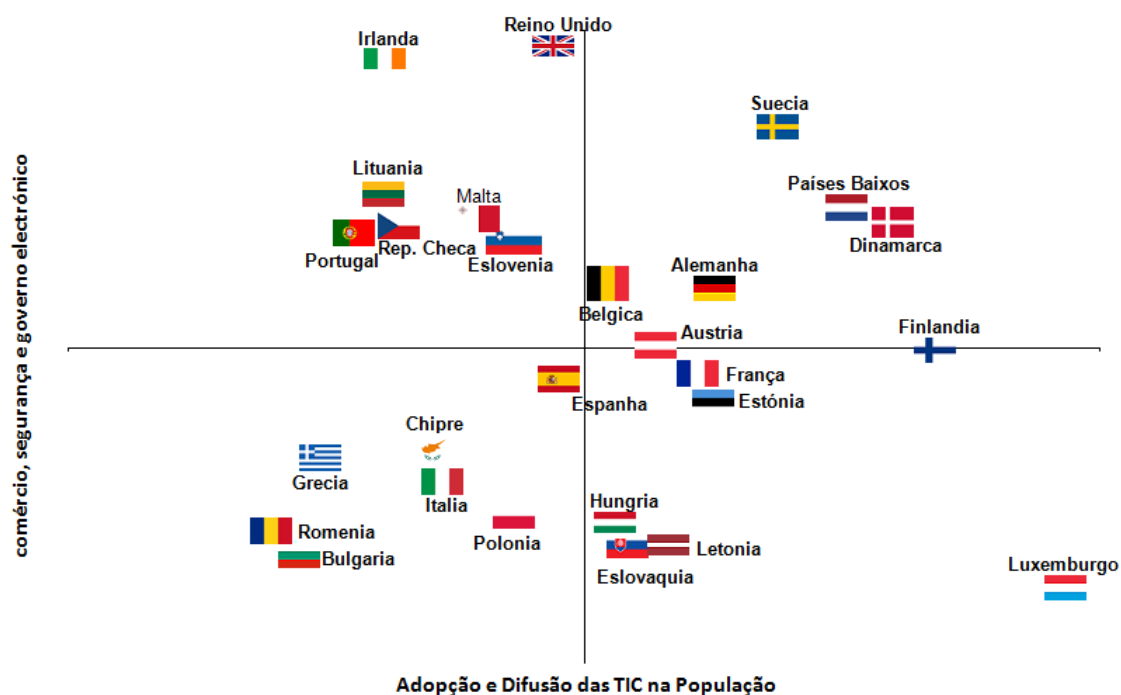


Figura 1 – Posição dos Países nas dimensões extraídas

### Análise de Clusters

Após o uso da AF em que encontramos duas dimensões latentes sobre a exclusão digital – foi realizada uma análise de clusters (AC) de forma a agrupar os países por um critério de similaridade, tanto para os factores extraídos como para as 13 variáveis originais. O uso da AC envolve dois métodos principais, os hierárquicos e os não hierárquicos. A metodologia utilizada para o clustering com base nos factores e nas 13 variáveis originais foi semelhante. Primeiro realizámos uma AC com métodos hierárquicos para definir o número de clusters a extrair dos dados, já que neste procedimento este número depende dos dados, o que significa que não precisamos de definir, *a priori*, quantos clusters queremos gerar. A solução gerada pelos procedimentos hierárquicos depende do tipo de algoritmo, bem como do tipo de distância usados. Em particular, usámos os métodos do vizinho mais próximo (Single), vizinho mais distante (Complete), Centroid e Ward. A solução foi escolhida com base na sua performance, ou seja, com base na análise do  $R^2$  e do dendrograma. De seguida, a melhor solução obtida pelos métodos hierárquicos foi utilizada para gerar as sementes iniciais dos métodos não hierárquicos (k-means), segundo Sharma (1996) esta é a melhor solução para realizar a AC. Finalmente os clusters gerados por este algoritmo foram classificados através de uma análise de perfil, ou seja, através dos valores médios de cada variável ou factor, no respectivo grupo.

