

**Análise Exploratória do Índice de Desenvolvimento Humano: uma
aplicação do Self Organizing Map na segmentação dos países**

por

Catarina Isabel Agostinho Cordas

Dissertação apresentada como requisito parcial para a obtenção do Grau de
Mestre em Estatística e Gestão de Informação

pelo

Instituto Superior de Estatística e Gestão de Informação

da

Universidade Nova de Lisboa

2011

**Análise Exploratória do Índice de Desenvolvimento Humano: uma
aplicação do Self Organizing Map na Segmentação dos países**

por

Catarina Isabel Agostinho Cordas

Dissertação apresentada como requisito parcial para a obtenção do grau de
Mestre em Estatística e Gestão de Informação

pelo

Instituto Superior de Estatística e Gestão de Informação

da

Universidade Nova de Lisboa

Dissertação orientada por
Professor Doutor Fernando Bação

2011

Agradecimentos

A elaboração da presente dissertação teve o contributo de diversas pessoas a quem não posso deixar de agradecer.

Um agradecimento especial à Fernanda Carvalho que me deu a ideia, à Professora Doutora Manuela Silva pela oportunidade e confiança depositada em mim, ao Professor Doutor Roberto Henriques pela disponibilidade e explicações do SOM e ao Professor Doutor Adolfo Yáñez.

Ao Professor Doutor Fernando Bação, o meu orientador, um agradecimento muito especial pela orientação, flexibilidade, disponibilidade e sugestões, durante a elaboração desta.

Agradeço também à minha família e amigos, pelo apoio e compreensão da minha ausência nos últimos meses.

Análise Exploratória do Índice de Desenvolvimento Humano: uma aplicação do Self Organizing Map na Segmentação dos países

Resumo

Em 1990 as Nações Unidas através do seu Programa para o Desenvolvimento, introduziu o Índice de Desenvolvimento Humano, que se tornaria um dos índices mais utilizados para comunicar o desenvolvimento dos países. Todos os anos este índice é publicado, classificando os países em quatro categorias: “Desenvolvimento Humano Baixo”, “Desenvolvimento Humano Médio”, “Desenvolvimento Humano Alto” e “Desenvolvimento Humano Muito Alto”.

A simplicidade deste índice (média da realização dos países em três dimensões: saúde, educação e rendimento), associado ao facto de ele defender que o desenvolvimento humano é mais do que apenas crescimento económico, são a chave do seu sucesso. No entanto, desde o seu lançamento que este índice tem originado um intenso debate e algumas críticas, entre as quais a escolha das dimensões de desenvolvimento, deixando dimensões importantes como a desigualdade (económica e de género) e a liberdade política.

O objectivo principal da dissertação que me proponho realizar é aplicar uma técnica de *Data Mining*, mais concretamente a rede neuronal Self Organizing Map, na segmentação dos países que compõem o Relatório do Desenvolvimento Humano das Nações Unidas, permitindo desta forma segmentar os países utilizando mais dimensões do que as actualmente utilizadas pelo Índice de Desenvolvimento Humano. Os dados utilizados estão publicados no website do Programa para o Desenvolvimento das Nações Unidas e será utilizado um software desenvolvido pelo ISEGI que aplica o algoritmo, o GeoSOM Suit. Os resultados obtidos sugerem que há diferenças de classificação de alguns países quando aplicadas outras dimensões de desenvolvimento humano, nomeadamente *Empowerment*, Desigualdade e Sustentabilidade.

Exploratory Data Analysis of Human Development Index: apply Self Organizing Map in the segmentation of the countries

Abstract

In 1990 the United Nations through its Development Program introduced the Human Development Index, which would become one of the most used indices to communicate the countries development. Every year this index is published, classifying the countries into four categories: “Low Human Development”, “Medium Human Development”, “High Human Development” and “Very High Human Development”.

The simplicity of this index (average of three dimensions: health, education and income), associated to the fact of it defends that the human development is more than just economic growth, are the key of its success. However, since it was launched, has originated a deep discussion with some criticism, like the selection of the human development dimensions that don't include important dimensions such as inequality and political freedom.

The main objective of this dissertation is to apply a Data Mining technique, more specifically the Self Organizing Map, in the segmentation of the countries of United Nations Human Development Report, allowing this way to target the countries using more dimensions than the ones used nowadays by the Human Development Index. The data used are published at the website of the United Nations Development Program and will be used a software developed by ISEGI that applies the algorithm – Geo SOM Suit. The results obtained suggest that there are differences of classification in some countries when applied others human development dimensions, namely Empowerment, Inequality and Sustainability.

Palavras Chave

Clustering

Desenvolvimento Humano

IDH

Índice de Desenvolvimento Humano

Modelo de Capacidades

Relatório de Desenvolvimento Humanidade

Self Organizing Map

SOM

Keywords

Clustering

HDI

Human Capabilities Model

Human Development

Human Development Index

Human Development Report

Self-Organizing Map

SOM

Índice do Texto

Agradecimentos	iv
Resumo	v
Abstract.....	vi
Palavras Chave.....	vii
Keywords.....	vii
Índice do Texto	viii
Índice de Figuras.....	x
Lista das Abreviaturas.....	xii
1. Introdução.....	1
1.1. Enquadramento.....	1
1.2. Problemática.....	3
1.3. Questões de investigação e objectivos.....	4
1.4. Organização da Dissertação	4
2. Índice de Desenvolvimento Humano	7
2.1. O IDH e o conceito de Desenvolvimento Humano	7
2.2. Fórmula de cálculo do (Antigo) IDH	10
2.3. Fórmula de Cálculo IDH 2010.....	13
2.4. As críticas mais frequente do IDH 1990 e como foram incorporadas.....	15
2.5. Impacto da nova fórmula de cálculo na classificação dos países.....	21
2.6. Índices Complementares	22
3. Algoritmo SOM	25
3.1. Descrição Sumária do Algoritmo SOM.....	25
3.2. Vantagens da Utilização do SOM na visualização de clusters	28
3.3. Estudo que aplicou o SOM ao Bem-estar e Pobreza no Mundo	29
3.4. Utilização do GeoSOM Suit.....	31
4. Metodologia.....	33
4.1. Compreensão dos Dados.....	33
4.2. Preparação dos Dados.....	34
4.3. Aplicação do Modelo.....	36

4.4.	Avaliação do modelo	42
5.	Resultados.....	45
5.1.	Interpretação do Plano de Componentes	48
5.2.	Identificação.....	49
5.3.	Caracterização dos Clusters	50
5.4.	Comparação dos segmentos com a classificação do IDH.....	54
6.	Conclusões.....	57
7.	Limitações e recomendações para trabalhos futuros.....	59
	ANEXOS	61
	Anexo A: Definição do Conceito de Desenvolvimento Humano do 1º Relatório.....	62
	Anexo B: Definição dos Indicadores Usadas	63
	Anexo C: Experiências com o SOM: Rede usando método K-means	64
	Anexo D: Comparação dos Clusters com a Classificação do IDH.....	66
	Referências Bibliográficas.....	69

Índice de Figuras

<i>FIGURA 1-1</i> PESQUISAS NA INTERNET DO IDH.....	2
<i>FIGURA 2-1</i> - DIMENSÕES REFERIDAS NOS DIFERENTES RELATÓRIOS.....	10
<i>FIGURA 2-2</i> - ALTERAÇÕES IDH ENTRE 1990 E 2010.....	11
<i>FIGURA 2-3</i> - FRACA RELAÇÃO ENTRE O CRESCIMENTO ECONÓMICO E AS ALTERAÇÕES DA SAÚDE E EDUCAÇÃO ENTRE 1970 E 2010.....	16
<i>FIGURA 2-4</i> - DISTRIBUIÇÃO DO IDH PUBLICADO PELO RDH ENTRE 1990 E 2005.....	21
<i>FIGURA 2-5</i> - IDH 2010: METODOLOGIA NOVA VS METODOLOGIA ANTERIOR.....	22
<i>FIGURA 3-1</i> – ILUSTRAÇÃO DOS PRINCÍPIOS DO SOM.....	26
<i>FIGURA 3-2</i> - EXEMPLO DE UMA GRELHA SOM.....	26
<i>FIGURA 3-3</i> – EXEMPLO DE UM SOM.....	29
<i>FIGURA 3-4</i> - UTILIZAÇÃO DA U-MATRIX PARA REPRESENTAR OS PAÍSES SEGUNDO O BEM-ESTAR (KOHONEN & S.KASKI, 1996).....	30
<i>FIGURA 3-5</i> – DISTRIBUIÇÃO DOS INDICADORES DOS PAÍSES DO SOM NAS VARIÁVEIS QUE CONSTROEM O ALGORITMO.....	30
<i>FIGURA 3-6</i> – EXEMPLO DA VISUALIZAÇÃO DAS VÁRIAS JANELAS DINÂMICAS DO GEOSOM SUIT QUE PERMITEM DE UMA FORMA DIRECTA E FÁCIL VISUALIZAR VÁRIOS ASPECTOS IMPORTANTES DA SEGMENTAÇÃO.....	31
<i>FIGURA 4-1</i> – CRUZAMENTO DOS PAÍSES DE ACORDO COM O VALOR DO IDH (EIXO DOS X) E DO NÍVEL DE DESIGUALDADE DE GÉNERO.....	33
<i>FIGURA 4-2</i> – TABELA DOS INDICADORES USADOS NA ANÁLISE POR CADA DIMENSÃO.....	35
<i>FIGURA 4-3</i> – PAÍSES NÃO INCLUÍDOS NA ANÁLISE.....	36
<i>FIGURA 4-4</i> – TABELA DESCRITIVA SUMÁRIA DOS DADOS.....	37
<i>FIGURA 4-5</i> – MATRIZ DE CORRELAÇÃO DAS VARIÁVEIS.....	39
<i>FIGURA 4-6</i> – FORMULÁRIO DE INTRODUÇÃO DOS PARÂMETROS NO GEOSOM SUIT.....	39
<i>FIGURA 4-7</i> – TIPO DE TOPOLOGIA DO SOM.....	40
<i>FIGURA 4-8</i> – DIFERENTES TIPOS DE FORMATO DO SOM.....	40
<i>FIGURA 4-9</i> – RESUMO DOS MODELOS FEITOS PARA DIFERENTES TAXAS DE APRENDIZAGEM.....	42
<i>FIGURA 4-10</i> – MAPA UMAP SELECCIONADO.....	42
<i>FIGURA 5-1</i> - CLUSTERS CRIADOS NO UMAP (ENCARNADO) E SUGERIDOS (AZUL ESCURO).....	45
<i>FIGURA 5-2</i> - CP NÚMERO DE PARTOS COM ASSISTÊNCIA MÉDICA (%).....	46
<i>FIGURA 5-3</i> – CP RÁCIO NA PARTICIPAÇÃO NO MERCADO DE TRABALHO (F/M).....	46
<i>FIGURA 5-4</i> – CP % DE LUGARES FEMININOS NO PARLAMENTO.....	46
<i>FIGURA 5-5</i> – CP COEFICIENTE GINI.....	46
<i>FIGURA 5-6</i> - CP IDH SEM RENDIMENTO.....	46
<i>FIGURA 5-7</i> – CP RENDIMENTO (IDH).....	46
<i>FIGURA 5-8</i> – CP TAXA DE FERTILIDADE ADOLESCENTE.....	46

