



INSTITUTO
UNIVERSITÁRIO
DE LISBOA

2ºSemestre Ano Letivo 2022/2023

Unidade Curricular de Métodos de Aprendizagem Não Supervisionada

Docente Jorge Sinval

Trabalho Final – PISA 2018

2º ano - LCD – CDB1

Eliane Susso Efraim Gabriel, Nº 103303

João Francisco Marques Gonçalves da Silva Botas, Nº104782

Marco Delgado Esperança, Nº 110451

Maria João Ferreira Lourenço, Nº 104716

Umeima Adam Mahomed, Nº99239

Lisboa, 15 de junho de 2023

Índice

1.	Introdução	3
2.	A educação na Estónia.....	4
3.	Objetivo da análise	6
4.	Tratamento dos dados	6
	Variáveis de INPUT	7
	Variáveis de PROFILE.....	7
5.	Identificação das dimensões da análise	8
6.	PCA`s caracterização	8
7.	Conclusões retiradas das componentes	10
8.	Análise dos PCA`s.....	11
9.	Identificação da heterogeneidade na base de dados	11
10.	Conclusões obtidas com o clustering	12
11.	Análise com imputação de dados omissos.....	14
12.	Conclusão	15
13.	Agradecimentos	17
14.	Referências	Erro! Marcador não definido.
15.	Anexos.....	19
	ANEXO A - Descrição das variáveis mais relevantes:	19
	ANEXO B - Países e Economias Participantes no PISA 2018	22
	ANEXO C -Imagem ilustrativa do ranking da Estónia	23
	ANEXO D - Education Lifelong Learning strategy 2020	23
	ANEXO E – Alguns pontos chave da entrevista com Gunda Tire	24
	ANEXO F – Matriz de correlação entre variáveis numéricas para a Estónia	26
	ANEXO G – Dados sobre <i>Bullying</i> na Estónia face a média dos países da OCDE	27
	ANEXO H – Gráfico ilustrativo da presença de <i>Bullying</i> na Europa em 2018.....	27
	ANEXO I – Rating da Estónia face outros países incluídos no PISA 2018.....	28
	ANEXO J – Gráficos diversos.....	29
	ANEXO K – Scree plot para o clustering	29
	ANEXO L – Tabelas alusivas a cada cluster e algumas variáveis de PROFILE	30
	ANEXO M –PAM <i>clustering</i>	31
	ANEXO N – Gráficos de Classificação e Incerteza	32

1. Introdução

O PISA (*Programme for International Student Assessment*), promovido pela OCDE (Organização para a Cooperação e Desenvolvimento Económico), é um estudo internacional que procura medir os níveis de desempenho na educação de jovens estudantes de 15 anos, sendo um dos programas internacionais mais influentes e reconhecidos a nível global. A sua abordagem visa avaliar não apenas as competências dos estudantes em três áreas principais: leitura, matemática e ciências, como também a capacidade que eles têm de aplicar os conhecimentos adquiridos nos desafios atuais da sociedade dentro e fora do contexto escolar. A recolha de informações desta organização é feita tendo em conta o ambiente educacional, fatores socioeconómicos, atitudes dos estudantes em relação à aprendizagem e os seus métodos de ensino.

O Programa Internacional de Avaliação de estudantes tem desempenhado um papel muito importante no que diz respeito ao desenvolvimento de políticas de educação e melhoria dos resultados educacionais em todas as partes do mundo. Esta iniciativa foi concebida em 2000 e ocorre a cada três anos. Em cada ciclo, uma área de conhecimento específica é enfatizada, com a maioria das questões aplicadas direcionadas a essa área.

Inicialmente, o PISA envolvia os 28 países membros da OCDE, mas devido aos seus interesses de avaliação, em 2001, mais países não pertencentes à OCDE juntaram-se à segunda fase de avaliação do PISA. Desde então, o número de países e economias participantes tem variado em cada ciclo, visto que novos países podem aderir ao programa ou alguns países membros podem optar por não participar em determinadas edições. Na edição do [PISA 2018](#), por exemplo, a literacia de leitura foi o domínio principal avaliado e participaram perto de 79 países e economias, sendo que destes 42 são países/economias parceiras. O foco deste relatório, no entanto, será na análise do desempenho da Estónia, que é um país membro da OCDE desde 2010.¹

Nos últimos anos, a Estónia tem-se destacado como um exemplo notável de sucesso no PISA. Este país tem passado por uma transformação significativa investindo de forma consistente na sua educação com o objetivo de promover a inovação, igualdade e inclusão, desde a sua independência, em 1991.

¹ Verificar [anexo B](#) para uma visualização num mapa.

Desde a sua primeira participação no PISA, em 2006, a Estónia tem alcançado resultados muito relevantes nas três áreas principais, estando sempre entre os melhores países avaliados. Em particular, o mesmo tem-se destacado na área de leitura com um alto desempenho em comparação com a média da OCDE. É de notar que o bom desempenho do país em estudo no PISA é resultado de um esforço coletivo que envolve não apenas o sistema educativo, como também fatores sociais, económicos e culturais. A dedicação dos professores, o envolvimento das famílias e o valor que a sociedade dá à educação, desempenham um papel de suma importância na performance dos estudantes estonianos.

A cultura da Estónia é de aprendizagem contínua e criticismo positivo (mentalidade “*could be better*”) e a responsabilidade da educação está sobre todos o que gera um melhor desempenho. Este país não adaptou o seu sistema de educação para impulsionar os resultados do PISA, mas foram os resultados do PISA que ajudaram a apoiar as medidas políticas adotadas pelo mesmo². São diversos os fatores que contribuem para o sucesso deste país no PISA e serão estudados mais a fundo no desenvolvimento deste projeto.

2. A educação na Estónia

Na Estónia, o sistema educativo é baseado principalmente em escolas públicas, com apenas cerca de 10% de escolas privadas. A confiança dos cidadãos no sistema público de ensino é geralmente elevada, e os pais estonianos não precisam de se preocupar em encontrar uma boa escola para os seus filhos, porque acredita-se que uma educação de qualidade está disponível em todo o país, inclusive nas escolas públicas.

Embora existam escolas públicas mais populares, especialmente em cidades como Tallinn e Tartu, onde mais pais desejam que os seus filhos estudem, algumas dessas escolas populares podem utilizar testes de aptidão como parte do processo de seleção, mas a aceitação na escola não está relacionada com o estatuto social dos pais, mas sim com os resultados dos testes. Isso significa que os alunos de diferentes origens socioeconómicas frequentam as mesmas escolas, o que promove a diversidade e a igualdade de oportunidades educacionais.

² No seguinte [link](#) pode-se encontrar mais em concreto as perguntas que foram feitas aos participantes do Estudo do PISA 2018.

Os currículos escolares da Estónia ³ são baseados no princípio de que os alunos devem ter uma ampla visão de mundo. Além de leitura, matemática e ciências, muitas escolas ensinam programação e robótica desde o primeiro ciclo. Ao mesmo tempo, o currículo inclui disciplinas criativas como música, arte e educação física como disciplinas obrigatórias para todos. No âmbito da educação física, a escola pode optar por ensinar dança a todos os alunos ou esquiar na floresta, caso haja neve. Todas as escolas têm uma aula de tecnologia onde ensinam seus filhos a cozinhar em cozinhas bem equipadas e tricotar uma meia ou cortar madeira. Normalmente, todas essas habilidades práticas são dominadas por meninos e meninas em grupos mistos. Fazer um prato, onde todos são responsáveis por determinados ingredientes e fazem sua parte durante o preparo, é um bom exemplo de um típico projeto comum que envolve habilidades como cooperação, resolução de problemas e todos são orientados para um objetivo.

Esta abordagem reforça a ideia de que a qualidade da educação não está necessariamente ligada ao setor privado ou ao escalão financeiro. Embora os alunos estonianos com vantagens socioeconômicas tenham superado os alunos desfavorecidos em 61 pontos no PISA 2018, a discrepância é menor do que a diferença média entre os dois grupos (89 pontos) nos países da OCDE. Em vez disso, a ênfase está na excelência educativa e nos elevados padrões académicos, e, por meio desse histórico, currículos exigentes e exames de alta qualidade diretamente integrados no currículo, a Estónia construiu um sistema educacional de alto desempenho. ⁴

Este entendimento pode ser aplicado no desenvolvimento de estratégias educativas e na promoção de escolas públicas de qualidade noutras regiões, o que pode desmistificar a noção de que somente escolas privadas podem fornecer uma educação prestigiosa. Isto pode ajudar a melhorar a confiança no sistema público de ensino e a garantir que todas as crianças, independentemente da sua situação socioeconómica, tenham acesso a uma educação da mesma qualidade e sejam oferecidas as mesmas oportunidades de aprendizagem. ⁵

³ Para mais informação sobre o currículo da Estónia, [visitar este link](#).

⁴ [Anexo](#) referente a políticas em vigor em 2018 – Lifelong Learning Strategy 2020

⁵ Neste [link](#), encontra-se uma análise sobre os resultados do PISA 2018 na Estónia.

3. Objetivo da análise

O Ministério da Educação da Estónia pediu à equipa responsável pelo estudo do PISA 2018 que caracterizasse os alunos que participaram no estudo de forma a determinar os principais fatores que levaram ao bom sucesso académico dos alunos, permitindo-lhes chegar aos excelentes resultados obtidos no PISA 2018, apenas ultrapassados pela China. Este relatório contém a compilação dos resultados mais interessantes e de todos os passos dados para conseguirmos chegar a conclusões práticas.

E, portanto, realizar-se-à uma análise mais aprofundada dos dados do PISA 2018, focada especificamente nos alunos estonianos e na construção de perfis que ajudarão a compreender os fatores que levaram ao sucesso académico excecional.

Além disso, serão explorados os pontos fortes do sistema educativo da Estónia, a fim de identificar as boas-práticas e os aspetos-chave que podem ser replicados em outros sistemas educativos. Poderá no futuro servir como exemplo para a adoção de novos standards em países que sentem carência em bases educacionais fortes bem como para uma análise de potenciais fragilidades no que toca a educação na Estónia e pontos que poderiam vir a ser melhorados.

Com a incidência em áreas da Ciência, Matemática e Leitura, o objetivo é intensificar e aprimorar políticas educacionais, com o intuito de melhorar ainda mais o sistema educacional estoniano.

4. Tratamento dos dados

Para cumprir o objetivo desta análise, começamos o estudo pelo tratamento do *dataset* que nos foi fornecido. Este contém os resultados do PISA 2018 com 612.004 linhas e 104 colunas inicialmente.