### Análise de Clusters com base nos Factores Extraídos

A solução dos métodos hierárquicos da AC sobre os factores é dada pelo dendrograma (ver Figura 3 em anexo), em que o eixo vertical representa a distância e o eixo horizontal os países da UE-27. De baixo para cima, o dendrograma mapeia o agrupamento de países. O primeiro grupo de países é formado pela Eslováquia e Letónia, o que significa que estes países são os mais semelhantes, do ponto de vista do perfil digital, dentro UE-27. A Roménia e a Bulgária são os próximos países a serem agrupados, seguidos pela Estónia e França. À medida que o algoritmo continua todos os países são colocados em clusters. O número de clusters é definido quando o dendrograma faz o primeiro “salto”, que neste caso, corresponde à solução de cinco clusters. O gráfico do  $R^2$  (ver Figura 5 em anexo) apresenta, também, a mesma solução para o



número de clusters a ser usado nos métodos não hierárquicos (k-means). Ainda pela análise do dendograma evidencia-se que o Luxemburgo é o país europeu com um perfil digital único, sendo o último a juntar-se aos restantes países europeus. Todos os métodos hierárquicos realizados apresentaram este resultado. O algoritmo k-means, com as sementes iniciais do método do vizinho mais distante (Complete), teve como resultado a mesma solução final apresentada pelo dendograma, o que significa que, neste caso, os métodos hierárquicos e não hierárquicos foram unânimes na composição dos clusters. Assim temos a Dinamarca, Finlândia, Países Baixos e Suécia a formar o cluster um. O cluster dois é formado pelo Luxemburgo sozinho. O terceiro grupo é consistido pela Alemanha, Áustria, Bélgica, Estónia, França, Hungria, Letónia, Eslováquia e Espanha. Em seguida, temos o cluster quatro com a República Checa, Irlanda, Lituânia, Malta, Portugal, Eslovénia e Reino Unido. Finalmente temos a Bulgária, Chipre, Grécia, Itália, Polónia e Roménia agrupados no cluster cinco. Esta solução pode ser visualizada na Figura 2, onde os clusters obtidos pelo k-means estão representados nas duas dimensões da AF.