<i>FIGURA 5-9 – CP PESSOAS SEM ACESSO A SERVIÇOS DE ÁGUA</i>	46
<i>FIGURA 5-10 – CP MÉDIA DE ANOS DE ENSINO</i>	46
<i>FIGURA 5-11 – CP MÉDIA DE ANOS ESCOLARES ESPERADA</i>	46
<i>FIGURA 5-12 – CP ESPERANÇA DE VIDA</i>	46
<i>FIGURA 5-13 – CP ÍNDICE DESENVOLVIMENTO HUMANO</i>	46
<i>FIGURA 5-14- CP NÍVEL DE VIOLAÇÃO DOS DIREITOS HUMANOS (1-5)</i>	47
<i>FIGURA 5-15 – CP NÍVEL SE SATISFAÇÃO COM A LIBERDADE DE ESCOLHA (%)</i>	47
<i>FIGURA 5-16 – CP NÍVEL DE DEMOCRACIA (0 NÃO DEM; 1 DEMS/ALTERNATIVA; 2 DEM)</i>	47
<i>FIGURA 5-17 % DO CRESCIMENTO DO NÚMERO DE UTILIZADORES DE INTERNET (2000-2008)</i>	47
<i>FIGURA 5-18 – CP % DA POPULAÇÃO EMPREGADA (15-64)</i>	47
<i>FIGURA 5-19 – CP % DA POPULAÇÃO URBANA</i>	47
<i>FIGURA 5-20 – CP ÁREA DO TERRITÓRIO PROTEGIDA (5 DO PAÍS)</i>	47
<i>FIGURA 5-21 – CP EMISSÃO DE CO2 PER CAPITA</i>	47
<i>FIGURA 5-22- CP % DE PESSOAS QUE OMITI A OPINIÃO A REPRESENTANTES OFICIAIS)</i>	47
<i>FIGURA 5-23 – CP JORNALISTAS PRESOS (NÚMERO)</i>	47
<i>FIGURA 5-24 – CP ÍNDICE LIBERDADE DE IMPRENSA (MAIS BAIXO MAIS LIBERDADE)</i>	47
<i>FIGURA 5-25 - SEGMENTAÇÃO DOS PAÍSES</i>	50
<i>FIGURA 5-26 – MÉDIA DOS INDICADORES DOS SEGMENTOS ENCONTRADOS</i>	52
<i>FIGURA 5-27 – QUADRO QUE CRUZA OS CLUSTERS DO SOM COM O IDH</i>	54
<i>FIGURA 5-28 - MÉDIA DOS INDICADORES DO IDH</i>	56

Lista das Abreviaturas

BMU – Best Matching Unit

HPI – Human Poverty Index

IDH – Índice de Desenvolvimento Humano

IDHAD – Índice de Desenvolvimento Humano Ajustado à Desigualdade

IDG - Índice de Desigualdade de Género

KDD – Knowledge Discovery in Databases

GDI – Gender Related Development Index

GEM – Gender Empowerment Measure

ONU – Organização das Nações Unidas

OMS – Organização Mundial de Saúde

PCS – Plano de Componentes

PIB per capita – Produto Interno Bruto per capita

PNB – Produto Nacional Bruto

PNUD – Programa das Nações Unidas para Desenvolvimento

RD – Rendimento Disponível

RDH – Relatório de Desenvolvimento Humano

RNA – Redes Neurais Artificiais

RNB – Rendimento Nacional Bruto

SOM – Self Organizing Map

UMAP – Matriz U

1. Introdução

1.1. Enquadramento

O Programa das Nações Unidas para o Desenvolvimento (PNUD) foi criado a partir da resolução da Assembleia Geral da Organização das Nações Unidas (ONU) em 1965 e é o principal órgão para o desenvolvimento desta Organização. O PNUD é também o responsável pela publicação dos Relatórios de Desenvolvimento Humano (RDH). Foi no âmbito destes relatórios que foi lançado o Índice de Desenvolvimento Humano (IDH) em 1990, que tem grande influência na teoria e prática do desenvolvimento humano em todo o mundo.

De forma a compreender o IDH é importante entender o conceito de desenvolvimento humano que o sustenta. O conceito de desenvolvimento humano do IDH foi baseado no trabalho de Amartya Sen (Sen, 1984), onde o Homem é colocado no centro da sua análise, sendo o rendimento apenas uma parte do bem-estar do ser humano e não o seu fim. O conceito de Desenvolvimento do RDH é assim definido como um processo que conduz ao aumento das capacidades oferecidas a cada um, sendo essas capacidades ilimitadas. Este conceito de desenvolvimento não é inovador, remonta pelo menos a Aristóteles e está presente nos fundadores da Economia Moderna (UNDP, 1990, p. 10).

Esta nova concepção desafiou a forma de medir as capacidades, por um lado devido à falta de dados, por outro, porque a inclusão de muitas variáveis num Índice poderia torná-lo demasiado complexo para ser compreendido e implementado. De forma a resolver este problema, o PNUD criou o IDH, que não é mais que uma média de três dimensões: Saúde, Educação e Rendimento, permitindo ordenar os países de acordo com o seu nível de desenvolvimento em 4 grupos: “Muito Alto”, “Alto”, “Médio” e “Baixo”.

O IDH não pretende ser uma medida de bem-estar, mas de capacidades. Desta forma o IDH pretende medir o acesso às diferentes capacidades nos países e comparar esse acesso, de forma a influenciar a direcção de políticas sociais internacionais.

O IDH é muito popular, segundo o New York Times, é a única medida alternativa ao PIB *per capita* bem-sucedida (Gertner, 2010). Se fizermos uma pesquisa no Google Insights sobre IDH, podemos verificar, que o RDH tem mais pesquisas que o seu concorrente (Relatório de Desenvolvimento do Banco Mundial), como se pode ver na Figura 1-1. A procura por IDH é superior ao RDH, e os picos referem-se aos períodos de lançamento, o que é altamente mediático, atraindo as atenções da sociedade civil, políticos e académicos.

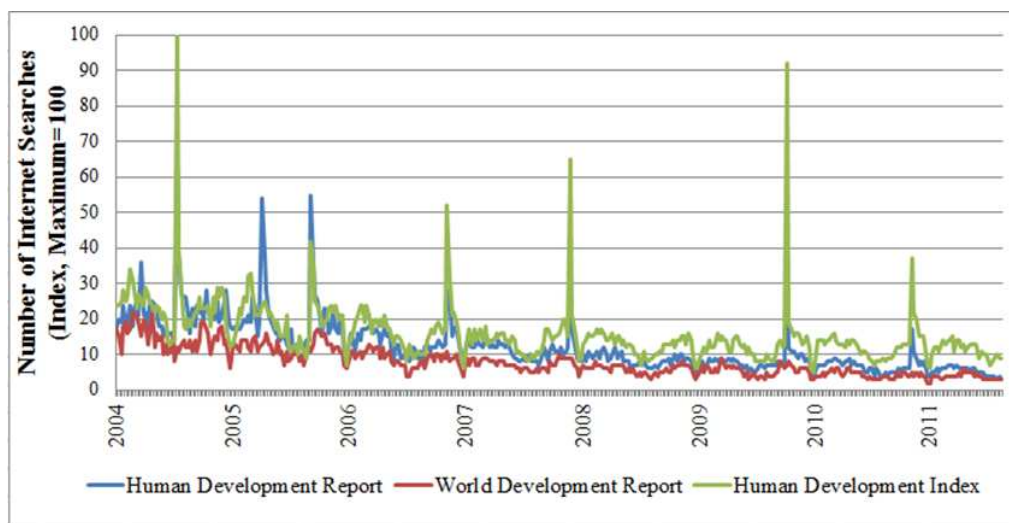


Figura 1-1 Pesquisas na internet do IDH

Fonte: Google Insights 15 de Agosto 2011

Nota: Os números no gráfico reflectem quantas pesquisas foram feitas desde 2004 em relação ao número total de pesquisas feitas no Google ao longo desse período. Eles não representam os números do volume de pesquisas absoluto, pois os dados são normalizados e apresentados numa escala de 0 a 100. Cada ponto no gráfico é dividido pelo ponto mais alto, ou seja, 100. Quando não temos dados suficientes, 0 é exibido. Data da Pesquisa: 15 de Agosto 2011.

Apesar de não ter sido criado com o fim de direccionar a política de cooperação entre o Norte e o Sul, de acordo com um artigo publicado pelo *Health Public Outlook* (Hendrik Wolff, 2010) a classificação da ONU é utilizada na diferenciação de preços por algumas empresas farmacêuticas de forma a ajudar os países com o Desenvolvimento Humano mais baixo. Desde 2001 que a empresa farmacêutica Merck vende os seus medicamentos com 90% de desconto aos países classificados como “Baixo Desenvolvimento” e com 75% de desconto aos que são classificados como “Médio. Desde o seu início que o IDH tem sido utilizado na construção e discussão

publica de políticas de desenvolvimento pelos governos e Organizações Não Governamentais, por exemplo, o protocolo de Kyoto interligou as propostas da responsabilidades dos países de acordo com a classificação do IDH.

1.2. Problemática

Apesar da sua popularidade, desde o seu lançamento que o índice tem originado um intenso debate académico com críticas e sugestões de melhoria. O PNUD incentiva e patrocina esse debate. Desde a primeira publicação do RDH que é salientando que o IDH não pretende ser uma medida fechada, estando aberto a críticas e melhorias. As contribuições chave foram publicadas no *Journal of Human Development and Capabilities* e discutidas na reunião anual da Associação *Human Development and Capabilities* (Klugman, et al., 2011) .

Este debate produziu alterações na fórmula de cálculo do IDH, mas as dimensões de desenvolvimento humano mantêm-se iguais, deixando de fora dimensões importantes como a equidade, sustentabilidade e algumas liberdades, sendo uma das principais críticas desde o início. De tal forma há consciência que este aspecto constitui uma limitação do índice, que no relatório comemorativo dos 20 anos este tema é explorado num capítulo completo com o seguinte título “Good things don’t always come together” (UNDP, 2010, p. 66) onde estas dimensões são analisadas.