Após filtrarmos o *dataset* inicial para a Estónia e eliminarmos as colunas completamente vazias, ficamos com 5316 linhas e 72 colunas, sendo que destas 67 são numéricas e 5 são categóricas. Relativamente às variáveis de INPUT e PROFILE, estas serão apresentadas de seguida, explicando-se o seu significado e algumas métricas interessantes, sendo que todas as que foram usadas são numéricas, considerando a tipologia de variáveis necessárias para aplicar PCA - *Principal Component Analysis* - e *clustering*.

Variáveis de INPUT

As variáveis de INPUT são variáveis numéricas utilizadas na técnica de análise de componentes principais (PCA). Das 67 variáveis numéricas, presentes na base de dados filtrada para a Estónia, selecionamos as 38 variáveis que serão descritas no [anexo A](#).

Antes de realizarmos os PCA's selecionamos um conjunto de variáveis que continham informação sobre as áreas de formação avaliadas/disciplinas, questões do ponto de vista de saúde mental dos alunos e até informações relativas aos pais e professores. A decisão de incluir estas variáveis passou pela visualização das correlações entre elas e pela incorporação de espírito crítico na seleção das mesmas. Por exemplo, achámos interessante incluir informações relativas aos pais do aluno porque achamos que tem influência no carácter cognitivo. Ainda, incluímos se a criança já tinha sido vítima de *bullying* sendo que este fator poderá refletir-se nos resultados obtidos na escola, mesmo não tendo uma correlação significativa com outras variáveis. (percentagem de estudantes que afirmaram ter sofrido *bullying* algumas vezes no mês foi de 25.4%, 2.7 pontos percentuais acima da média da OCDE).

Para além do que já foi referido, ao ver algumas medidas descritivas achou-se interessante ver a assimetria e a curtose das colunas que continham os resultados das crianças nas três grandes áreas (Leitura, Matemática e Ciências). Estas continham valores negativos, próximas de 0, quando comparadas com as outras variáveis, o que informa que os estudantes estão mais desviados para resultados melhores (assimetria negativa), mas também em pequena escala de frequência (grau de achatamento negativo/curtose). Já nas variáveis referentes ao tempo de aprendizagem dos alunos, em cada uma destas áreas, vê-se que a medida de assimetria tem um comportamento completamente oposto, estando os alunos mais deslocados à esquerda - menos tempo de aprendizagem por semana (assimetria positiva alta).

Variáveis de PROFILE

As variáveis de PROFILE são as que caracterizam os alvos deste estudo, ou seja, os estudantes estonianos. Neste caso, as variáveis de PROFILE que escolhemos, e que achamos que se encaixariam nesta tipologia, seriam as variáveis: género (**ST004D01T**), **GRADE** (nota obtida comparada à nota modal no país) e **HISCED** (educação mais

elevada, considerando os 2 pais). Assim, é nos possível diferenciar os alunos que fizeram parte deste estudo.⁶

5. Identificação das dimensões da análise

Para realizarmos a PCA, começamos por eliminar as linhas que continham valores omissos, ficando somente com as observações completas da base de dados, mais precisamente com 4035 linhas e 41 colunas, onde 3 eram de PROFILE, portanto, 38 para o PCA. Em contrariedade à utilização dos “complete cases” na análise, iremos realizar uma outra abordagem [explicada mais à frente no relatório](#), onde são imputados os valores omissos, de forma a não perder tantas linhas da base de dados e, ao mesmo tempo, verificar se tiramos conclusões diferentes.

Referente às correlações entre variáveis numéricas presentes, pudémos concluir que as mesmas eram divididas de forma a gerar pequenos grupos de variáveis correlacionadas entre si, no entanto, de forma geral, as correlações eram baixas. Este fenómeno pode ser observado no [anexo E](#).

Mais adiante, fizemos o Teste de Bartlett, sendo que H_0 tem de ser rejeitada para que se possa realizar a PCA. Como $X^2_{(741)} = 100620.7$; $p < 0.001$. Assim sendo, o teste sugere que se pode proceder com a análise em componentes principais (i.e., a amostra é adequada). Obtivemos um valor de 0.68 para KMO, o que é um valor aceitável. De seguida, estandardizamos as variáveis, para garantir que estão todas na mesma escala, logo têm todas o mesmo peso e relevância na análise.

6. PCA's caracterização

Primeiramente, começou-se por fazer PCAs para todas as variáveis escolhidas, que eram 38. Analisaram-se os valores próprios, o KMO, o teste de Bartlett (onde se rejeita a hipótese nula), verifica-se se são todas quantitativas e se fazem sentido para o objetivo em estudo. Para este caso, testou-se para 6, 7 e 11 componentes, não decidindo fazer com outros valores superiores, uma vez que se verificou a existência de valores próprios superiores a 1, até 11 componentes.

Em seguida, apresenta-se o scree plot para os componentes principais do PISA, onde verifica-se a existência de um “cotovelo” para 7 ou 11 componentes.

⁶ Para uma maior compreensão das variáveis, verificar o [anexo A](#).

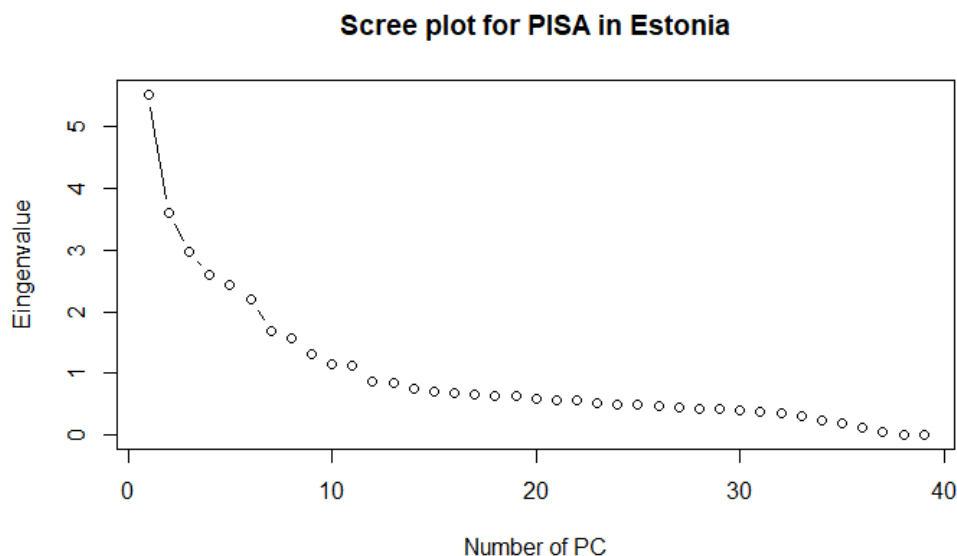


Gráfico 1- Scree Plot para descobrir o valor "cotovelo"

Quanto à variância acumulada, apesar de se possuir uma maior variância acumulada para 7 componentes, decidiu-se escolher menos componentes, neste caso, 6, de forma a conseguir-se uma melhor interpretação das mesmas (seguindo a navalha de Ockham), sendo que quanto mais componentes se tinha, mais difícil se tornava a explicação e resultante caracterização das mesmas. O valor da variância acumulada das 38 variáveis originais de INPUT foram de cerca de 0.5 (metade da explicação das variáveis, um pouco inferior ao aceitável (0.6), pelas razões descritas acima).

Para suportar esta escolha, foram realizados testes com os 6 e 7 componentes, onde vimos dificuldade para nomear um sétimo componente que remetia para informações bastante idênticas a outros componentes que já tínhamos (variáveis relativas ao tempo).

Quanto às comunicações, há uma elevada explicabilidade com a solução obtida, apesar de haver algumas exceções. Os valores mais frequentes foram acima de 0.4.

A variável “BeingBullied” foi incluída na amostra devido à sua significância em termos de interesse no estudo, mesmo possuindo uma comunalidade de 0.15. Esta variável é de interesse sendo que ao entrevistarmos a responsável pelo estudo do PISA na Estónia desde 2007, a mesma sublinhou a presença deste fator e a sua influência na Estónia. De forma a corroborar este fenómeno, pode-se também consultar o [anexo E](#) e [F](#), bem como [visitar este website](#).

7. Conclusões retiradas das componentes

Após a obtenção das componentes principais, observou-se os *loadings* resultantes e aplicou-se vários métodos de rotação: VariMax, QuartiMax e Oblimin de forma a analisar quais variáveis eram mais retidas em cada componente gerada. Desta maneira, conseguiu-se concluir o seguinte sobre cada componente, considerando as variáveis⁷ mais predominantes em cada e uma possível denominação para cada uma:

- **RC1 (ambition)** - principais variáveis retidas: variáveis referentes às disciplinas de matemática, ciências e leitura, sensação de pertença, **RESILIENCE** (resiliência), **SCREADCOMP** (perceção do aluno sobre a importância da leitura), **MASTGOAL** (ter o objetivo de aprender e dominar uma tarefa de acordo com os standards pessoais), **PERFEED** (capacidade de compreender feedback e como isso poderá afetar o aluno), **ADAPTATIVITY** (adaptabilidade), **TEACHSUP** (auxílio do professor nas aulas de teste de línguas) -> equivale ao **PC1**.
- **RC2 (achievement)** - variáveis referentes às disciplinas de matemática, ciências e leitura, **DIRINS** (onde os professores utilizam técnicas específicas para o ensino de uma competência específica), **SMINS** (tempo de aprendizagem por semana para ciências), **ST059Q03TA** (total de horas por semana de aulas de matemática) -> equivale ao **PC2**.
- **RC3 (ICT)** - variáveis relacionadas com **ICT** (ou TIC), **ADAPTIVITY** (adaptabilidade), **ENTUSE** (uso de aparelhos tecnológicos para atividades lúdicas), **USESCH** (uso de aparelhos tecnológicos para fins escolares), **HOMESCH** (uso de aparelhos tecnológicos para trabalhos de casa) -> equivale ao **PC3**.
- **RC4 (learning-time)** - relacionadas com **ICT** (ou TIC), **SMINS**, **MMINS**, **LMINS**, **ST059Q03TA**, **ST059Q01TA** (total de horas por semana de leitura, é uma característica que vai afetar o desempenho nas outras disciplinas) -> equivale ao **PC4**.
- **RC5 (mental-health)** - **EUDMO** (significado da vida), **SWBP** (bem-estar subjetivo), **BELONG** (bem-estar subjetivo), **RESILIENCE**, **EMOSUPS**, **ST016Q01NA** (satisfação geral com a vida), **BEINGBULLIED** negativo-> equivale ao **PC5**.

⁷ Verificar o [anexo A](#) de para uma maior clarificação sobre a descrição de cada variável.

- **RC6 (support)** - **ADAPTATIVITY**, **WEALTH** (riqueza familiar), **HOMEPPOS** (bens em casa), **ESCS** (índice sócio económico), **EMOSUPS** (perceção de suporte emocional dos pais), **CULTPOSS** (bens culturais em casa) -> equivale ao **PC6**.