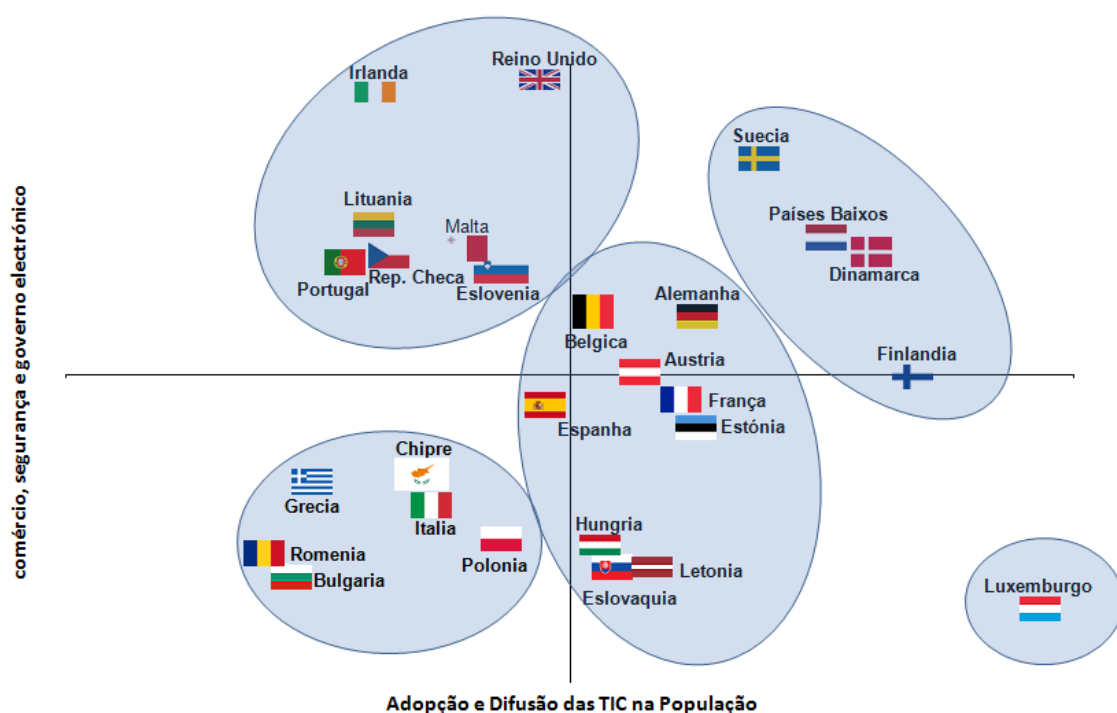


Figura 2 – Análise de Clusters sobre os Factores Estimados

Através de uma análise de perfil de cada cluster (Tabela 5), temos que: O cluster um representa os países com melhor performance no conjunto dos dois factores dentro da UE-27, em termos de desenvolvimento digital. Este cluster é composto por países com altos níveis de desenvolvimento digital em ambas as dimensões latentes. Tal como o desenvolvimento económico, o desenvolvimento digital deve ser equilibrado e homogéneo, pelo que, efectivamente, se a Europa quer ver os objectivos da agenda digital e da Estratégia 2020 alcançados, todos os outros países europeus devem ter como objectivo atingir os níveis de desenvolvimento digital destes países. O cluster dois, Luxemburgo, apresenta o nível mais elevado na *adopção e difusão das TIC na população* e, simultaneamente, o nível mais baixo no *comércio, segurança e governo electrónicos*. Contrariamente ao cluster um, o Luxemburgo, tem um desenvolvimento digital altamente desequilibrado e heterogéneo sobre as dimensões latentes do desenvolvimento digital. Esta situação, no entanto, não é nova considerando outros estudos no passado que apresentaram resultados semelhantes para este país (Cuervo & Menéndez, 2006). Uma das explicações para esse facto pode residir nas características únicas, como tamanho e número populacional, que o Luxemburgo apresenta. O terceiro cluster é, juntamente

com o cluster um, um grupo bastante homogéneo e equilibrado, apesar de apresentar níveis de desenvolvimento digital mais baixos. É o cluster que melhor representa a média da totalidade dos países europeus. Além disso, é o cluster com um maior número de países (8). No entanto, apesar da presença de Eslováquia, Hungria e Letónia neste cluster, estes países apresentam níveis mais baixos para o *comércio, segurança e governo electrónicos* do que os restantes membros. O quarto cluster inclui os países com baixos níveis sobre a *adoção e difusão das TIC na população* e, simultaneamente, altos níveis no *comércio, segurança e governo electrónicos*. O Reino Unido é o país com maior nível nesta dimensão dentro de toda a UE-27. Considerando as variáveis que compõem esta dimensão parece que, apesar da população não usar de forma intensiva as TIC, os governos destes países, juntamente com as empresas, estão a fazer esforços significativos para a implementação das infra-estruturas necessárias para alcançar uma sociedade baseada na informação. Será interessante ver, no futuro, se esta estratégia irá resultar ou não. No entanto, ao contrário do que se sucede no Luxemburgo, o desenvolvimento digital tem de ser homogéneo e equilibrado em todas as suas dimensões. Finalmente o cluster cinco contém os países digitalmente menos desenvolvidos na UE-27. Estes apresentam os níveis mais baixos para ambas as dimensões. Dentro deste grupo temos a Roménia e a Bulgária como os países digitalmente com maior grau de subdesenvolvimento.

Cluster	Factor 1		Factor 2	
	Média	D. P.	Média	D. P.
1	0,33	0,27	-0,38	0,66
2	-0,98	0,47	-0,95	0,26
3	-0,70	0,37	1,08	0,54
4	1,36	0,33	0,77	0,58
5	2,37	-	-1,52	-

Tabela 5 – Estatísticas Descritivas para os Clusters Extraídos (Factores)

### Análise de Clusters com base nas Variáveis Originais

Depois de usadas as duas dimensões extraídas na AF para a identificação de grupos de países homogéneos, em termos de perfis digitais, dentro UE-27, foram usadas as 13 variáveis originais. Uma vez mais foram usados métodos hierárquicos para definir o número de clusters e não-hierárquicos para obter a solução final.