Para fazer face a este problema do IDH deixar de fora algumas dimensões do desenvolvimento humano na classificação dos países, o presente trabalho pretende testar através de um técnica de *Data Mining* uma forma de segmentar os países segundo o seu nível de desenvolvimento humano, tendo como base várias dimensões. Devido às suas capacidades de *clustering*, o algoritmo SOM parece apropriado para resolver este problema e desta forma será utilizado. Utilizar-se-á os dados disponíveis no site do PNUD e um software gratuito disponível no site do Instituto Superior de Estatística e Gestão de Informação, o GeoSOM Suit¹ , o qual permite uma visualização interligada e dinâmica para definir os clusters.

¹ www.isegi.unl.pt/labnt/GeoSOM Suit

O trabalho desenvolvido e aqui apresentado, pretende ser uma contribuição de uma área recente do conhecimento, o *Datamining*, na forma de analisar o problema do desenvolvimento humano.

1.3. Questões de investigação e objectivos

Na sequência do problema identificado, foi definido um objectivo genérico. Este consiste na aplicação do SOM a um conjunto de dados que representam várias dimensões do desenvolvimento humano e segmentar os países de acordo com o seu nível de desenvolvimento. Este objectivo genérico incorpora os seguintes objectivos específicos:

1. Analisar de uma forma critica a forma de calcular o IDH desde a sua criação e como este segmenta os países;
2. Identificar, justificando o critério, as dimensões relevantes na classificação do desenvolvimento humano dos países;
3. Encontrar segmentos de países de acordo com o seu tipo de Desenvolvimento Humano;
4. Comparar a segmentação apurada pelo trabalho com a classificação actual do IDH identificando as principais diferenças.

1.4. Organização da Dissertação

Após uma introdução ao tema, é feita uma breve revisão do conceito que sustenta o IDH (2.1), a forma como este era inicialmente calculado (2.2) e a fórmula de cálculo em 2010 (2.3). No capítulo 2 são ainda sintetizadas as críticas mais frequentes ao IDH original e de que forma estas foram incorporadas no IDH de 2010 (2.4). Por fim é analisado o impacto da nova fórmula na classificação dos países (2.5) e os indicadores complementares criados em 2010 para analisar mais dimensões do que as englobadas pelo IDH (2.6). No ponto 3 são sintetizadas as principais características do algoritmo SOM e são revistos, de uma forma breve, estudos que utilizaram o SOM para segmentar

países de acordo com o Bem-estar (3.3). Neste capítulo é também feita uma breve explicação do GeoSOM Suit. No ponto 4 é apresentada a metodologia desenvolvida nas fases de um processo de análise exploratória de dados, que pretende sobretudo extrair conhecimento a partir dos mesmos. Assim, é explicado o processo que se seguiu para se compreender os dados (4.1), as técnicas que foram usadas para os preparar (4.2) e os parâmetros usados na aplicação do SOM (4.3). É também feita a avaliação do modelo. No ponto 5 os resultados são analisados, recorrendo ao Plano de Componentes (5.1), são identificados os clusters (5.2) e descrito o perfil desses clusters extraindo informação relevante e original. No ponto 6 são apresentadas as conclusões, comparando o modelo com a classificação do IDH. Por fim, no ponto 7 são enumeradas as limitações do estudo e feitas recomendações para trabalhos futuros.

2. Índice de Desenvolvimento Humano

2.1. O IDH e o conceito de Desenvolvimento Humano

O primeiro Relatório de Desenvolvimento Humano foi publicado em Londres no dia 24 de Maio de 1990 como se pode ler no livro de reflexões do seu principal impulsionador Mahbub ul Haq (Haq, 1995, p. 26). O primeiro capítulo do Relatório intitula-se “*Defining and Measuring Human Development*” e começa com a famosa frase, que sintetiza o objectivo do conceito de desenvolvimento proposto:

(UNDP, 1990, p. 10) “*People are the real wealth of a nation. The basic objective of development is to create an enabling environment for people to live long, healthy and creative lives. This may appear to be a simple truth but it is often forgotten in the immediate concern with the accumulation of commodities and financial wealth*”.

Mais à frente neste capítulo o conceito de Desenvolvimento Humano é definido:

(UNDP, 1990, p. 10) “*Human development is a process of enlarging people’s choices. In principle, these choices can be infinite and change over the time. But at all levels of development the three essential ones are for people to lead a long and healthy life, to acquire knowledge and to have access to resources needed for a decent standard of living. If these essential choices are not available, many other opportunities remains inaccessible*”.

É neste sentido que o relatório propõe um índice (IDH) composto por três “capacidades” básicas:

1. **Vida saudável e longa** – utilizando o indicador de esperança de vida à nascença;
2. **Educação** – inicialmente utilizou-se a taxa de alfabetização. A partir de 1991 passou-se a calcular uma média ponderada em que a taxa de alfabetização tinha o ponderador de 2/3 e a média de anos de escolaridade 1/3;
3. **Nível de Vida** – PIB *per capita* ajustado pela Paridade de Poder de Compra.

Apesar da sua simplicidade, o Relatório deixa claro que o IDH tem limitações (UNDP, 1990, p. 16) “*To stress again an earlier point, the human development index captures a few of people's choices and leaves out many that people may value highly*

- *economic, social and political freedom (box 1.5), and protection against violence, insecurity and discrimination, to name but a few. The HDI thus has limitations.*”

De acordo com ul Haq o conceito de Desenvolvimento Humano é um conceito holístico que visa cobrir todos os aspectos do desenvolvimento. No entanto, o ponto principal segundo o autor é *“the widening of people’s choices and the enrichment of their lives. All aspects of life – economic, political or cultural – are viewed from that perspective”* (Haq, 1995, p. 20).

Para Haq o desenvolvimento humano tem quatro características essenciais:

1. **Equidade:** o acesso às possibilidades deve ser equitativo, mas isto não garante resultados iguais;
2. **Sustentabilidade:** a próxima geração deve ter as mesmas oportunidades que a actual. Segundo o autor esta dimensão é confundida com a renovação de recursos naturais, que é apenas um dos seus aspectos;
3. **Produtividade:** esta característica é fundamental e por isso requer investimento nas pessoas e um ambiente macroeconómico que lhes permita alcançar o seu potencial máximo;
4. **Empowerment:** as pessoas são agentes activos das políticas de desenvolvimento humano.

Para ul Haq, os pontos dos modelos tradicionais são abordados no seu modelo, mas numa perspectiva que coloca as pessoas no centro da análise.

As bases teóricas do IDH estão no modelo de capacidades do Nobel de Economia Amartya Sen como se pode ler numa entrevista dada por este na página do PNUD (Sen, 2010). Sen era um antigo colega e amigo de ul Haq que se baseou no seu livro (Sen, 1984), para a construção do IDH. Neste livro de Sen o modelo das capacidades é explicado e de uma forma muito resumida, segundo este modelo os bens possuem características e as pessoas ao possuírem os bens, passam a comandar tais características. Sen defende, que não se deve considerar apenas as características dos bens, mas também o que as pessoas são capazes de fazer a partir dos bens de acordo com as suas capacidades e é neste sentido que este conceito de desenvolvimento coloca o Homem no centro da sua análise.

O IDH centrou-se em três capacidades: viver uma vida longa e saudável, ter acesso ao conhecimento e ter um nível de vida decente. Este índice segue seis princípios descritos pelo criador do relatório: (i) medir as capacidades básicas do modelo de capacidades; (ii) manter um número limitado de variáveis de forma a permitir que seja perceptível e simples a sua compreensão; (iii) ser um índice composto e não um conjunto de vários índices; (iv) cobrir dimensões sociais e económicas; (v) ser suficientemente flexível para permitir a reformulação incorporando as alternativas de melhoramento propostas; (vi) não estar limitado à falta de informação/indicadores. (Haq, 1995)

Para Sen, esta era uma forma incompleta de medir todas as capacidades. (Sen, 2010), no entanto, ao contrário de outros autores como por exemplo (Nussbaum, 2000), Sen nunca construiu uma lista com essas capacidades, sendo mesmo contrário ao conceito de desenvolvimento que fechasse e tornasse estáticas essas capacidades.

De forma a resolver este problema, o RDH tenta incluir mais dimensões do que as que estão no IDH da seguinte forma:

1. Desenvolver um tema diferente sobre desenvolvimento humano anualmente, com informação detalhada sobre esse tema;
2. Incluir em apêndice várias tabelas com informação de outras dimensões;
3. Desenvolver índices complementares, de forma a medir outras dimensões de desenvolvimento humano.

A Figura 2-1 sintetiza as várias dimensões que são mencionadas em cada relatório. As dimensões Saúde, Educação e Rendimento são sempre mencionadas, uma vez que fazem parte do IDH, todas as outras dimensões como a Equidade, Liberdade, Sustentabilidade, são referidas apenas em alguns.

Dimensions mentioned in HDR by year	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	07	09
Long healthy life	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Knowledge	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Resources for decent std of life	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Political freedom	X	X						X							X			X
Guaranteed Human Rts	X					X		X		X	X							
Self Respect	X					X		X		X	X							
Good physical environment		X	X		X													
Freedom of Action & Expression			X															
Participation				X	X								X					
Human Security					X		X											
Political, Social & Econ Freedoms						X		X	X	X								
Being creative						X		X	X	X	X							
Being Productive						X		X	X	X	X							
Freedom							X	X										
Democracy							X											
Dignity & Respect of others								X										
Empowerment									X		X							
A sense of belonging to a community									X	X	X							
Security											X							X
Sustainability											X							
Enjoying political and civil freedoms to participate in the life of one's community.														X				
Cultural liberty															X			
Social & Political Participation																X		
Civil & Political Rights																	X	

Figura 2-1 - Dimensões referidas nos diferentes relatórios

Fonte: (Alkire, 2010)

2.2. Fórmula de cálculo do (Antigo) IDH

Até 2010 o IDH era calculado através de uma média aritmética simples de índices normalizados das dimensões saúde, educação e rendimento:

$$- \quad (1)$$

Onde os H_i são os “sub-índices” normalizados utilizando máximos e mínimos pré definidos, para as dimensões saúde (h - *health*), educação (e) e nível de vida (ls - *living standard*), respectivamente. Estes índices eram normalizados de forma que variassem de 0 a 1, permitindo que na média aritmética da fórmula (1) tivessem todos a mesma relevância.