8. Análise dos PCA's

Após a realização da análise de componentes principais foram extraídas as correlações com as novas componentes criadas. Estes adquiriram valores muito significativos, considerando o método utilizado, podendo dizer-se que assumem todos valores de 0.

Para a fase de *clustering*, que se segue, estes valores são bastante positivos, pois garantem que as componentes criadas não possuem multicolinearidade, um dos objetivos a colmatar quando criamos PCA's. Somado a isto, ao realizar uma análise bivariada no novo conjunto de variáveis, vemos que estas não aparentam ter associações fortes, conduzindo a um bom indicativo para a fase de *clustering*.

Apesar da fraca associação entre os componentes, explorou-se algumas relações em separado para melhor comparar. Foi o exemplo das seguintes⁸:

- Ambition com Achievement.
- Learning Time com Achievement.
- Ambition com Support.
- Mental health com Achievement

Olhando para estas associações, pode-se validar o que foi mencionado acima, e assim, proceder para o *Clustering*.

9. Identificação da heterogeneidade na base de dados

A fim de se criar perfis para os diferentes alunos, mediante as variáveis mais presentes em cada um, e consequentemente, a redução de linhas presentes na base de dados, procedeu-se à técnica de *clustering* (descoberta de padrões nos dados). Foram testadas diferentes metodologias ao longo desta fase, de forma a poder encontrar a que melhor se encaixa e capta a variação entre os dados.

A partir do **Clustering Hierárquico**, utilizou-se primeiramente o “Ward-D2”, onde se experimentou várias partições do dendrograma gerado. A partir disto, foi possível a

⁸ Verificar o [anexo I](#) para a visualização dos mesmos.

criação de uma silhueta, de forma a conseguir perceber-se qual a robustez dos *clusters*, ou conglomerados, criados. As experiências feitas foram com 2, 3, 4 e 5 *clusters*, de forma que a média obtida para a média da largura da silhueta foi 0.09, 0.07, 0.05, 0.06, respetivamente.

Para além do método “Ward-D2”, também foram estudados outros métodos como o “complete”, onde se testou com 3 *clusters*. Não foram testadas mais divisões, sendo que se percebeu que com este método não havia uma visualização tão propícia para a determinação de cortes a fazer, de forma que, apenas se considerou o método “Ward-D2” como mais apropriado para tal.

Ademais, procedemos à utilização do [K-Means](#), para tentarmos encontrar outro ponto de corte que permita a criação de conglomerados de forma a ser muito semelhantes entre si, mas distantes uns dos outros, sendo que este é um procedimento cíclico, (ou seja, requer várias tentativas).

Para isso, foi utilizado o *scree-plot*, de forma a poder encontrar este valor, conhecido como o “cotovelo”. Este encontra-se [anexado](#).

Ao observar este mesmo gráfico, achámos que o valor que se encontrava no ponto ideal de corte seria o 4, portanto utilizou-se o $k = 4$ para o K-Means. Os 4 conglomerados, *clusters*, criados foram guardados num *dataframe*, de forma a poderem ser utilizados em futuras análises.

10. Conclusões obtidas com o clustering

Através da técnica aplicada com o *clustering*, mais especificamente, com o K-Means, com $k = 4$, foi possível a criação de 4 *clusters* que compara cada um com as componentes criadas anteriormente.

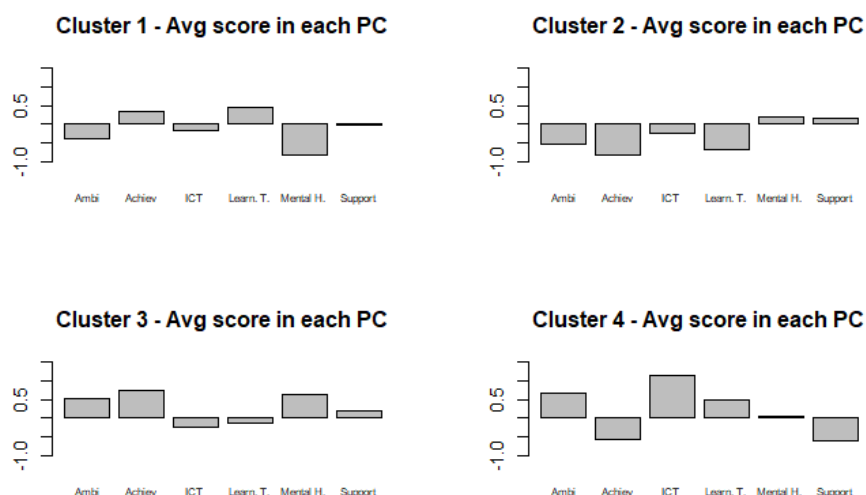


Figura 1- Clusters criados e a sua associação com cada componente

As conclusões que pudemos retirar da criação dos *clusters* e a sua respetiva alocação de observações mediante as componentes existentes, foram:

- **Cluster 1** (pessimistas/pessimistic) – Observa-se pouca ambição e uma baixa saúde mental, sendo que este podem indicar uma menor satisfação na vida.
- **Cluster 2** (os desinteressados/ the uninterested) – Observa-se um baixo desempenho e ambição escolar e pouco tempo de aprendizagem - o que poderá ser indicativo de uma baixa assiduidade, ou de pontualidade.
- **Cluster 3** (the achievers/os bem-sucedidos) – Observa-se ambição e um bom alcance em termos de desempenho escolar. Também se observa uma saúde mental alta, o que poderá ser indicativo de satisfação na vida e sentido de pertença (bem-estar).
- **Cluster 4** (tecnófilos/ technophile) – Há uma maior incidência dentro das TIC, ou ICT, em comparação com os restantes *clusters* sendo que poderá ser indicativo de um interesse e gosto pelo uso de tecnologias tanto a nível de uso em sala de aula como em trabalhos de casa.

Observando para a associação dos *clusters* com as variáveis de PROFILE escolhidas, presentes em anexo, conseguimos notar que os estudantes do sexo masculino tendem a ser mais esforçados (*cluster 2*), ao contrário do sexo feminino que foram classificados como pessoas mais bem-sucedidas (*cluster 3*). Ainda, olhando para o maior nível de escolaridade dos pais- o mais alto do pai e mãe juntos- vemos que a maior parte dos pais que têm o maior grau presente na base de dados (Bacharelato), tendem a estar presentes no cluster de estudantes mais bem-sucedidos. Isto pode ser explicado da maneira em que

os pais com maior grau obtido consigam incentivar e apoiar os filhos melhor com os estudos. Por último, na variável **GRADE**, é bastante complicado a tiragem de conclusões, uma vez que a maior parte dos estudantes possuem valor 0 ou -1.

Para o PAM foram experimentadas várias soluções para os *clusters*, 2, 3 e 4, sendo este último o escolhido em seguimento do K-Means. As conclusões possíveis do PAM é que se pode ver que há alguma sobreposição entre *clusters*, o que poderá indicar que os alunos não têm muitos fatores distintivos. Para além de termos em conta a visualização dos *clusters*, também foi levada em conta a explicabilidade dos mesmos, e sendo assim, independentemente de ter sido feita a divisão em 2, 3 ou 4 *clusters*, a explicabilidade foi de 33.33%.

Também foram efetuados o BIC e o GMM de forma a complementar a nossa análise, e podemos concluir o mesmo que foi indicado acima, isto porque os gráficos resultantes da classificação e da incerteza, acabaram por ser indicativos de uma sobreposição bastante elevada e de uma incerteza considerável, sendo que havia uma dificuldade da atribuição das observações a cada *cluster*. Os resultados e os gráficos análogos a esta parte encontram-se em [anexo](#).

11. Análise com imputação de dados omissos

Antes de procedemos a esta análise, e logo depois de seleccionamos apenas as linhas da base de dados referentes à Estónia, foi realizado o teste de MCAR (*missing completely at random*), da biblioteca “*nanian*”. O valor resultado na aplicação deste teste foi de 14693.35, com 12297 graus de liberdade. O *p-value* foi < 0.001 , que indica que rejeitamos a hipótese nula, afirmando que os dados omissos não são completamente aleatórios. Foram identificados 392 padrões de omissos nos dados. De salientar que esta análise foi efetuada, sobretudo, porque existe um valor de *threshold* associado a dados omissos no PISA. O limiar desse valor corresponde a 20% de omissos, o que é aceitável, levando em conta que os dados da Estónia não têm uma taxa de dados omissos tão significativa.

Desta maneira, será relevante imputar os valores na base de dados relativa à Estónia, numa análise complementar e, para isso, irá recorrer-se à biblioteca “*mice*”. Após os dados serem imputados por regressão linear, para 5 iterações, ficámos com 5316 linhas e 41 colunas, sendo 3 destas de PROFILE. De lembrar que são dados sintéticos/criados a partir de outros, mas que mesmo assim iremos observar como estes se comportam à mesma análise efetuada.

Quanto ao número de PCA's, à semelhança da análise dos casos completos, concluímos que seriam 6 componentes, onde cada um destes teria o mesmo significado que os outros já criados. Quanto ao número de *clusters*, de um modo severo, concluímos que no método hierárquico, com o *Ward.D2*, temos 3 *clusters* e no *K-Means* e *PAM* 4 *clusters*, levando em consideração o *screeplot* efetuado. Os *clusters* resultantes diferiam um pouco dos efetuados acima, estando os alunos menos sobrepostos em relação uns aos outros, mas, mesmo assim, ainda existiam alguns pontos difíceis de classificar entre *clusters* (que traziam alguma incerteza). De um modo geral, no que toca a PCA's e *clusters*, tendo a imputação dos omissos ou apenas as observações em omissos, não afeta significativamente as nossas conclusões.

12. Conclusão

Em primeiro lugar, foi necessária uma análise sobre a educação na Estónia e o seu posicionamento em relação aos outros países participantes do estudo do PISA 2018. Com base nesta, percebemos que os resultados obtidos foram fruto de um trabalho contínuo, tanto pelo governo estoniano, como pelos responsáveis pela educação na Estónia e os professores. Isto levou-nos a perceber mais sobre a evolução da educação e de como esta foi alavancada. Conseguiu-se perceber que esta evolução foi constante, sendo que não houve grandes aumentos ou diminuições em termos do seu posicionamento face a média na OCDE nos últimos anos.

Pode-se dizer que, face os resultados apresentados pelos estudos anteriores e os estudos feitos sobre o PISA 2018, a Estónia tem mantido uma boa performance.

Em termos de políticas adotadas pelo governo, é nos possível dizer que o governo estoniano tem-se orientado pelos resultados do PISA de forma a poder desenvolver e melhorar em aspetos que possam ser mais fracos a nível nacional. Pode-se destacar uma estratégia aplicada, a de Aprendizagem ao Longo da Vida de 2020, que foi publicada em 2014, sendo esta muito direcionada para a integração de diferentes faixas etárias no domínio da aprendizagem bem como a promoção da literacia digital.