Pela análise do dendograma (ver Figura 4 em anexo) verifica-se que a Bulgária e a Roménia são os primeiros países a formar um cluster, seguidos pela Polónia, Chipre e República Checa. O primeiro “salto” no dendograma ocorre na solução de cinco clusters. O gráfico do  $R^2$  (ver Figura 6 em anexo) confirma esta solução. Assim foi utilizado o algoritmo k-means com as sementes iniciais de cinco clusters obtidas pelo método de Ward (melhor solução hierárquica). A solução obtida foi a seguinte: Dinamarca, Finlândia, Luxemburgo, Países Baixos e Suécia agrupados no cluster A. Em seguida temos a Áustria, a Bélgica, a Estónia, a França, a Alemanha e o Reino Unido reunidos no cluster B. O cluster C é formado pela Irlanda, Malta, Portugal, Eslovénia e Espanha. Chipre, República Checa, Hungria, Itália, Letónia, Lituânia, Polónia e Eslováquia formam o cluster D. Finalmente o cluster E é dado pela Bulgária, Grécia e Roménia.

Através da análise das médias de cada variável em cada cluster obtido (ver Tabela 6) verifica-se que: O cluster A tem níveis médios superiores em 12 das 13 variáveis incluídas. A única excepção encontra-se na percentagem de serviços públicos online. Este cluster é formado por todos os países que estavam anteriormente no cluster um, a partir da análise de clusters, baseada nos factores e pelo Luxemburgo (cluster dois). Este é o grupo de países que mais se destaca de todos os outros formados. O cluster B tem os segundos maiores níveis para todas as variáveis, com excepção da percentagem de empresas que vende online (ecom). O cluster C apresenta valores médios para a maioria das variáveis. No entanto, a percentagem de empresas que vende online nestes países é a segunda maior dentro de toda a UE. Além disso, é o grupo de países com a maior percentagem de serviços públicos online. Os países que formam o cluster D

apresentam o segundo valor médio mais baixo em todas as variáveis, excepto para a percentagem da população que utiliza regularmente a Internet (IntPop). Finalmente, temos o Cluster E que, sem surpresa, é composto pelos países com valores mais baixos para todas as variáveis. As mudanças na estrutura dos clusters são devidas à perda de variância e informação na AF (13%).

Cluster	Cluster A		Cluster B		Cluster C		Cluster D		Cluster E	
	Média	D. P.	Média	D. P.	Média	D. P.	Média	D. P.	Média	D. P.
HsInt	84,80	4,5	69,83	6,9	59,40	8,0	56,75	3,5	35,33	4,6
HsBro	75,40	3,0	62,33	4,5	54,00	6,3	47,38	4,5	27,67	4,7
IntPop	83,20	2,9	69,33	3,9	53,80	7,0	54,00	7,9	36,33	4,7
IntSrc	75,60	2,4	60,00	5,8	47,60	5,0	42,88	8,3	20,67	11,0
ebank	67,20	7,9	45,17	9,1	25,40	5,9	23,25	9,6	3,00	1,7
elearn	54,40	14,7	40,00	8,9	39,00	1,0	29,88	5,6	19,33	6,1
email	81,00	3,7	66,17	5,4	50,40	6,1	49,25	8,3	31,00	3,0
ehealth	48,52	8,0	36,70	5,6	29,36	3,6	25,29	6,4	13,40	3,1
egovinf	50,40	8,5	34,33	5,5	24,80	5,2	20,50	3,8	8,33	2,5
egovint	57,20	5,7	37,50	4,4	27,00	4,5	22,38	4,5	9,33	3,1
ecom	17,20	5,3	13,83	3,3	14,00	4,5	8,13	5,3	4,00	1,7
esafe	7,40	2,7	6,00	3,2	5,80	2,4	3,38	2,3	1,67	1,2
govsup	83,00	10,3	85,67	13,0	91,60	9,5	59,50	6,6	43,33	2,9

Tabela 6 – Estatísticas Descritivas para a AC (Solução de cinco clusters)

As mudanças na estrutura de clusters, mais evidentes, dizem respeito ao Luxemburgo, Reino Unido e à separação da Bulgária, Grécia e Roménia, que formaram o cluster com níveis de subdesenvolvimento digital mais elevados.