HDR	Bounds	Indicators			Treatment of Income	Aggregation Formula	
		Health	Education	Income			
1990	Observed	Life expectancy at birth (UN Population Division)	Adult literacy rate (25+) (UNESCO)	Real GDP per capita PPP \$ (log) (World Bank)	Logarithmic transformation with a cap	Arithmetic mean	
1991 ~ 1993			(2/3)Adult literacy rate (UNESCO) (1/3)Mean years of schooling (UNESCO)		Atkinson formula with threshold value derived from poverty line		
1994			(2/3)Adult literacy rate index (UNESCO) (1/3)Mean years of schooling index (UNESCO)	Real GDP per capita PPP \$ (adjusted) (World Bank)	Atkinson formula with threshold value derived from global average		
1995 ~ 1998	(2/3)Adult literacy rate index (UNESCO) (1/3)Combined gross enrolment ratio index with a cap starting to bind in 1996 (UNESCO)						
1999							
2000 ~ 2009	(2/3)Adult literacy rate (15+) index (UNESCO) (1/3)Combined gross enrolment ratio index with a cap starting to bind in 1996 (UNESCO)		Real GDP per capita PPP \$ (log) (World Bank)	Logarithmic transformation with a cap starting to bind in 2001			
2010	Upper: Observed Lower: Fixed		(1/2)Mean years of schooling index (Barro-Lee) (1/2)Expected years of schooling index (UNESCO)	Real GNI per capita PPP \$ (ln) (World Bank)	Natural logarithmic transformation without a cap		Geometric mean

Figura 2-2 - Alterações IDH entre 1990 e 2010

Fonte: (Klugman, et al., 2011)

O IDH permite ordenar os países de acordo com o seu nível de desenvolvimento, agrupando-os em três grupos mutuamente exclusivos – “Alto Desenvolvimento” (IDH superior a 0,8), “Médio Desenvolvimento” (IDH entre 0,5 e 0,8) e “Baixo Desenvolvimento” (IDH inferior a 0,5).

Nos últimos 20 anos a IDH sofreu algumas alterações, que estão sintetizadas na Figura 2-2, que podemos destacar:

1. **A fixação do máximo e do mínimo na normalização das variáveis.** No primeiro relatório era o máximo e mínimo observado, a partir de 1994 o máximo e o mínimo passaram a ser valores fixos, motivado pela ideia que há valores máximos e mínimos aceitáveis para estas variáveis e desta forma é possível comparar historicamente o IDH².
2. **Alterações da dimensão da educação.** No primeiro relatório esta dimensão era medida pela Taxa de Alfabetização, mas em 1995 passou a agregar também a Taxa de Escolarização, uma vez que em alguns países, sobretudo os mais desenvolvidos, a Taxa de Alfabetização é muito alta e não diferencia.³
3. **Alteração da dimensão do rendimento.** Esta foi a dimensão que sofreu mais alterações ao longo do tempo. A grande alteração foi dada com a substituição da fórmula de Atkinson's⁴ de rendimento pela simples fórmula logarítmica com um máximo e mínimos fixos.

Conforme foi referido no ponto 2.1, uma forma que o RDH encontrou para analisar as dimensões de desenvolvimento que não estão no IDH, como por exemplo a Desigualdade de Género, foi desenvolver índices complementares. Assim, em 1995 o RDH introduziu dois índices complementares relacionados com a Desigualdade de Género:

² Nota técnica (UNDP, 1994)

³ Nota técnica (UNDP, 1995)

⁴ $W(y) = y^*$ for $0 < y \leq y^*$
 $= y^* + 2[(y-y^*)/2]$ for $y^* < y \leq 2y^*$
 $= y^* + 2(y^*/2) + 3[(y-2y^*)/3]$ for $2y^* < y \leq 3y^*$
 $= y^* + 2(y^*/2) + 3(y^*/3) + \dots + n[(y-(n-1)y^*)/n]$ for $(n-1)y^* < y \leq ny^*$.

1. Gender Related Development Index (GDI) usa as mesmas variáveis que o IDH, ajustando-as à desigualdade entre géneros. Quando comparado com o IDH permite identificar se há ou não desigualdade de género nas capacidades analisadas pelo IDH;
2. Gender Empowerment Measure (GEM) pretende reflectir de que forma homens e mulheres são capazes de participar de uma forma activa na vida política e económica dos países.⁵

Em 1997 o RDH introduziu os Human Poverty Indices (HPI-1 e HPI-2) de forma a responder à necessidade de medir privações dos mais desfavorecidos. O HPI-2 era aplicado aos países mais industrializados, com targets ligeiramente superiores ao HPI-1, que era aplicado aos países em vias de desenvolvimento e incluindo o indicador de exclusão social.⁶

Outros índices foram criados, de uma forma inconsistente e que não se repetiu mais do que um ano. Por este motivo, o presente trabalho não os irá abordar.

No último relatório publicado à data (UNDP, 2010), o PNUD introduziu várias alterações em simultâneo ao IDH, bem como novos índices complementares. Veremos de seguida as alterações que foram feitas ao índice.

2.3. Fórmula de Cálculo IDH 2010

Em 2010 o Relatório do Desenvolvimento Humano introduziu várias alterações à fórmula de calcular o IDH, passando a ter a seguinte fórmula:

$$IDH = \sqrt[3]{(H_{Health} \times H_{Education} \times H_{Living\ Standard})} \quad (2)$$

Os índices H continuam a ser índices de realização normalizados para cada dimensão, mas os indicadores de algumas dimensões foram alterados. Assim cada dimensão é calculada da seguinte forma:

⁵ Para mais detalhes sobre a fórmula de cálculo deste índice complementar consultar nota técnica (UNDP, 1995)

⁶ Para mais detalhes sobre a fórmula de cálculo deste índice complementar consultar nota técnica de (Alkire, 2010) (UNDP, 1997)

$$H_{Health} = \frac{le - le_{min}}{le_{max} - le_{min}} \quad (3)$$

$$H_{Education} = \sqrt{\frac{mys - mys_{min}}{mys_{max} - mys_{min}} \times \frac{eys - eys_{min}}{eys_{max} - eys_{min}}} \quad (4)$$

$$H_{Living\ Standard} = \frac{\log GNI - \log GNI_{min}}{\log GNI_{max} - \log GNI_{min}} \quad (5)$$

A esperança de vida (*le*) continua a ser o indicador para a dimensão da saúde, o Rendimento Nacional (GNI) substitui o PIB per capita na dimensão do nível de vida e o nível de educação é medido pela média dos anos de escolaridade (*mys*) e anos de escolaridade esperados (*EYS*), em vez da taxa de iliteracia e nível de escolaridade. A acrescentar a estas alterações, três Índices Complementares foram introduzidos em 2010 de forma a captar privações e desigualdades: IDH ajustado à desigualdade, Índice de Desigualdade de Género e Índice de Pobreza Multidimensional.

Outra alteração foi a fixação dos valores máximos e mínimos. Os valores máximos foram fixados nos valores máximos observados dos indicadores dos países na série cronológica entre 1980 e 2010. Os valores mínimos são os valores considerados de subsistência ou zeros “naturais” para as variáveis. Para a esperança de vida foi fixado como valor mínimo 20 anos baseado em dados históricos de longo prazo. O valor fixado para mínimos nas duas variáveis da educação foi 0 anos, porque uma sociedade ao limite pode subsistir sem educação formal. O mínimo do GNI per capita fixado foi de USD 163, que é o valor mais baixo atingido por qualquer país desde que se começou a calcular o IDH⁷ e que corresponde a 45 cêntimos por dia, pouco acima de um terço da linha de pobreza de USD 1,25/dia do Banco Mundial.

Por fim a forma como os países são classificados também sofreu alterações. No passado, a classificação do IDH baseava-se em pontos de corte predefinidos dos valores do IDH (0,5; 0,8). Em 2010 as classificações baseiam-se em quartis e indicam um IDH muito alto, alto, médio e baixo. Como há 169 países, um grupo tem de ter mais um país do que os outros, o país extra foi atribuído ao grupo de IDH alto.

Exemplo de cálculo do IDH de 2010:

⁷ Zimbabwe 2008

Indicador	Value
Life expectancy at birth (years)	73.5
Mean years of schooling(years)	7.5
Expected years of schooling(years)	11.4
GNI per capita(PPP US\$)	7,263

Fonte: (UNDP, 2010, p. 216)

$$\text{Life expectancy index} = \frac{73.5 - 20}{83.2 - 20} = 0.847 \quad (3)$$

$$\text{Mean years of schooling index} = \frac{7.5 - 0}{13.2 - 0} = 0.568$$

$$\text{Expected years of schooling index} = \frac{11.4 - 0}{20.6 - 0} = 0.553$$

$$\text{Education index} = \frac{\sqrt{(0.568 \times 0.55)} - 0}{0.951 - 0} = 0.589 \quad (4)$$

$$\text{Income index} = \frac{\log(7,263) - \log(163)}{\log(108,211) - \log(163)} = 0.584 \quad (5)$$

$$\text{Human Development Index} = \sqrt[3]{0.847 \times 0.589 \times 0.584} = 0.663 \quad (2)$$

2.4. As críticas mais frequente do IDH 1990 e como foram incorporadas

Desde a sua primeira publicação que o IDH tem sido alvo de inúmeras críticas. Alguns autores, como por exemplo (Lind, 1992), (Dasgupta & Weale, 1992) e (Sagar & Najam, 1998), são mesmo radicais dizendo que o IDH dá uma imagem distorcida do mundo. Outro grupo de críticos (McGillivray, 1991), (Srinivasan, 1994), (Ravallion, 1997) e (Ogwan, 1994) não vêem a utilidade do IDH, uma vez que segundo eles um indicador centrado no PIB seria suficiente devido à forte correlação do rendimento com as restantes dimensões.

De facto, como se pode ver no primeiro gráfico da figura 2-3, há uma forte correlação entre o crescimento económico, as alterações da saúde e da educação, mas quando o rendimento é retirado do IDH (segundo gráfico da figura 2.3), a correlação passa a ser baixa entre esse índice e o PIB, salientando a importância de ter mais dimensões para além do rendimento na análise do desenvolvimento humano, sendo este o grande contributo do IDH para o debate político (Klugman, et al., 2011).