Em relação aos resultados obtidos pela nossa análise, pode-se salientar que se criaram 6 compósitas, sendo estas representativas de cerca de 50% da explicabilidade dos dados, de forma que se consegue obter algumas conclusões sobre que tipo de variáveis estariam mais representadas em cada uma e consequentemente, nomear as componentes.

Já sobre os conglomerados, ou *clusters*, foram criados 4 diferentes, e pode-se aferir que os mesmos tinham uma explicabilidade de 33.33% e que, ao visualizar os mesmos, é possível se perceber uma grande sobreposição - indicativo de haver pouca diferença entre os alunos e validando a convicção da transversalidade de oportunidades da obtenção de educação na Estónia.

Para além destas conclusões, podemos também nomear os *clusters* em relação à incidência das observações e o valor das componentes em cada um, visto que, cada um, tinha as suas particularidades.

Ao somar todas estas conclusões pode-se observar um *cluster* que está mais virado para a tecnologia, o *cluster* dos tecnófilos, que poderá vir a ser um indicativo da eficácia das medidas adotadas, mais especificamente, a “Education Lifelong Learning strategy 2020”, sendo que, um dos objetivos da mesma era o reforço na importância da literacia digital.

Também é possível alocar alguma causalidade quando se observa o *cluster* 3, referente aos alunos bem-sucedidos, e esta estratégia, sendo que um dos objetivos da mesma era a qualificação reforçada dos professores. Isto poderá ser observado, na componente “ambition” que era bastante elevada, algo que poderá ser indicativo de valores altos na variável que menciona a capacidade de integração de feedback dado pelos professores, bem como na variável que mede a resiliência do aluno.

Ainda sobre a caracterização dos conglomerados, pode-se referir o *cluster* dos desinteressados, o *cluster* 2, sendo este com a performance mais baixa em termos da literacia, e uma resiliência inferior. Isto poderá vir a indicar que há margem para melhorias no que toca a motivação dos estudantes e a inspiração para adquirir resultados melhores, não só pelos professores, como pelos pais, bem como, uma necessidade de maior comunicação entre as três partes, alunos, pais e professores, de forma a perceber melhor que fatores levam a estes resultados.

Para além destas medidas, pode-se também acrescentar uma maior necessidade de atenção prestada ao bem-estar dos alunos, sendo que a componente da saúde mental tinha valores bastante inferiores ao desejado, em especial, na primeira conglomeração. Isto poderá vir a ser indicativo da presença de *bullying* na Estónia, algo que teria sido já estudado a partir de outras fontes, bem como mencionado pela responsável pelo PISA desde 2007, Gunda Tire. Pode também servir como um sinal para o governo estoniano para dar mais importância e investir na saúde mental e o acompanhamento das crianças,

principalmente, em termos de questões existenciais, necessárias para uma boa saúde mental, e consequentemente, bom desempenho escolar.

Outro apontamento em relação à alocação de observações em cada cluster, é a existência de alguma disparidade em termos do tempo de aulas para as diferentes disciplinas, isto poderá sugerir alguma diferença em termos de não só o currículo e as metodologias individuais de cada escola, sendo que há alguma liberdade nas decisões das mesmas, como poderá vir a ser um ponteiro da assiduidade dos estudantes participantes no estudo. Algo que, se o governo estoniano reconhecer como relevante estudar, será benéfico no sentido de uniformizar a educação em todas as regiões da Estónia (caso este fator esteja a ser influenciado pela fraca assiduidade dos alunos).

Em último, é de se notar que todas estas conclusões referidas, são apenas sugestões de medidas e ponteiros que poderão vir a ser seguidas como linhas de orientação para o continuo melhoramento da performance no âmbito da educação na Estónia.

13. Agradecimentos

Os autores gostariam de expressar os seus agradecimentos a Gunda Tire, responsável pelos estudos do PISA 2018 da Estónia, por todo o apoio prestado ao longo da realização do trabalho, nomeadamente, pela sua disponibilidade para a realização de uma entrevista e de esclarecimentos dados em relação à educação na Estónia. Ademais, gostariam de agradecer ao professor Jorge Sinval pelo tempo disponibilizado para sessões de esclarecimento prestadas e pela disponibilização do contacto da Gunda Tire.

14. Bibliografia

- Caucasian Journal. (2020, 02). *Gunda Tire: "Estonians believe in education, and this belief has been essential for centuries"*. Retrieved from <https://english.caucasianjournal.org/2020/02/gunda-tire-estonians-believe-in.html>
- Crato, N. (2021). *Improving a Country's Education: PISA 2018*. Lisbon: Springer.
doi:<https://doi.org/10.1007/978-3-030-59031-4>
- Education GPS. (14/06/2023, 18:44:59). *OECD*. Retrieved from <https://gpseducation.oecd.org/>
- IAVE. (2019). *RELATORIO_NACIONAL_PISA2018_IAVE*. Lisboa: IAVE.
[://www.cnedu.pt/content/noticias/internacional/RELATORIO_NACIONAL_PISA2018_IAVE.pdf](http://www.cnedu.pt/content/noticias/internacional/RELATORIO_NACIONAL_PISA2018_IAVE.pdf)
- IAVE. (n.d.). *Estudo Internacional Pisa*. Retrieved 06 11, 2023, from IAVE:
<https://iave.pt/estudo-internacional/pisa/>
- Jürimäe, M. (2022, 06 14). *Insights from the Education Nation: The Case of Estonia*. Retrieved from Green European Journal - The European Venue for Green Ideas:
<https://www.greeneuropeanjournal.eu/insights-from-the-education-nation-the-case-of-estonia/>
- OCDE. (2016). *PISA 2015 Results in Focus*. IAVE. Retrieved from https://iave.pt/wp-content/uploads/2020/02/Relatorio_PISA2015.pdf
- OCDE. (2019). *Assessment PISA Results 2018*. OCDE. Retrieved from https://www.oecd.org/pisa/publications/PISA2018_CN_EST.pdf
- OECD. (2019). *PISA 2018 Assessment and Analytical Framework*. PISA. Paris: OECD Publishing. doi:<https://doi.org/10.1787/b25efab8-en>.
- Organization for Economic Cooperation and Development (OECD). (2018-2019). *PISA 2018 Worldwide Ranking – average score of mathematics, science and reading*. Retrieved from FactsMaps: <https://factsmaps.com/pisa-2018-worldwide-ranking-average-score-of-mathematics-science-reading/>
- UNESCO. (n.d.). *Estonia: Lifelong Learning Strategy 2020, issued in 2014*. Retrieved from <https://uil.unesco.org/document/estonia-lifelong-learning-strategy-2020-issued-2014>

15. Anexos

ANEXO A - Descrição das variáveis mais relevantes:

- **PV1SCIE**: valor plausível a Ciência, varia de 197.0 a 830.1, como a mediana é 530.5, cerca de 50% dos alunos estão ligeiramente acima da média, sendo que a escala de pontuação vai de 0 a 1000;
- **PV1READ**: valor plausível em Leitura, varia entre 224.8 e 870.9, sendo a que tem o valor plausível mais elevado das três disciplinas/ componentes avaliadas;
- **PV1MATH**: valor plausível em Matemática, varia de 210.1 a 816.1 e a média é de 523.3;
- **SOIAICT**: TIC como tópico nas interações sociais (WLE – Warm Likelihood Estimation) admite valores de -2.1763 a 2.3635 e tem 517 valores omissos, apenas 25% dos alunos admite valores superiores a 0.5149;
- **AUTICT**: autonomia percebida no uso das TIC (WLE) varia entre -2.5144 e 2.0258 e apresenta 501 valores omissos, a média é de -0.0341;
- **COMPICT**: competência percebida no uso das TIC (WLE) varia de -2.6033 a 1.9885, tem 510 valores omissos e a mediana é de 0.0537;
- **INTICT**: interesse nas TIC (WLE) varia entre -2.9505 e 2.6672, com 486 valores omissos e 75% dos alunos têm mais de -0.6385 de interesse nas TIC;
- **ENTUSE**: uso das TIC fora da escola por lazer (WLE) varia de -3.5940 a 4.2444, apresenta 417 valores omissos e a média é 0.0734;
- **BEINGBULLIED**: experiência do aluno de sofrer bullying (WLE) varia entre -0.7823 e 3.8591, tem 536 valores omissos e mediana de 0.1462;
- **BELONG**: bem-estar subjetivo: sentimento de pertença à escola (WLE) varia de -3.2367 a 2.7249, com 227 valores omissos, 25% dos alunos têm a sensação de pertença superior a 0.3018;
- **MASTGOAL**: mestria na orientação dos objetivos (WLE) varia entre -2.5252 e 1.8524, apresenta 239 valores omissos, 75% dos alunos admitem um valor superior a -0.7346;
- **RESILIENCE**: resiliência (WLE) varia de -3.16750 a 2.36930, tem 219 valores omissos e média de -0.01867;
- **SWBP**: bem-estar subjetivo: afeto positivo (WLE) varia entre -3.0666 e 1.2386, com 242 valores omissos e mediana de -0.5928;

- **EUDMO**: eudemonia: significado na vida (WLE) varia de -2.14640 a 1.74110, apresenta 206 valores omissos, 25% dos alunos tem um valor acima de 0.26210;
- **GFOFAIL**: medo geral de fracasso (WLE) varia entre -1.8939 e 1.8905, tem 208 valores omissos, 75% dos alunos admitem um valor acima de -0.6870;
- **WORKMAST**: mestria do trabalho (WLE) varia de -2.7365 a 1.8164, com 246 valores omissos e média de -0.3154;
- **SCREADCOMP**: autoconceito de leitura: percepção de competência (WLE) varia de -2.4403 a 1.8839, tem 215 valores omissos, 25% dos alunos admitem um valor superior a 0.1222;
- **JOYREAD**: gosto em ler (WLE) varia entre -2.73160 e 2.65740, com 135 valores omissos, 75% dos alunos assumem valor superior a -0.60860;
- **TEACHINT**: percepção do interesse do professor (WLE) varia entre -2.21770 e 1.82450, apresenta 131 valores omissos e média igual a -0.09531;
- **ADAPTIVITY**: adaptação da instrução (WLE) varia de -2.2652 a 2.0073, tem 156 valores omissos e mediana igual a -0.1716;
- **STIMREAD**: estimulação por parte do professor da leitura percebida pelos alunos varia entre -2.3003 e 2.0871, com 145 valores omissos e 25% dos alunos admitem valor superior a 0.5102;
- **EMOSUPS**: percepção por parte do aluno do suporte emocional dos pais (WLE) varia de -2.4468 a 1.0346, apresenta 437 valores omissos, 75% dos alunos assume um valor superior a -0.6576;
- **PERFEED**: feedback percebido (WLE) varia entre -1.6391 e 2.0165, tem 157 valores omissos e média igual a -0.1487;
- **DIRINS**: instrução dirigida pelo professor (WLE) varia de -2.94250 a 1.82020, com 121 valores omissos e mediana de -0.04240;
- **TEACHSUP**: apoio do professor em aulas de teste de línguas (WLE) varia entre -2.7106 e 1.3411, apresenta 121 valores omissos e 25% dos alunos admitem um valor superior a 0.4810;
- **WEALTH**: riqueza familiar (WLE) varia de -6.90160 a 4.12190, tem 103 valores omissos, 75% dos alunos admitem um valor superior a -0.53950;
- **CULTPOSS**: bens culturais em casa (WLE) varia de -1.79290 a 2.05370, apresenta 122 valores omissos e 25% dos alunos assumem um valor acima de 0.51200;
- **HOMEPOS**: bens em casa (WLE) varia entre -8.58050 e 4.98480, tem 102 valores omissos e 75% dos alunos assumem um valor superior a -0.49320;