## 5. Conclusões e Limitações do Estudo

Aparentemente o desenvolvimento digital, medido pelos indicadores incluídos na nossa análise, inclui duas dimensões latentes, sendo estas a “*adoção e difusão das TIC na população*” e o “*comércio, segurança e governo electrónicos*”. Estas duas dimensões são independentes uma da outra, considerando que um país pode, simultaneamente, apresentar níveis bastante elevados para uma delas, e níveis bastante reduzidos para a outra. Adicionalmente também foram encontrados cinco grupos de países na UE com diferentes níveis de desenvolvimento digital.

Algumas acções políticas podem ser propostas a partir dos resultados da nossa análise, são: O facto de alguns países apresentarem níveis de desenvolvimento digital numa das dimensões extraídas e, ao mesmo tempo, níveis de desenvolvimento bastante reduzidos na outra, tem que ser encarado como um desafio crítico, visto que o desenvolvimento digital, tal como o económico, tem que ser homogéneo e transversal a todas as suas dimensões. Outra situação que deverá ser resolvida está relacionada com a enorme diferença entre os países digitalmente mais desenvolvidos com os menos desenvolvidos. Ou seja, se compararmos a situação da Bulgária e Roménia com a da Suécia ou Países Baixos, é perceptível o grande caminho que aqueles países ainda têm que percorrer. Este caminho só poderá ser finalizado com sucesso se forem aplicadas medidas estruturais e multifacetadas como, por exemplo, criando as infra-estruturas necessárias para as TIC e fomentando a adopção destas pela população, entidades públicas e empresas.

Apesar do nosso esforço para oferecer uma análise completa e multidimensional, algumas limitações devem ser, no entanto, consideradas: A análise incidiu sobre dados do ano de 2009, pelo que futuras alterações susceptíveis de acontecer em poucos anos, não foram consideradas. Devido a um número limitado de variáveis passíveis de serem incluídas, neste caso 13, existe a possibilidade de alguns aspectos do desenvolvimento digital não estarem incluídos. Finalmente foram analisadas as disparidades digitais dentro da UE com base em indicadores agregados às respectivas realidades nacionais, o que significa que disparidades domésticas não foram consideradas.