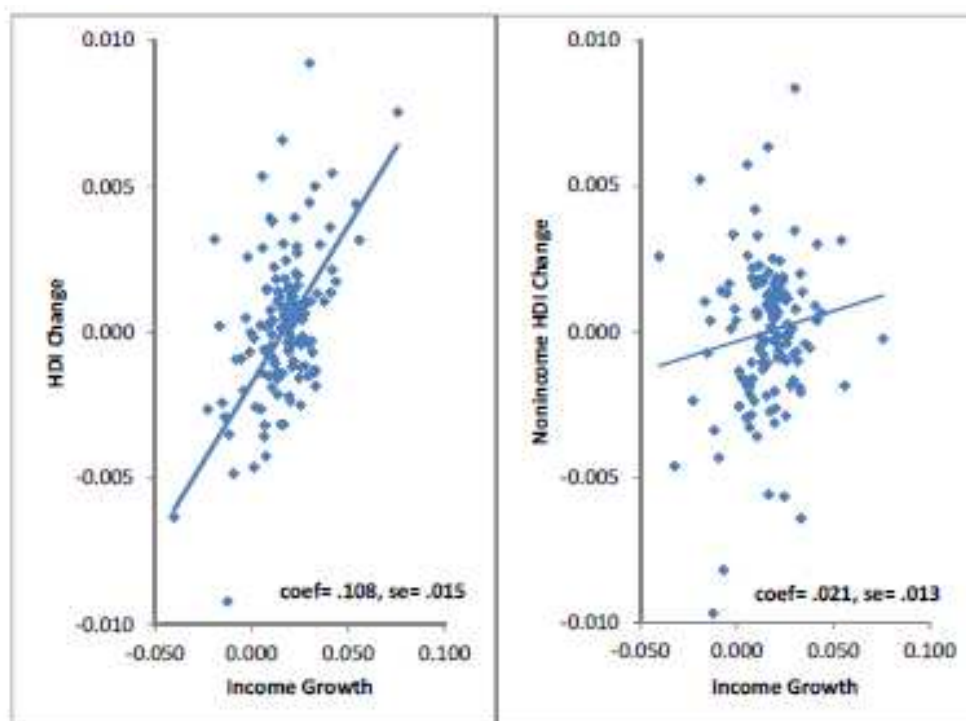


Figura 2-3- Fraca Relação entre o crescimento económico e as alterações da saúde e educação entre 1970 e 2010

Fonte: (Klugman, et al., 2011)

Alguns autores como por exemplo (Noorbakhsh, 1998), (Lind, 2004), (Chakravarty, 2003), (Despotis, 2005) e (Herrero, et al., 2010) sugerem outros índices, mas que são basicamente o IDH com algumas alterações, o presente trabalho não os utilizará.

De seguida iremos sintetizar as críticas que foram feitas nos últimos anos e a forma pela qual o último IDH de 2010 as incorporou.

2.4.1. Escolha das Variáveis

Dentro do grupo que critica a escolha das variáveis, dois tipos de críticas em relação às variáveis usadas pelo IDH podem ser identificados.

O primeiro grupo critica o facto de algumas dimensões importantes, como o ambiente (Sagar & Najam, 1998), não estarem incluídas. Nussbaum construiu em 2000 um modelo com uma lista de capacidades⁸ (Nussbaum, 2000), que na sua maioria não estão incluídas no IDH. Seria muito difícil ter todas as capacidades de Nussbaum, dado que algumas nem têm indicadores.

A falta da dimensão “liberdade política” tem sido mencionada ao longo dos últimos anos, mas há questões diplomáticas sensíveis sempre que este tema é abordado pelo Relatório, como por exemplo em 2010 a publicação de tabelas no RDH de indicadores sobre liberdade política e direitos civis, que levou a uma queixa formal de Cuba e da Liga dos Países Árabes.

Uma crítica que também é feita desde a primeira publicação, (Hicks, 1997) (Foster, et al., 2005) e (Seth, 2009), é o facto de a desigualdade não estar reflectida no índice. No último relatório foram publicados três índices complementares⁹ de desigualdade com o objectivo de reflectir esta questão, no entanto, o IDH por si não responde a esta questão.

O segundo grupo de críticas prende-se com o facto da escolha das variáveis para representar as dimensões; em 2010 foram alterados alguns indicadores do IDH em resposta a estas críticas.

Na dimensão de “viver com saúde”, os críticos afirmam que a esperança de vida não dá informação sobre viver com saúde; em vez deste indicador, deveria ser usado um outro, como por exemplo a Esperança de Vida com Saúde da Organização Mundial de Saúde (OMS). O problema da utilização desta medida reside no facto de não estar previsto o seu cálculo regular pela OMS e desta forma não seria possível actualizar

⁸Nussbaum Capabilities: Life; Bodily Health; Bodily Integrity; Senses, Imagination, and Thought; Emotions; Practical Reason; Affiliation; Having the social bases of self-respect and non-humiliation; Other Species (environment); Play; Control over one's Environment.

⁹ IDH ajustado à desigualdade, Índice de Desigualdade de Género e Índice de Pobreza Multidimensional

anualmente o IDH. O relatório de 2010 manteve a Esperança de Vida, para esta dimensão.

O mesmo tipo de críticas é feita à forma como a educação é medida pois não dá informação sobre a qualidade do ensino. Idealmente dever-se-ia utilizar uma avaliação transversal do ensino, que medisse a qualidade das várias matérias nos diversos países, como se fez por exemplo para a matemática em 2007¹⁰. O problema seria a dificuldade em fazê-lo para todos os países e de uma forma contínua, por exemplo a avaliação da qualidade de ensino da matemática só foi feita para 36 países. No entanto, o IDH apresentou profundas alterações nesta dimensão em 2010, deixando de utilizar a taxa de analfabetismo na dimensão da educação, uma vez que esta medida já não diferencia os países, sendo superior a 90% em 102 países¹¹. Para medir a educação, o IDH em 2011 passou a adoptar a média de anos de escolaridade esperados, ou seja, o tempo de escolaridade que as crianças podem esperar ter na idade adulta se a taxa de participação escolar se mantiver. Para além desta medida, também é utilizado o número de anos escolares, usando o método de Barro (Barro & J.W.Lee, 2010) baseado em dados da UNESCO que podem ser actualizados anualmente e é um indicador bem estabelecido e relativamente bem aceite.

Ao longo dos últimos 20 anos muitos autores têm criticado a utilização do PIB, considerando uma medida inadequada para medir capacidades ou bem-estar, uma vez que não tem em conta os bens e serviços não transaccionáveis e a economia informal. Outra crítica que é feita a esta medida é não incluir as remessas dos emigrantes. O peso dos bens não transaccionáveis, como por exemplo os bens agrícolas e artesanais, a economia informal e as remessas dos emigrantes são elevados em muitas economias, especialmente nos países menos desenvolvidos. É certo que a utilização do PIB no IDH não pretende medir directamente uma capacidade, como acontece com a Saúde e Educação, mas sim medir o meio para atingir a capacidade de “ter um nível de vida decente”, muito diferente do conceito “quanto a economia é capaz de produzir”, desta forma o RN é mais adequado e passou a ser utilizado.

¹⁰ Trends in International Mathematics and Science Study (TIMSS) pelo International Association for the Evaluation of Educational Achievement (IEA),

¹¹ (UNDP, 2008)

2.4.1. Fórmula

Os autores (Desai, 1991), (Palazzi & Lauri, 1998), (Nathan, et al., 2008) e (Sagar & Najam, 1998) criticaram a fórmula do IDH por ser uma média aritmética. O facto de ser uma média aritmética implicava a substituição simples entre as dimensões. Isto contraria a intuição de que quanto pior é a privação numa dimensão particular, mais urgente são os esforços para melhorar os resultados dessa dimensão. Segundo a fórmula inicial, uma sociedade ao limite podia se centrar apenas numa dimensão e ter tão ou melhores resultados, do que uma que tentava melhorar as três dimensões.

Talvez a inovação mais radical da nova fórmula tenha sido a substituição da média aritmética pela média geométrica que é a única forma de permitir que a substituição entre as dimensões não seja linear e não é sensível ao tipo de variável (anos, rendimento), tornando-a mais sensível aos mínimos que são estabelecidos, uma vez que na formula antiga um país que se destacasse positivamente ou negativamente numa dimensão podia ser compensado de uma forma directa pelas outras dimensões, podendo assim esse resultado não se reflectir no IDH de uma forma muito acentuada. Com a nova fórmula se houver um grande diferencial entre a variável do país e o mínimo numa das dimensões, isso será reflectido com maior impacto no IDH através de uma média geométrica (Klugman, et al., 2011).

2.4.2. Peso das dimensões

O peso das dimensões tem sido uma das críticas mais frequentes ao longo dos anos, mas alguns autores demonstraram recentemente (Klasen, et al., 2011) fazendo uma análise de componentes principais aos dados de 1975 e 2005, que os pesos de cada dimensão deveriam ser iguais. De uma forma mais prática, (Haq, 1995) já argumentava que *"os pesos iguais se justificam pela simples premissa que todas as escolhas têm a mesma importância e não há uma razão a priori para dar mais peso a uma escolha do que a outra"*. Por este motivo o peso manteve-se.

2.4.3. Robustez e redundância

Uma das grandes preocupações, e que está relacionada com a crítica anterior, é verificar se o IDH é robusto a diferentes pesos e alterações dos indicadores subjacentes.

De alguma forma, a robustez é o oposto da relevância, dado que se houver uma correlação perfeita entre as variáveis, qualquer peso das dimensões levará ao mesmo resultado e o índice sendo robusto é irrelevante. Ao contrário, uma correlação baixa entre as variáveis implica que a escolha dos pesos seja muito importante na determinação do ranking dos países e o índice dificilmente será robusto aos pesos das dimensões.

Como se referiu no ponto 2.4.2 e como se pode ver na Figura 2-3, há uma forte correlação entre o Rendimento e o IDH, mas quando se compara os rankings do IDH com o PIB per capita, 79 dos 169 países analisados, apresentam diferenças superiores a 10 posições. Se comparamos as alterações nos rankings ao longo do tempo, o argumento da redundância desaparece.

Deve se ainda salientar que os críticos da redundância do IDH não estão a ter em conta a eficácia da comunicação de um índice agregado que alerta para outras dimensões para além do rendimento. Além disso, se tentarmos responder às questões mais frequentes sobre o desenvolvimento humano, como por exemplo “quem são os países mais desenvolvidos?” obteremos respostas diferentes se olharmos para cada um dos índices que compõem o IDH individualmente.

2.4.4. Classificação dos Países

Wolff demonstra que as categorias usadas para classificar os países em Alto, Médio e Baixo desenvolvimento não são robustas às alterações da fórmula e às actualizações dos dados (Wolff, et al., 2010). Segundo o seu estudo, 34% dos países do RDH estão mal classificados. O estudo também demonstra que muitos dados usados no IDH não estão revistos, conforme é demonstrado nas notas de rodapé das tabelas e quando são revistos têm valores completamente diferentes. Um exemplo que é dado é o caso da taxa de alfabetização de Portugal em 1975 influenciando o IDH. Em 2000 o IDH₁₉₇₅ era de 0,75 e em 2006 o mesmo índice era de 0,79. Outra crítica que também é feita refere-se às alterações da fórmula, sem alterar a forma como se classifica os países. O estudo demonstra, como se pode ver na figura 2-4 que quando se criou o IDH em 1990 a distribuição dos países era normal, mas com as alterações da fórmula e sem recalculer os valores que separam a classificação dos países, verifica-se que a

distribuição já não era normal e apresentava picos nos valores pré-estabelecidos para classificar os países.