- **ESCS:** índice do estatuto económico, social e cultural varia de -5.1849 a 3.4450, com 114 valores omissos e média de 0.1004;
- **SMINS:** quanto tempo de aprendizagem em minutos por semana de ciência varia entre 0.0 e 2380.0, apresenta 576 valores omissos e mediana de 213.1;
- **LMINS:** quanto tempo de aprendizagem em minutos por semana de leitura varia de 0.0 a 2240.0, tem 559 valores omissos e 25% dos alunos assumem valores superiores a 225.0;
- **MMINS:** quanto tempo de aprendizagem em minutos por semana de matemática varia entre 0.0 e 2070.0, com 552 valores omissos e 75% dos alunos assumem valores superiores a 180.0;
- **ST059Q03TA:** quantas horas são tipicamente exigidas que o estudante frequente por semana de matemática, varia de 0.00 a 34.00, apresenta 566 valores omissos e média de 4.67;
- **ST059Q02TA:** quantas horas são tipicamente exigidas que o estudante frequente por semana de ciência; varia entre 0.000 e 34.000, tem 537 valores omissos e mediana de 5.000.
- **ST059Q01TA:** quantas horas são tipicamente exigidas que o estudante frequente por semana de leitura varia de 0.000 a 32.000, com 545 valores omisos e 25% dos alunos assumem valores superiores a 4.000.
- **ST016Q01NA:** de forma geral quão satisfeito o estudante está com a sua vida ultimamente varia entre 0.000 e 10.000, apresenta 187 valores omissos e 75% dos alunos assumem valores superiores a 6.000;
- **HISCED:** educação mais elevada, considerando os 2 pais, segundo a Classificação Internacional Normalizada da Educação varia de 0.00 a 6.00, a mediana apresenta valor 5.00. Esta variável não foi considerada na PCA, sendo que foi determinada como uma variável de PERFIL;
- **ST004D01T:** género do estudante estandardizado varia entre 1.000 e 2.000 e a média é 1.501. Esta variável não foi considerada na PCA, sendo que foi determinada como uma variável de PERFIL;
- **USESCH:** uso das TIC na escola em geral (WLE) varia de -1.72 a 3.30, 75% dos alunos assumem valores superiores a -0.61;
- **HOMESCH:** uso das TIC fora da escola, para uso em atividades escolares (WLE) varia entre -2.30 e 3.31, 25% dos alunos assumem valores superiores a 0.51.

- **GRADE**: nota obtida comparada à nota modal no país, varia entre -2.0000 e 2.0000.
A mediana é de 0.0000, logo 50% dos alunos está dentro da moda do país

ANEXO B - Países e Economias Participantes no PISA 2018

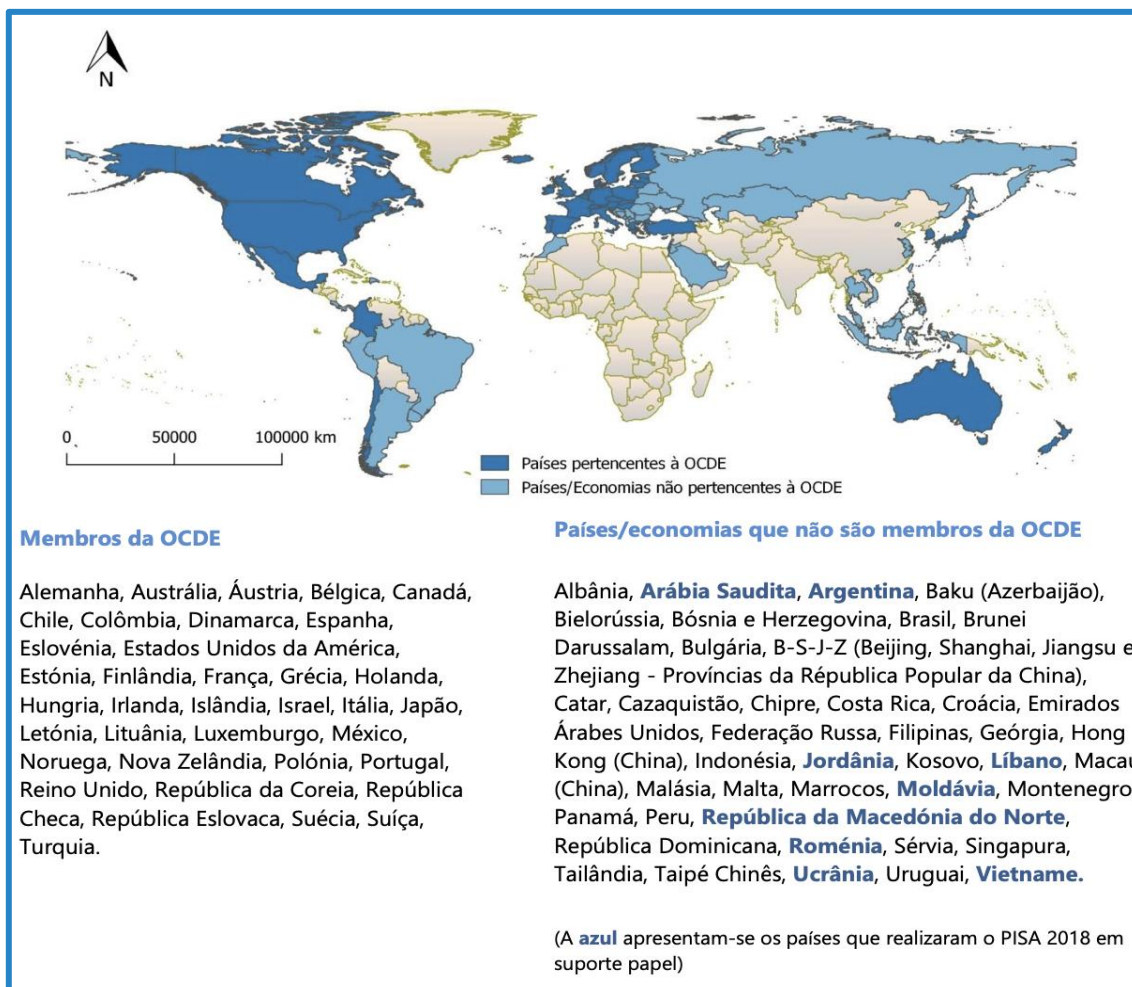


Figura 2- Países e Economias participantes no PISA 2018

Fonte: IAVE, a partir de OCDE (2019) Programme for International Student Assessment – PISA 2018

ANEXO C -Imagem ilustrativa do ranking da Estónia

Face outros países e economias participantes do PISA para as variáveis referentes a matemática, ciências e leitura

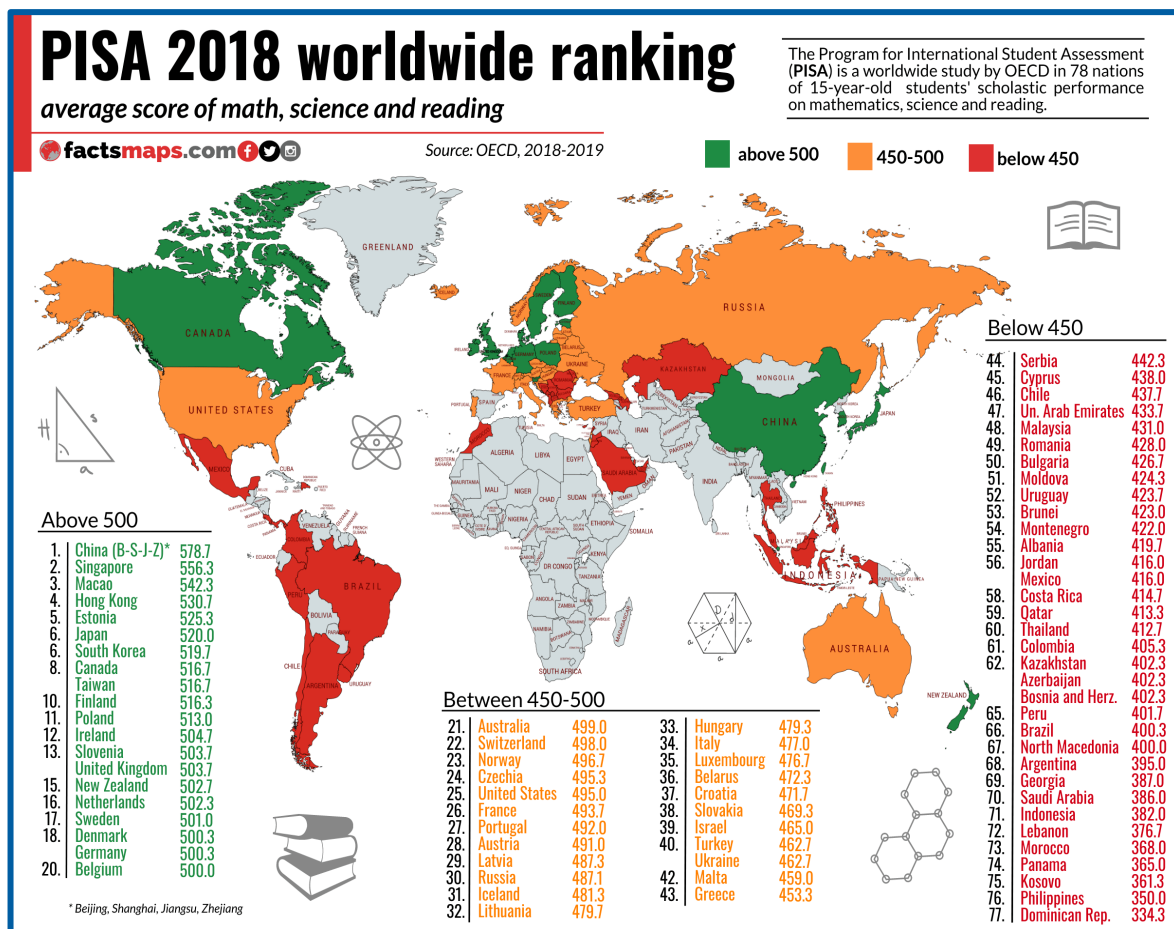


Figura 3- Ranking da Estónia no PISA 2018

Fonte: [PISA 2018 Worldwide Ranking – average score of mathematics, science and reading.](#)

ANEXO D - Education Lifelong Learning strategy 2020

A Estratégia Estoniana de Aprendizagem ao Longo da Vida abrange todos os níveis de educação e estabelece objetivos estratégicos importantes. Esta estratégia tem como principal objetivo: promover o desenvolvimento individual e social, garantir a competência e motivação dos docentes, alinhar necessidades do mercado de trabalho de forma a aumentar o número de qualificados independentemente da idade, promover a participação na educação ao longo da vida na Estónia, destacar a importância do uso digital tanto no ensino como na aprendizagem e por fim aumentar a igualdade de oportunidade.