## 6. Referências

- Bakker, A. R. (2002). Health care and ICT, partnership is a must. *International Journal of Medical Informatics*, 66(1-3), 51-57.
- Bertot, J. C. (2003). The Multiple Dimensions of the Digital Divide: More than the Technology "Haves" and "Haves Nots". *Government Information Quarterly*, 20, 191.
- Brooks, S., Donovan, P., & Rumble, C. (2005). Developing Nations, the Digital Divide and Research Databases. [doi: DOI: 10.1016/j.serrev.2005.09.002]. *Serials Review*, 31(4), 270-278.
- Castells, M. (2007). *The Network Society* (Vol. III). Lisbon: Fundação Calouste Gulbenkian.
- Castells, M., & Himanen, P. (2007). *A Sociedade da Informação e o Estado Providência - O Modelo Finlandês* (T. d. M. Soares, Trans. 1 ed.). Lisboa: Fundação Calouste Gulbenkian.
- Chinn, M. D., & Fairlie, R. W. (2007). The determinants of the global digital divide: a cross-country analysis of computer and internet penetration. [Article]. *Oxford Economic Papers-New Series*, 59(1), 16-44.
- Chinn, M. D., & Fairlie, R. W. (2010). ICT Use in the Developing World: An Analysis of Differences in Computer and Internet Penetration. *Review of International Economics*, 18(1), 153-167.
- Çılan, Ç. A., Bolat, B. A., & Coskun, E. (2009). Analyzing digital divide within and between member and candidate countries of European Union. [doi: DOI: 10.1016/j.giq.2007.11.002]. *Government Information Quarterly*, 26(1), 98-105.
- Cuervo, M. R. V., & Menéndez, A. J. L. (2006). A multivariate framework for the analysis of the digital divide: Evidence for the European Union-15. [doi: DOI: 10.1016/j.im.2006.05.001]. *Information & Management*, 43(6), 756-766.
- Cukusic, M., Alfirovic, N., Granic, A., & Garaca, Z. (2010). e-Learning process management and the e-learning performance: Results of a European empirical study. *Computers & Education*, 55(2), 554-565.
- Dragulanescu, N.-G. (2002). Social Impact of the "Digital Divide" in a Central-Eastern European Country. *The International Information & Library Review*, 34(2), 139-151.
- European Comission (2006). *Bridging the Broadband Gap*. Brussels.
- European Comission (2010a). *A Digital Agenda for Europe*. from [http://ec.europa.eu/information\\_society/digital-agenda/index\\_en.htm](http://ec.europa.eu/information_society/digital-agenda/index_en.htm).
- European Comission (2010b). *Europe 2020 - A strategy for smart, sustainable and inclusive growth*. from [http://europa.eu/press\\_room/pdf/complet\\_en\\_barroso\\_007\\_-\\_europe\\_2020\\_-\\_en\\_version.pdf](http://europa.eu/press_room/pdf/complet_en_barroso_007_-_europe_2020_-_en_version.pdf).
- Hargittai, E. (1999). Weaving the Western Web: explaining differences in Internet connectivity among OECD countries. *Telecommunications Policy*, 23(10-11), 701-718.
- J. C. Nunnally (1978). *Psychometric Theory*. New York: McGraw-Hill.
- Kiiski, S., & Pohjola, M. (2002). Cross-country diffusion of the Internet. *Information Economics and Policy*, 14(2), 297-310.
- OECD (2001). *Understanding the Digital Divide*, 32. Paris: OECD Publications.
- OECD (2009). *Guide to Measuring the Information Society*. Paris: Secretary-General of the OECD.
- Okazaki, S. (2006). What do we know about mobile Internet adopters? A cluster analysis. *Information & Management*, 43(2), 127-141.
- Ono, H., & Zavodny, M. (2007). Digital inequality: A five country comparison using microdata. *Social Science Research*, 36(3), 1135-1155.
- Peres-Neto, P. R., Jackson, D. A., & Somers, K. M. (2005). How many principal components? stopping rules for determining the number of non-trivial axes revisited. [doi: DOI: 10.1016/j.csda.2004.06.015]. *Computational Statistics & Data Analysis*, 49(4), 974-997.

Pohjola, M. (2003). The Adoption and Diffusion of ICT Across Countries: Patterns and Determinants. *The New Economy Handbook*, Academic Press, 230–247.

S. Sharma (1996). *Applied Multivariate Techniques*. New York: John Wiley & Sons, Inc.

Selwyn, N., & Facer, K. (2007). Beyond the digital divide: Rethinking digital inclusion for the 21st century: Futurelab.

The World Bank (2006). Information and Communications for Development: Global Trends and Policies. Washington DC: The World Bank.

Unesco, I. f. S. (2003). Measuring and monitoring the information and knowledge societies: a statistical challenge. Retrieved 23 May 2010, from [http://portal.unesco.org/ci/en/ev.php-URL\\_ID=12851&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/ci/en/ev.php-URL_ID=12851&URL_DO=DO_TOPIC&URL_SECTION=201.html)

Unwin, T., & de Bastion, G. (2009). Digital Divide. In K. Rob & T. Nigel (Eds.), *International Encyclopedia of Human Geography* (pp. 191-197). Oxford: Elsevier.

Vehovar, V., Sicherl, P., Husing, T., & Dolnicar, V. (2006). Methodological Challenges of Digital Divide Measurements. *The Information Society*, 22, 12.

Vicente, M. R., & Gil-de-Bernabé, F. (2010). Assessing the broadband gap: From the penetration divide to the quality divide. *Technological Forecasting and Social Change*, 77(5), 816-822.

Vicente, M. R., & Lopez, A. J. (2010). A Multidimensional Analysis of the Disability Digital Divide: Some Evidence for Internet Use. *The Information Society: An International Journal*, 26(1), 48 - 64.

Warschauer, M. (2002). Reconceptualizing the Digital Divide. *First Monday*, 7 (7), on-line.