Com o método de 2010, os países já não são agrupados de acordo com valores fixos, mas sim em quartis e a última questão já não se coloca, porque os países são classificados de acordo com a sua posição no ranking e não tendo em conta o valor obtido no IDH.

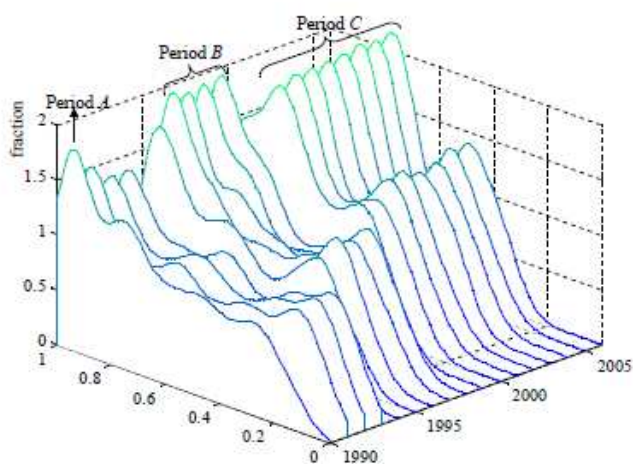


Figura 2-4- Distribuição do IDH publicado pelo RDH entre 1990 e 2005

Fonte: (Wolff, et al., 2010)

2.5. Impacto da nova fórmula de cálculo na classificação dos países

As alterações da fórmula de cálculo do IDH₂₀₁₀ produzem alterações substanciais no IDH e na classificação dos países como se pode ver na figura 2-5. A utilização da média geométrica implica valores mais baixos do IDH e com maiores alterações nos países com maior desigualdade nas dimensões que compõem o IDH. O impacto da alteração da classificação é moderado e sobretudo nos países com IDH médio.

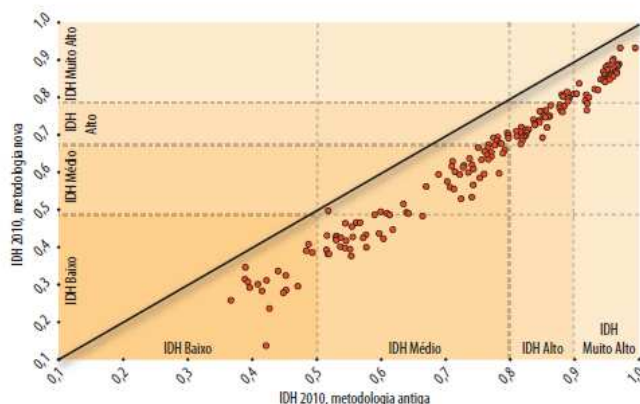


Figura 2-5 - IDH 2010: Metodologia nova vs metodologia anterior

Fonte: (UNDP, 2010)

2.6. Índices Complementares

Em 2010, dois Índices foram criados para dar resposta à necessidade de compreender o desenvolvimento humano na dimensão da Desigualdade de cada Dimensão do IDH (Índice de Desenvolvimento Humano Ajustado à Desigualdade) e de Género (Índice de Desigualdade de Género), os quais iremos analisar de seguida.

- **Índice de Desenvolvimento Humano Ajustado à Desigualdade**

Desde o seu lançamento que o RDH reconheceu que a Desigualdade era uma dimensão importante do Desenvolvimento Humano. Desde a primeira publicação que os dados dos países de várias instituições, entre elas o Banco Mundial, têm melhorado significativamente assim como os dados para comparar esta dimensão, sendo possível hoje criar um índice que reflecta esta dimensão. O IDHAD ajusta o IDH à desigualdade na distribuição de cada dimensão pela população. O IDHAD é calculado como uma média geométrica de médias geométricas, calculadas separadamente para cada dimensão em relação a toda a população. O IDHAD reflecte as desigualdades nas dimensões do IDH “descontando” o valor médio de cada dimensão segundo o seu nível de desigualdade. O IDHAD é igual ao IDH quando não existe desigualdade entre as pessoas, mas diminui em relação ao IDH à medida que a desigualdade cresce. Neste

sentido, o IDHAD é o nível real de desenvolvimento humano (levando em conta essa desigualdade), ao passo que o IDH pode ser visto como um índice de desenvolvimento humano “potencial” (ou o nível máximo de IDH) que pode ser alcançado se não existir desigualdade. A “perda” no desenvolvimento humano potencial devida à desigualdade é dada pela diferença entre o IDH e o IDHAD e pode ser expressa como uma percentagem (UNDP, 1990, pp. 11-12).

Exemplo de cálculo do IDHAD de 2011:

	Indicator	Dimension Index	Inequality measure (A1)	Inequality-adjusted Index
Life expectancy	78.8	0.930	0.043	$(1-0.043) \cdot 0.930 = 0.890$
Mean years of schooling	9	0.682		
Expected years of schooling	16.7	0.811		
Education index		0.782	0.040	$(1-0.040) \cdot 0.782 = 0.751$
Logarithm of GNI	10.16	0.780		
GNI	25,857	0.238	0.122	$(1-0.122) \cdot 0.238 = 0.209$

	Human Development Index	Inequality-adjusted Human Development Index	Percent loss
HDI with unlogged Income	$\sqrt[3]{0.930 \cdot 0.782 \cdot 0.238} = 0.557$	$\sqrt[3]{0.890 \cdot 0.751 \cdot 0.209} = 0.519$	$1 - 0.519 / 0.557 = 0.068$
HDI	$\sqrt[3]{0.930 \cdot 0.782 \cdot 0.780} = 0.828$	$(0.519 / 0.557) \cdot 0.828 = 0.772$	

Fonte: (UNDP, 2010, p. 216)

- **Índice de Desigualdade de Género**

O Índice de Desigualdade de Género (IDG) é a nova medida para reflectir a desvantagem das mulheres em três dimensões – capacitação, actividade económica e saúde reprodutiva. A disponibilidade de informação é uma limitação a este Índice, não estando disponível para 30 países.

O índice mostra a perda no desenvolvimento humano resultante da desigualdade entre as realizações femininas e masculinas nestas dimensões. Varia entre 0, que indica

que mulheres e homens desfrutam de igualdade, e 1, que indica que as mulheres sofrem uma desigualdade tão grande quanto possível em todas as dimensões medidas. É calculado usando uma medida de desigualdade sensível à associação sugerida por (Seth, 2009). O índice baseia-se na média geral das médias gerais de ordens diferentes – a primeira agregação é pela média geométrica entre dimensões; estas médias, calculadas separadamente para mulheres e homens, são depois agregadas usando uma média harmónica entre géneros.

Exemplo de cálculo do IGD de 2010

	Reproductive health		Empowerment		Labour market
	Maternal mortality ratio	Adolescent fertility rate	Parliamentary representation	Attainment at secondary and higher education	Labour market participation rate
Female	110	75.6	0.094	0.488	0.640
Male	na	na	0.906	0.463	0.852
(F+M)/2	$(\sqrt{(1/110) - (1/75.6) + 1})/2 = 0.50$		$(\sqrt{0.094 \cdot 0.488 + \sqrt{0.906 \cdot 0.463}})/2 = 0.431$		$(0.640 + 0.852) / 2 = 0.746$

(UNDP, 2010, p. 230)

$$G_F = 0,115 = \sqrt[3]{\sqrt{\left(\frac{1}{110} \frac{1}{75,6}\right)} \times \sqrt{0,094 \times 0,488} \times 0,64}$$

$$G_M = 0,82 = \sqrt[3]{1 \times \sqrt{0,906 \times 0,463} \times 0,852}$$

$$Harm(G_F, G_M) = 0,201 = \left[\frac{1}{2} \left(\frac{1}{0,115} + \frac{1}{0,82}\right)\right]^{-1}$$

$$G_{\overline{FM}} = 0,546 = \sqrt[3]{0,505 \times 0,431 \times 0,746}$$

$$\text{Índice de Desigualdade de Género} = 1 - 0,201/0,546 = 0.632$$

3. Algoritmo SOM

3.1. Descrição Sumária do Algoritmo SOM

Nos últimos 20 anos, fruto dos avanços tecnológicos, novas técnicas de extracção de informação a partir dos dados têm sido desenvolvidas. As Redes Neurais Artificiais (RNA) são uma aproximação quantitativa da análise de grandes volumes de dados, inspiradas na forma como se pensa que o cérebro processa a informação, um processo dinâmico cuja propriedade mais importante é a própria capacidade de aprender durante o processo. (Fayyad, et al., 1996)

As RNA são compostas por múltiplos métodos, entre eles o Self Organizing Map (SOM), um método não supervisionado e que por isso não tem nenhum agente externo que lhe indique a resposta desejada.

O SOM foi proposto por Tuevo Kohonen na década de 80 (Kohonen, 1982). Desde essa altura várias trabalhos e livros têm sido publicados sobre este tema, mas o primeiro livro continua a ser a referência do assunto.

Segundo Kohonen *“The SOM is a new, effective software tool for the visualization of high-dimensional data. It converts complex, nonlinear statistical relationships between high-dimensional data items into simple geometric relationships on a low-dimensional display. As it thereby compresses information while preserving the most important topological and metric relationships of the primary data items on the display, it may also be thought to produce some kind of abstractions”* (Kohonen, 1982).

O algoritmo pode ser consultado no Software gratuito SOM_PAK (Kohonen, et al., 1995), mas de uma forma muito resumida a sua essência é: reduzir a dimensionalidade dos dados para um mapa ou grelha de neurónios com um menor número de dimensões.

O algoritmo de treino do SOM envolve essencialmente dois passos como está representado na Figura 3-1 que representa o processo de aprendizagem da rede e a sua projecção.

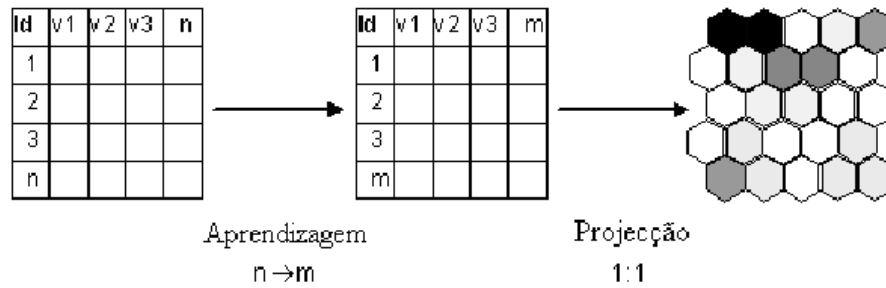


Figura 3-1– Ilustração dos princípios do SOM

Fonte: (Henriques & Bação, 2004)

Como se pode ver na Figura 3-2, o objectivo do SOM é agrupar grupos similares de vectores ou neurónios que estão perto uns dos outros no espaço *input* (x_i), em unidades próximas no mapa ou grelha.