Fonte: [UNESCO](#)

ANEXO E – Alguns pontos chave da entrevista com Gunda Tire – responsável pelo estudo do PISA na Estónia desde 2007:

1 -What do you think sets Estonia’s education system apart from the rest of the world?

Not exactly but the system is very complex. Give students autonomy to improve to check the system. Spain doesn’t have external information apart from PISA and in the end is not that good. Germany don't get so much information before studying PISA. School is interested in improving. Estonian have divorced situations and the results are not affected.

2- What do you think are the most influential factors in the education of an estonian child, we read that the Estonian education values soft skills a lot, and has for example, sewing, cooking, and coding activities that allow children to develop not only skills that they can use later but also, cooperation, goal orientation and problem solving.

Not only learn outcomes but also learning these skills because the teachers have autonomy. If we can put the old-style learning and the new style learning (we can put these soft skills from the basic skills). Basic roles like passing the street are a thing “inata”.

3- We understand that with this system the “poorest students in Estonia” perform “better than the top quarter with the most affluent background in many countries”, like you mentioned before in the 2018 study, what exactly makes this social disparity not felt in the Estonian children's education?

The richness and poorness are in the country. Finding a problem in the first steps because if you don’t the kid doesn’t learn anything and waits for the next year. Finding the students who need help at the early stages.

- In the same schools you have both poor and rich kids, there is no distinction, there are only 11% private schools;
- Find the problem at an early stage.

4- We had the opportunity to look at the “PISA 2018 Results in 10 Countries” book, more specifically in Estonia, and we got this understanding that Estonian people are

always searching for ways to improve their educational systems and methods of teaching. What do you think was or is the biggest challenge in this regard and how is Estonia tackling it?

We can always be better... Political situation changed our education a lot where more students. Portuguese students in 1st grade are the number of students in Estonia (in proportion)

5- Are there significant differences in the development of students from different regions that students are from?

In the islands the students perform very well. Russian border is not performing very well compared (Narva). Rural and urban differences are not that big. Areas close to Tallinn get better results.

6- What would you say are the most consistent patterns throughout the years in Estonia's education?

Policies in the country that apply to everybody but depends on the school. Estonia is in the same situation of many countries, but they decrease their results.

- Girls better in reading
- Boys better in mathematics
- Similar in science

We haven't improved, we remained consistent, other countries around us may have changed/went down.

7- From what we have been discussing, what other factors would you think that we could also consider for example parents' support, time spent learning each subject or the student's life goals and aspects such as the meaning in life?

- Mathematics
- The kids are not that competitive like they say themselves.
- Growth mindset, bullying tardiness, skipping school, sense of belonging.

ANEXO F – Matriz de correlação entre variáveis numéricas para a Estónia

De forma a facilitar a visualização da geração de grupos de variáveis correlacionadas:

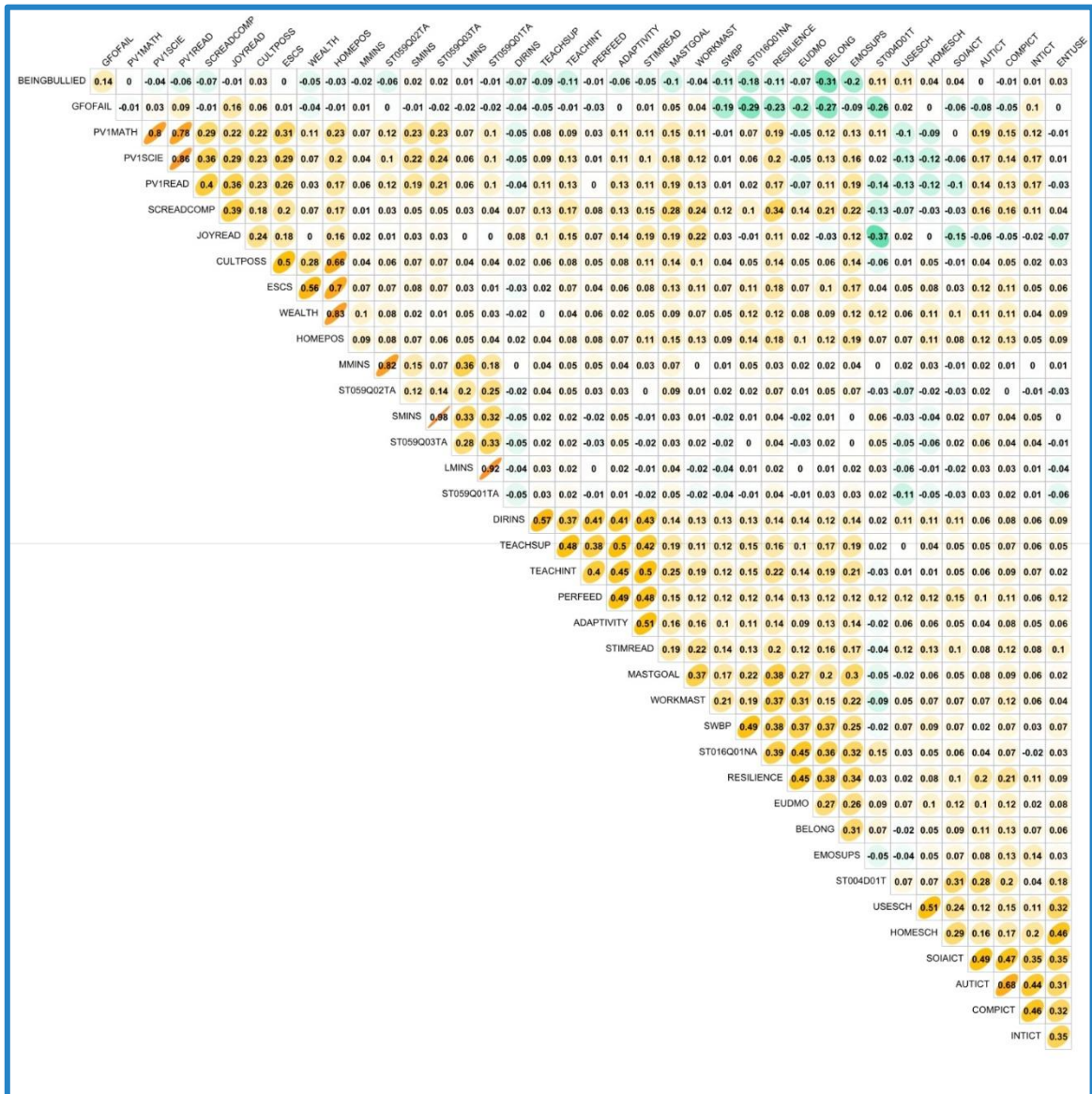


Figura 4 - Correlações entre variáveis numéricas

ANEXO G – Dados sobre *Bullying* na Estónia face a média dos países da OCDE

Indicator	Sort	Estonia EST	OECD average OCD	Portugal PRT
BULLYING				
Students who reported being bullied (includes any type of bullying act) at least a few times a month (%)	(2018) Download Indicator	25.4	22.7	13.7
Students who reported strongly agreed that it is a good thing to help students who can't defend themselves (%)	(2018) Download Indicator	28.7	36.5	42.1

Figura 5 - Informação sobre o *Bullying* na Estónia face a média OCDE

Fonte: [Education GPS: Estonia](#)

Na Estónia, cerca de 25% dos alunos relataram sofrer *bullying* pelo menos algumas vezes por mês, em comparação com 23% em média nos países da OCDE. Ao mesmo tempo, 89% dos alunos na Estónia (e 88% dos alunos em média nos países da OCDE) concordaram ou concordaram fortemente que é bom ajudar os alunos que não podem se defender. Além disso, foi estimado que 17% dos alunos faltaram um dia de aula e 45% dos alunos chegaram atrasados à escola durante esse período. Na maioria dos países e economias, os alunos frequentemente vítimas de *bullying* tinham maior probabilidade de faltar à escola, enquanto os alunos que valorizavam a escola, desfrutavam de um clima disciplinar melhor e recebiam maior apoio emocional dos pais tinham menos probabilidade de faltar à escola. Este foi um ponto também frisado pela Gunda Tire, e por isso, tentamos explorar mais sobre esta variável.

ANEXO H – Gráfico ilustrativo da presença de *Bullying* na Europa em 2018

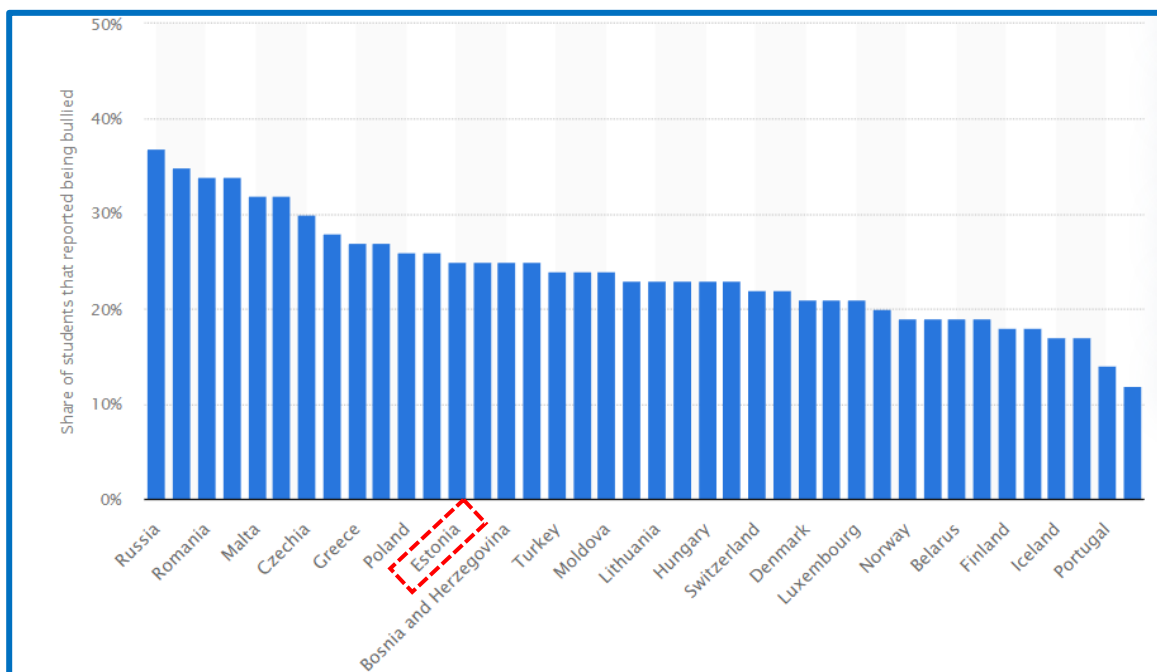


Gráfico 2- *Bullying* na Estónia face a Europa

Fonte: [Bullying in European schools 2018 | Statista](#)