WSIS (2003). *World Summit on the Information Society: Declaration of Principles*. Paper presented at the World Summit on the Information Society. Retrieved 27 May 2010, from [http://www.itu.int/wsis/documents/doc\\_multi.asp?lang=en&id=11610](http://www.itu.int/wsis/documents/doc_multi.asp?lang=en&id=11610)

WSIS (2005). *Tunis Commitment*. Paper presented at the Second Phase of the World Summit on the Information Society. Retrieved 27 May 2010, from [http://www.itu.int/wsis/documents/doc\\_multi.asp?lang=en&id=22660](http://www.itu.int/wsis/documents/doc_multi.asp?lang=en&id=22660)

## Anexos

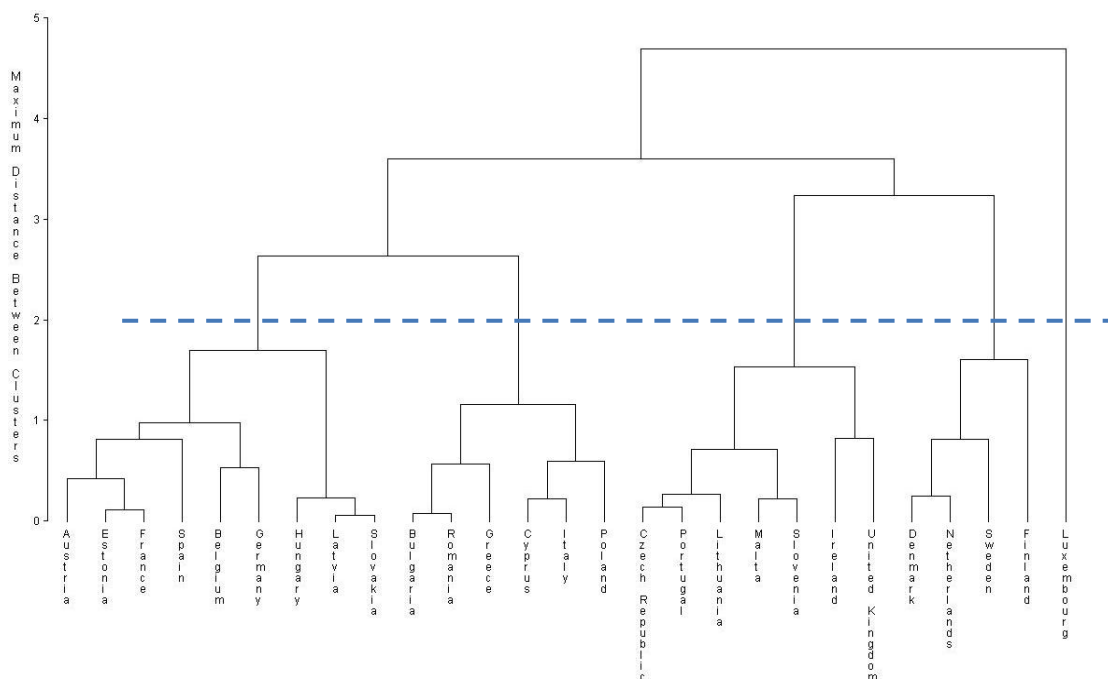


Figura 3 – Dendrograma do Método vizinho mais distante (Complete) para a AC sobre os Factores