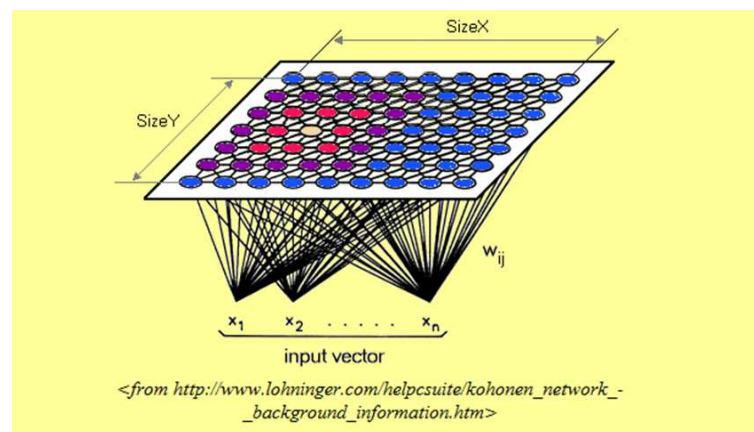


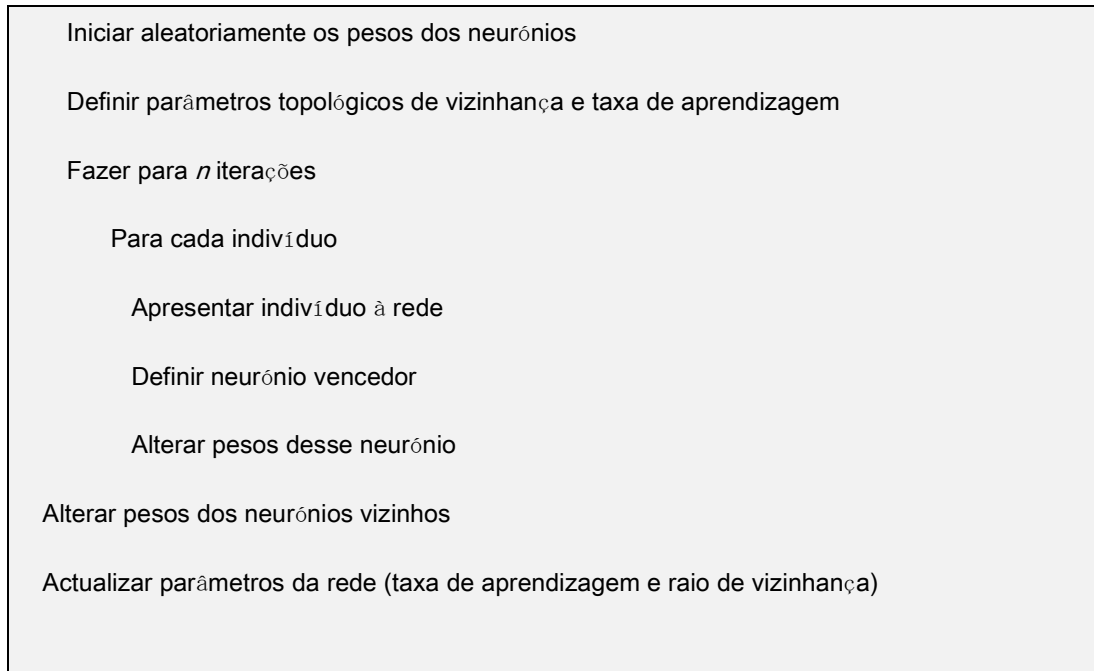
Figura 3-2 - Exemplo de uma grelha SOM

Fonte: (Ahn & Syn, 2005)

Assim, um padrão *input* é apresentado à rede e as unidades da camada de *output* competem entre elas pelo direito de se declararem vencedoras. O número de unidades de processamento na camada *output* é determinado pelo utilizador, baseado na forma e tamanho que se pretende do mapa. A unidade *output* vencedora (Best Matching Unit ;BMU) será aquela cujos pesos nas conexões de entrada são os mais próximos do padrão de *input* em termos de distância euclidiana. O peso da conexão do vencedor é

então ajustado em direcção ao padrão de *input* por um factor denominado por taxa de aprendizagem.

De seguida é descrito de forma sucinta o algoritmo do SOM:



Fonte: (Henriques & Bação, 2004)

A grande vantagem da utilização do algoritmo SOM é que este cria um mapa topológico ajustando não só as ponderações dos vencedores, mas igualmente as ponderações dos outputs vizinhos. Como resultado do treino da rede, o tamanho da vizinhança à volta das unidades vencedoras vai diminuindo ao longo do tempo. Durante o processo de treino, o número de unidades de output que sofre alterações vai diminuindo. Por fim só a unidade vencedora sofre alterações. A taxa de aprendizagem vai igualmente diminuindo durante o processo.

No Youtube há vários filmes que explicam de uma forma interactiva este processo, bastando fazer uma busca por “Self Organizing Map”, mas o filme do seguinte endereço parece explicar bem este processo de interacção no algoritmo (Anonimo, 2010) ¹².

¹² <http://www.youtube.com/watch?v=-6a7LATC-9g&feature=related>

3.2. Vantagens da Utilização do SOM na visualização de clusters

O objectivo central da Análise Exploratória de Dados ou *Datamining* é descobrir padrões e relações em grandes quantidades de dados, apresentando-os de uma forma compreensiva e simultaneamente preservando-os na forma original ao máximo.

Um dos métodos de Análise Exploratória de Dados é o *Clustering*, que consiste em agrupar uma população heterogénea em subgrupos (*clusters*) homogéneos. A diferença entre o *Clustering* e a Classificação, é que ao contrário da Classificação, o *Clustering* não depende de classes pré-definidas, os registos são agrupados apenas com base nas suas semelhanças (Kohonen & S.Kaski, 1996).

Há inúmeros algoritmos para segmentar registos, normalmente de distância. A aplicação destes algoritmos requer não só um bom conhecimento destes, como das relações existentes entre os dados, de forma que a sua aplicação seja eficaz e que se consiga identificar esses *clusters*.

O SOM por ser uma rede neuronal não-supervisionada, é um método único de visualização de *clusters*, pois permite não só identificar estes no espaço de output, como também preserva as relações não lineares entre eles e mapeia os dados perto dos vizinhos. Mesmo que não exista uma relação explícita de *clustering* nos dados, as sombras do mapa de output permitem construir esses clusters, identificando-os como se fossem mapas topográficos, onde a maior concentração de dados são as “montanhas” e os vales, o que os pode separar, como se pode ver na Figura 3-3.

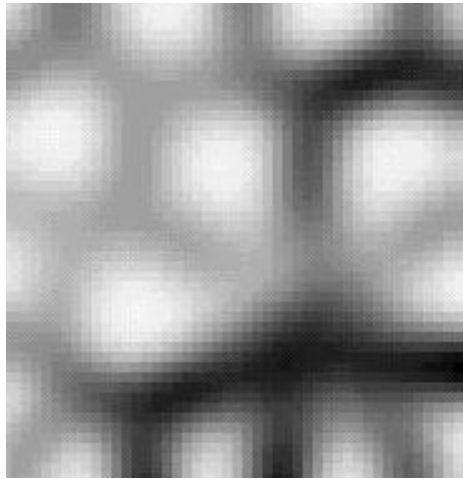


Figura 3-3– Exemplo de um SOM

Observação: Podemos ver na Figura 3-3 que se a distância média é alta, os pesos ao redor são muito diferentes e uma cor escura é atribuída. Pelo contrário, se a distância média é baixa, uma cor mais clara é atribuída. O mapa da Figura 3-3 mostra clusters nas sombras brancas e as linhas pretas representam as distâncias.

Nesta fase a experiência e conhecimento do problema por parte do utilizador é importante para a qualidade da interpretação da projecção e construção dos clusters.

3.3. Estudo que aplicou o SOM ao Bem-estar e Pobreza no Mundo

Kohonen e S. Kaski publicaram em 1996 um estudo que pretendia ser uma receita de como se podem usar variáveis macro económicas e micro económicas através do SOM para descrever fenómenos complexos, como o bem-estar dos países (Kohonen & S.Kaski, 1996). No estudo mencionado, são utilizadas 39 variáveis do Relatório do Banco Mundial e que na sua maioria existem no RDH. Estas variáveis foram escolhidas para reflectir aspectos como consumo, saúde, educação e assistência social. O PIB foi deixado de fora, uma vez ter uma forte correlação com as restantes variáveis. O mapa do SOM utilizado tinha 13 colunas por 9 linhas e duas dimensões.

Os *missing values* foram corrigidos pelo SOM, mas países com mais de 12 variáveis em falta (30% das variáveis) não foram tidos em conta.

Os países foram ordenados pelo SOM como se pode ver na Figura 3-4. A ordem dos países reflecte a organização geográfica dos países e serem ou não membros da OCDE.

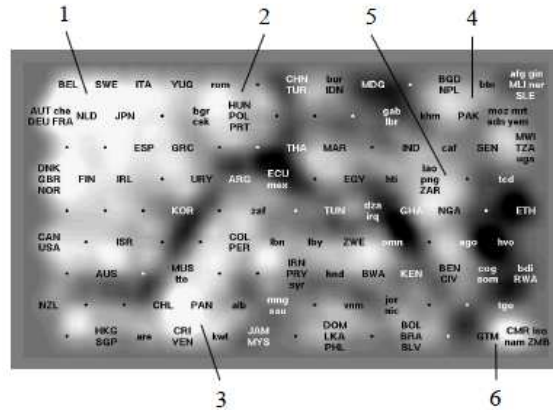


Figura 3-4 - Utilização da U-matrix para representar os países segundo o bem-estar (Kohonen & S.Kaski, 1996)

Nota: A ordem das abreviaturas dos países, indica a semelhança do nível de vida. As áreas mais claras significam áreas com um nível alto de clustering e as áreas mais escuras a separação. Os países do Cluster 1 correspondem aos países Muito Desenvolvidos, os países da América do Sul encontram-se no Cluster 3, que estão claramente separados do Cluster 1 por um vale escuro. Os países Asiáticos estão no Cluster 4, próximos dos Mais Ricos, mas separados por um vale. Os países Africanos encontram-se no Cluster 5 e 6.

O estudo também permite visualizar como as variáveis contribuem para o cluster final (Figura 3-5).