ANEXO I – Rating da Estónia face outros países incluídos no PISA 2018

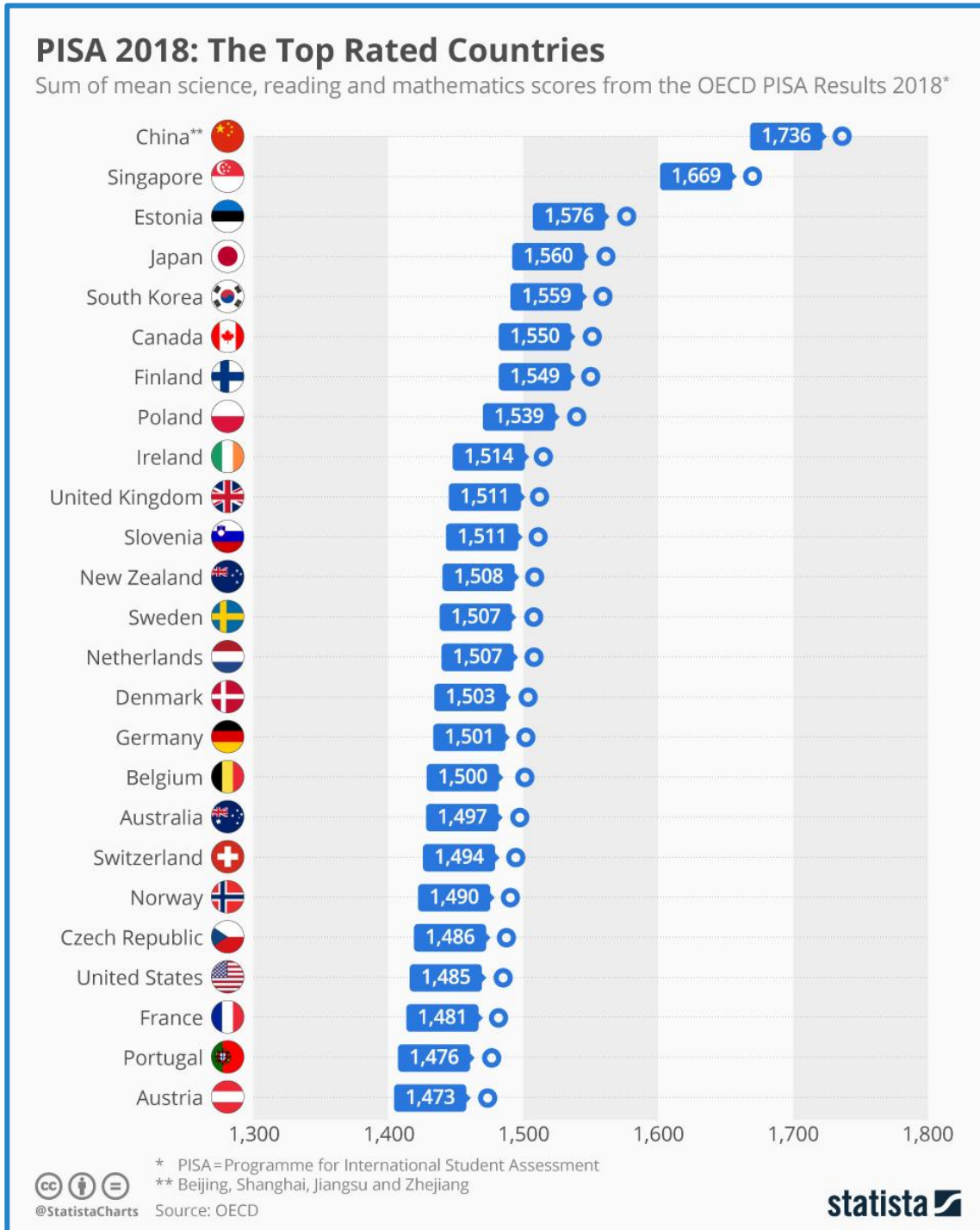
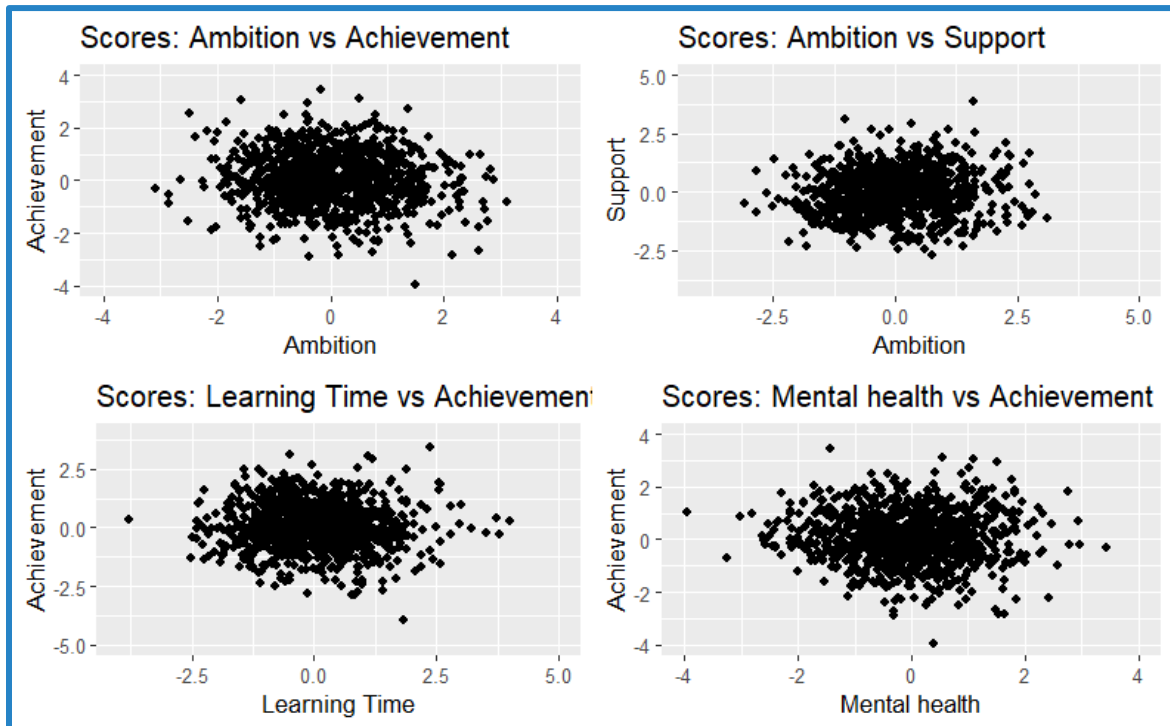


Figura 6- Estónia entre os Top Rated no PISA 2018

ANEXO J – Gráficos diversos

Os seguintes gráficos são alguns dos que teriam sido feitos e apresentados no ficheiro *RMarkdown*.



ANEXO K – Scree plot para o clustering

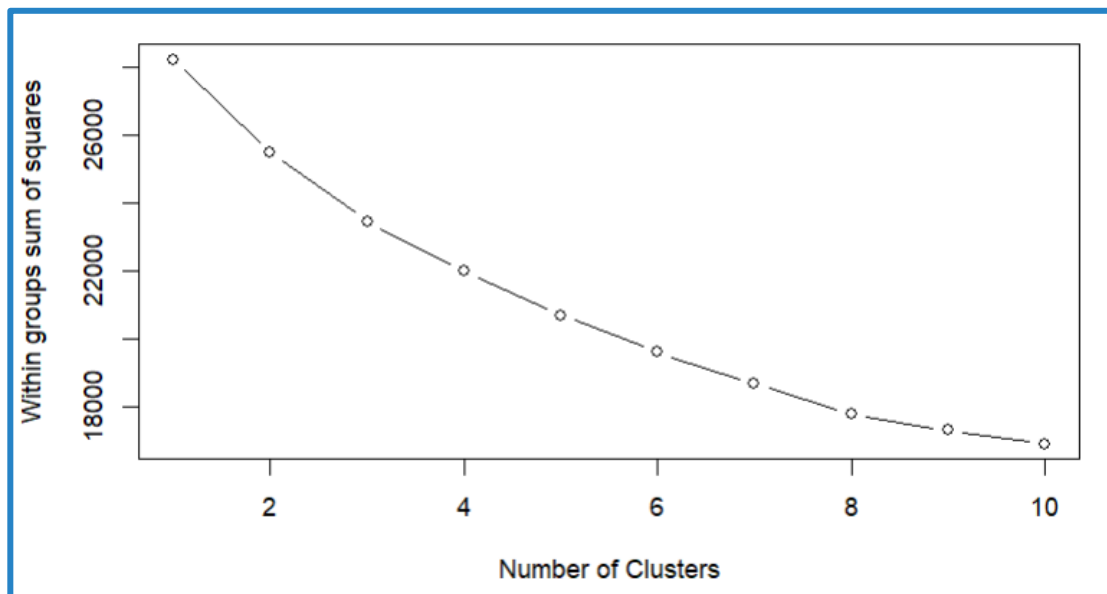


Gráfico 3 - Scree Plot para o clustering

ANEXO L – Tabelas alusivas a cada cluster e algumas variáveis de PROFILE

Tabela 1- Nível ISCED mais alto dos pais (**HISCED**) e a sua alocação pelos 4 clusters

	1	2	3	4
Primary education	0	1	0	1
Lower secondary education	57	32	10	12
Upper secondary education	16	12	10	8
Post-secondary non-tertiary education	366	332	217	171
Short-cycle tertiary education	300	242	178	142
Bachelor's or equivalent level	432	474	718	304

Tabela 3- Divisão entre o género pelos clusters

	1	2	3	4
Female	584	637	737	155
Male	587	456	396	483

Tabela 2- Nota dos alunos em cada cluster (**GRADE**)