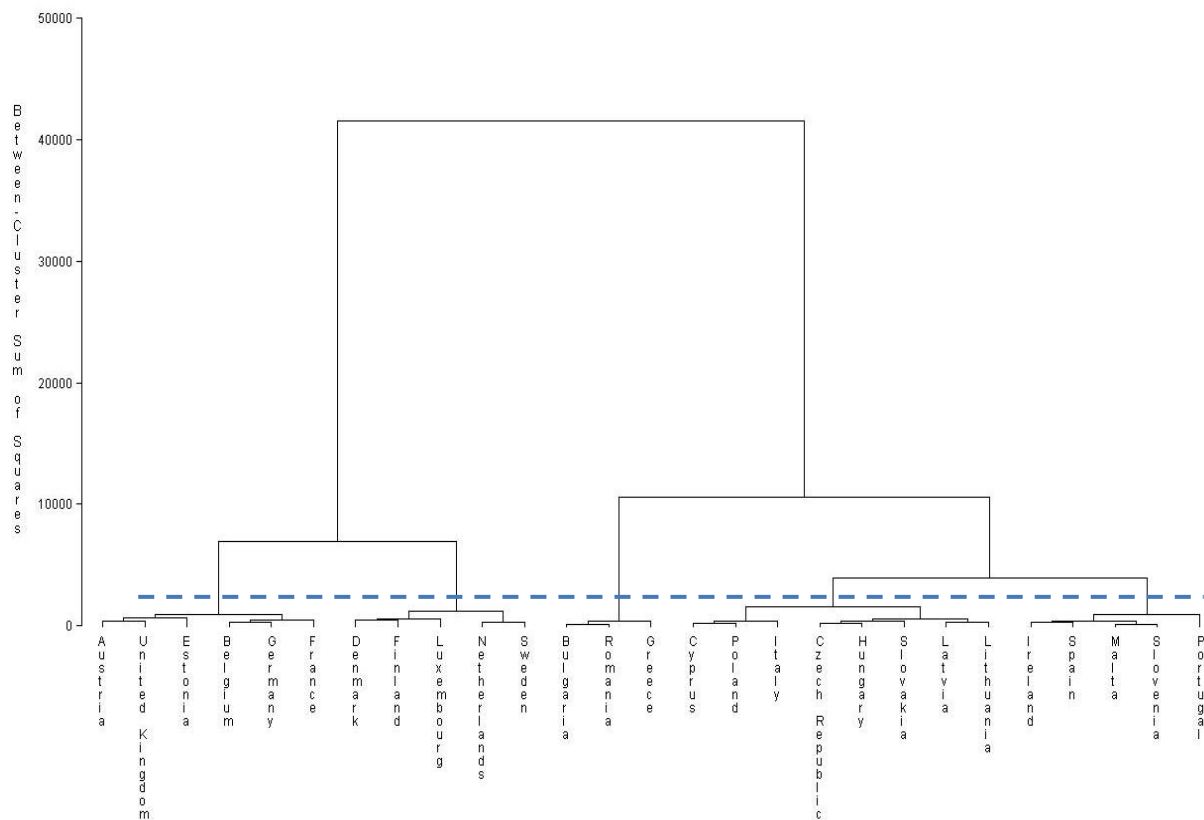


Figura 4 – Dendrograma do Método Ward para a AC sobre as Variáveis

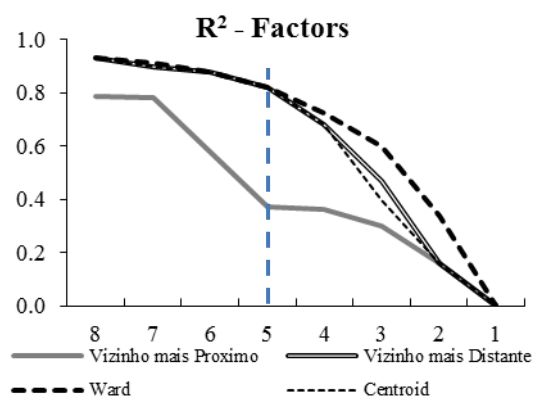


Figura 5 – Gráfico do  $R^2$  para a AC sobre os Factores Extraídos

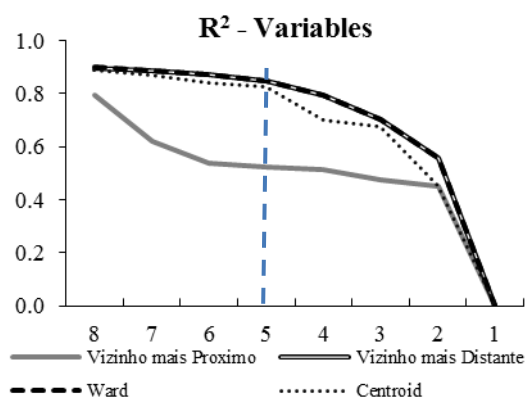


Figura 6 – Gráfico do  $R^2$  para a AC sobre as Variáveis Originais