	1	2	3	4
-2	6	2	0	1
-1	218	299	168	139
0	940	777	945	491
1	6	15	20	7
2	1	0	0	0

ANEXO M – PAM clustering

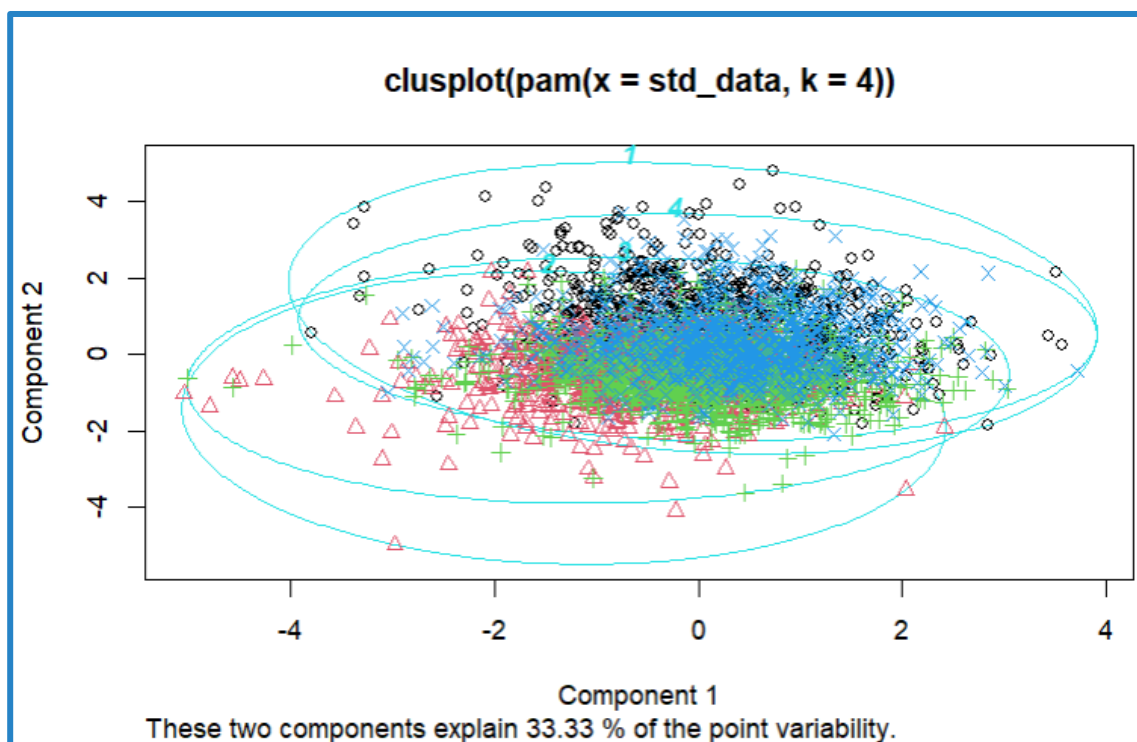


Gráfico 4- PAM com 4 clusters

ANEXO N – Gráficos de Classificação e Incerteza

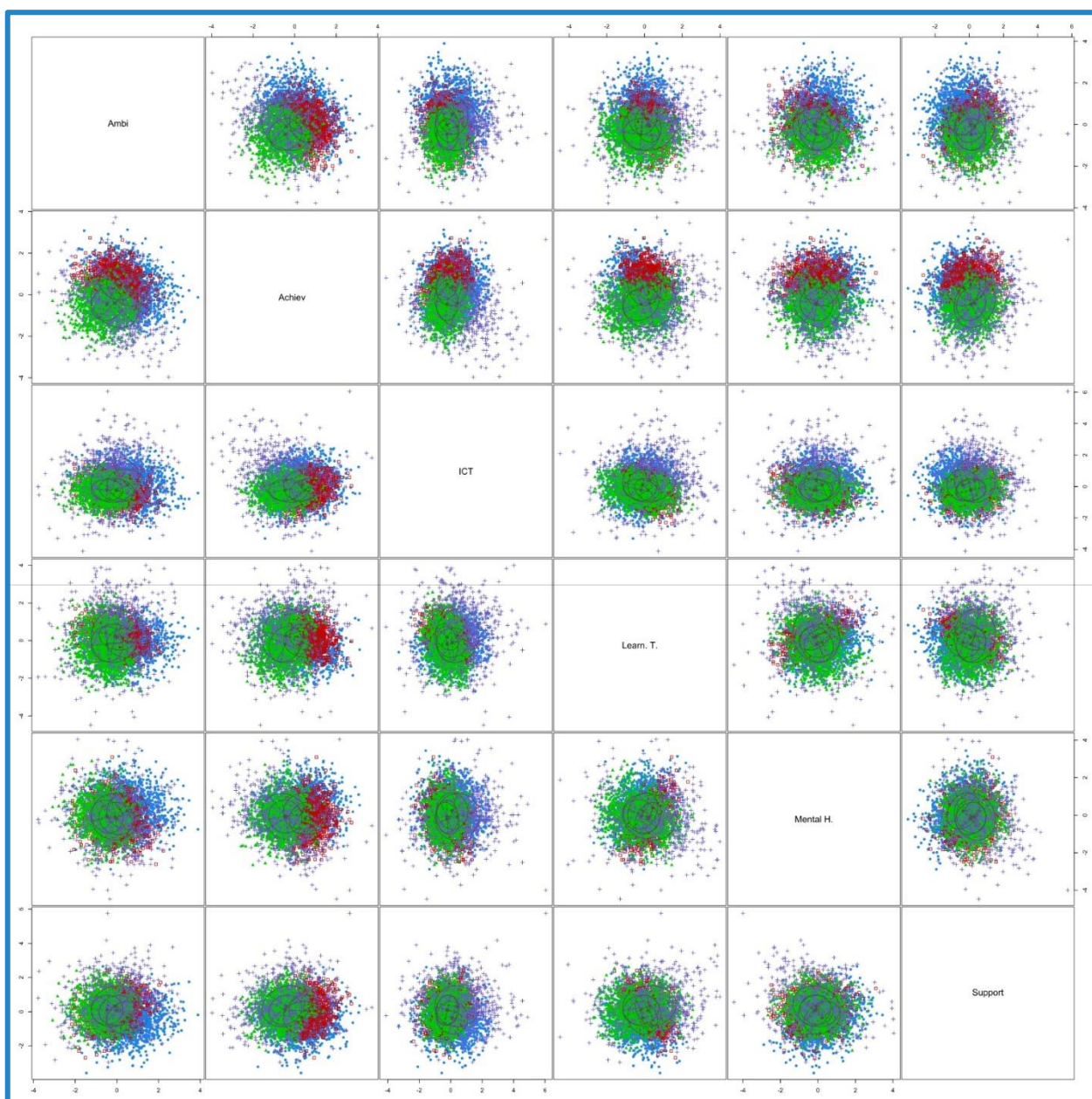


Gráfico 5- Gráfico de Classificação

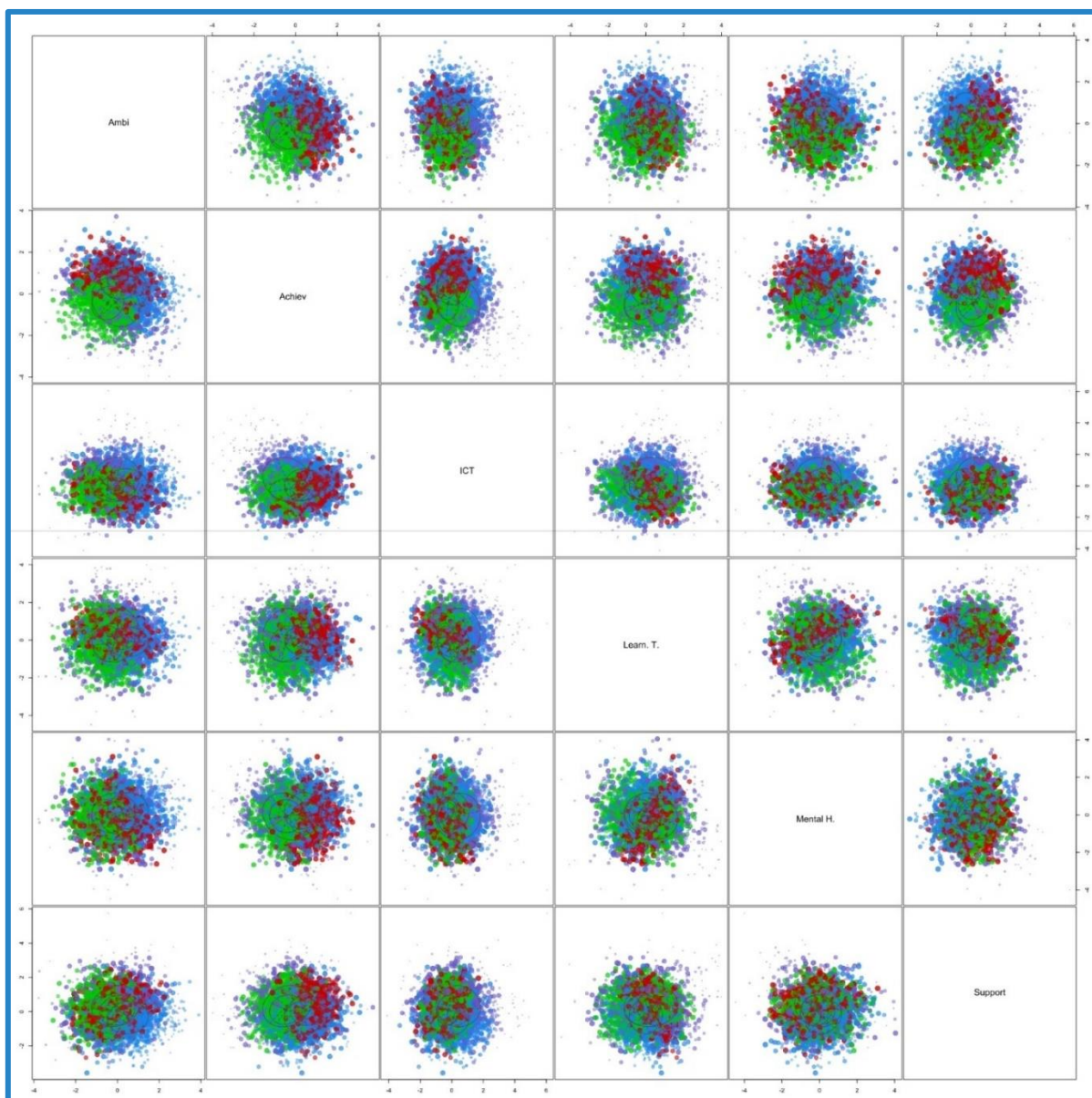


Gráfico 6 - Gráfico de Incerteza