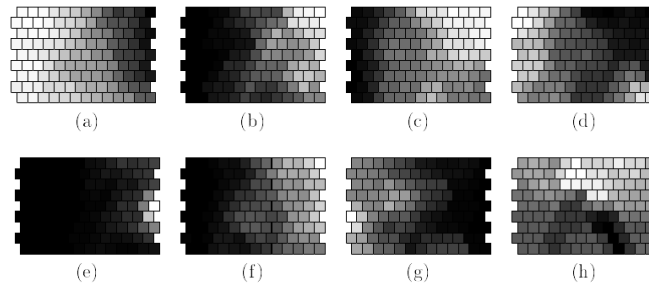


Figura 3-5 – Distribuição dos indicadores dos países do SOM nas variáveis que constroem o algoritmo

Nota: Podemos ver a distribuição nos indicadores que compõem os *clusters* da seguinte forma: A Esperança de Vida (a) diminui para os *clusters* da direita (claro é o maior valor, escuro o menor). As variáveis são: Taxa de Iliteracia (b); % de gastos em comida no orçamento familiar (c); % de despesas em saúde no consumo das famílias (d); População por Agregado (e); Taxa de Mortalidade Infantil (f); Taxa de Ensino Secundário (g)

3.4. Utilização do GeoSOM Suit

Existem vários softwares disponíveis para aplicar o SOM, nomeadamente o SAS, o GeoSOM Suit e o SOM_PAK, sendo estes dois últimos gratuitos.

Optou-se pelo GeoSOM Suit por permitir visualizar de uma forma dinâmica e com ligações em MatLab de várias janelas que apresentam diferentes aspectos importantes na construção dos clusters. Esta ferramenta também permite de uma forma manual construir os *clusters*, como se pode ver na Figura 3-6.

Mais informações sobre este software podem ser encontradas no website respectivo. (ISEGI, 2009)

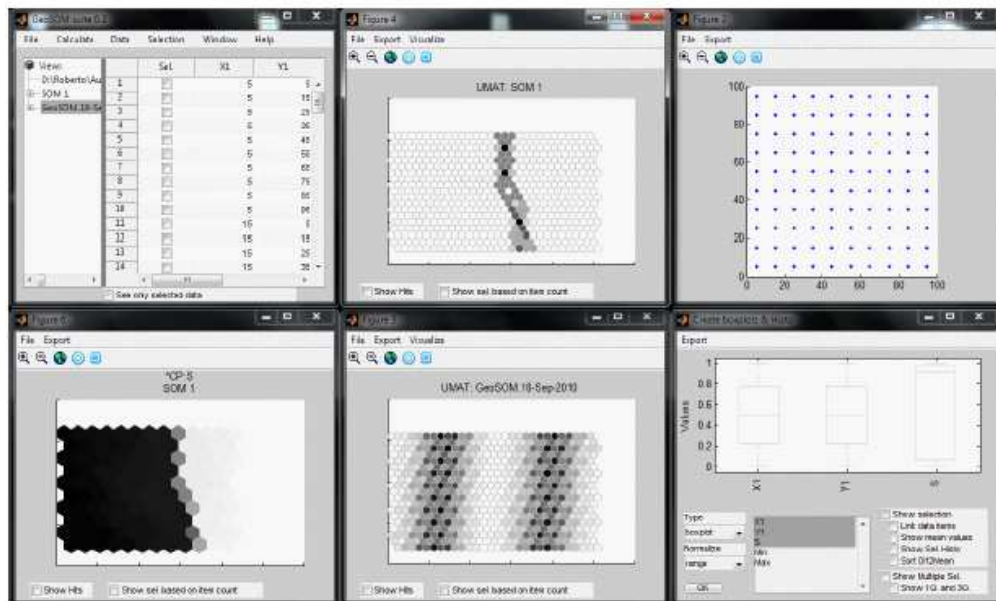


Figura 3-6 – Exemplo da visualização das várias janelas dinâmicas do GeoSOM Suit que permitem de uma forma directa e fácil visualizar vários aspectos importantes da segmentação.

4. Metodologia

De forma a atingir os objectivos propostos, foram seleccionados vários indicadores de várias dimensões do Desenvolvimento Humano, disponíveis nas tabelas dos Anexos do Relatório de Desenvolvimento Humano e que estão em formato digital no website do PNUD.

Seguiu-se os passos de uma Análise Exploratória, nomeadamente: (i) formulação do problema; (ii) compreensão dos dados; (iii) preparação dos dados; (iv) aplicação do modelo; (v) avaliação do modelo; (vi) apresentação dos resultados.

Nos pontos seguintes é apresentado em pormenor cada umas destas fases, excepto a formulação do problema, que foi apresentada no Capítulo 1 e a apresentação dos resultados, que será apresentada no capítulo seguinte.

4.1. Compreensão dos Dados

Nesta fase o objectivo é compreender o significado das variáveis e a forma como estas se relacionam. Foi muito importante nesta fase a utilização de uma ferramenta criada pelo Google (Google Public Data Explorer), disponível no site do PNUD e onde é possível visualizar de uma forma rápida e fácil as variáveis mais importantes do RDH das seguintes dimensões: Saúde, Rendimento, Desigualdade, Pobreza, Sustentabilidade, Educação e Género. Um exemplo é mostrado na Figura 4-1.

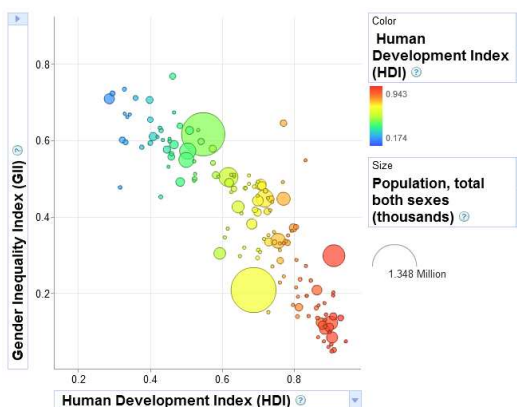


Figura 4-1 – Cruzamento dos países de acordo com o valor do IDH (eixo dos x) e do Nível de Desigualdade de Género

O conceito de Desenvolvimento Humano que esteve na base da selecção das variáveis para o presente trabalho foi o do criador do Relatório Humano do Desenvolvimento, segundo o qual o desenvolvimento humano deve ter quatro características essenciais: Equidade, Sustentabilidade, Produtividade e Capacitação (Haq, 1995).

O Capítulo IV do Relatório Humano, foi tido como referência nesta selecção, de forma a analisar várias dimensões do Desenvolvimento Humano, utilizando os indicadores disponibilizados nas tabelas (UNDP, 2010, p. 65). Foi também utilizado um Relatório do PNU que analisa a Desigualdade de Género (Gaye, et al., 2010).

Assim as variáveis escolhidas foram todas variáveis usadas pelo IDH para as três dimensões básicas (“Viver com Saúde”, “Ter Acesso a uma vida digna” e “Educação”). Seleccionaram-se todos os indicadores disponíveis no RDH para as seguintes dimensões de Género: “Desigualdade”, “*Empowerment*”, “Sustentabilidade” e “Igualdade de Género” (UNDP, 2010). Usou-se também o Coeficiente de Gini, o indicador usado pelo Banco Mundial para medir a Desigualdade de Rendimento. Para além destes indicadores, também se seleccionou a Taxa de Emprego, por ser um importante indicador de segurança, o Crescimento do Acesso à Internet, por ser um importante indicador de *Empowerment* como se viu na recente primavera Árabe e a Percentagem da População Urbana.

4.2. Preparação dos Dados

Para as dimensões Desigualdade de Género, *Empowerment* e Sustentabilidade, as variáveis foram escolhidas por terem menos de 25 países omissos (15% do total) e por existirem em intervalos temporais de um ano.

Assim, não se irá usar o IDHA e as suas componentes. Na dimensão do *Empowerment* não se utilizará o Valor de Vitimas de Corrupção. Na dimensão da Sustentabilidade não se utilizará a Percentagem da Poupança Ajustada no PIB, a Pegada Ecológica e a Percentagem de Energia de Origem Fóssil e Renovável. Nesta dimensão

também não se usou o Número de Mortes devido a Desastres Naturais por se encontrarem desactualizados (valores de 2004).

De forma a melhorar a compreensão do modelo foram criadas algumas variáveis, sobretudo na Dimensão de Género, de forma a obter Rácios de Desigualdade e serão esses os valores usados no modelo e não as variáveis originais das tabelas. De uma forma resumida as variáveis seleccionadas encontram-se na Figura 4.2, bem como os códigos usados. O Anexo B contém a definição destas variáveis.

Dimensão	Indicador	Código
HDI	Human Development Index (HDI) value	IDH
	Rank	Rank
	Life expectancy at birth (years)	LifExp
	Mean years of schooling (of adults) (years)	MYSchool
	Expected years of schooling (of children under 7) (years)	ExpYSchool
	GNI per capita in PPP terms (constant 2005 international \$)	GNI
	Non Yield IDH	IDHNR
Desigualdade	Income Gini coefficient	Gini
	Adolescent fertility rate (births per 1,000 women aged 15-19)	AdoFert
	Shares in parliament, female-male ratio	SeatParl
	Population with at least secondary education (female/male ratio)	EdGR
	Labour force participation rate (female-male ratio)	LFR
	Births attended by skilled health personnel	BirthsAten
Empowerment	Political freedom, democracy	Democ
	Human rights violations	HRVio
	Press freedom	PressFre
	Journalists in prisoned	JourImpr
	Political engagement	PoliEng
Sustentabilidade	Carbon Dioxide Emissions per capita (tonnes)	CO2
	Protected area	ProtAre
	Population Without Access to Improved Services of Water	Water
Outras	Employment to population ratio	Empl
	Internet Users Growth	Inter
	Urban Population	Urban

Figura 4-2– Tabela dos Indicadores Usados na Análise por cada Dimensão

Os países com mais de 5 Indicadores em falta não foram considerados, reduzindo-se assim o número de países de 169 para 144. Na Figura 18 apresentam-se os países não considerados agrupados de acordo com a classificação do IDH. A maior parte dos países não considerados estão classificados pelo IDH como de Desenvolvimento Médio (11 países não seleccionados num total de 42).

Desenvolvimento Humano	Número de Países Excluídos	Nome dos Países
Muito Alto	5	Andorra; Barbados; Brunei; Darussalam; Hong Kong; Liechtenstein
Alto	5	Mauritius; Montenegro; Serbia; Tonga; Bahamas
Médio	11	Cape Verde; Equatorial Guinea; Fiji; Maldives; Micronesia; Sao Tome and Principe; Solomon Islands; Suriname; Swaziland; East Timor; Turkmenistan
Baixo	4	Comoros; Djibouti; Guinea-Bissau; Lesotho

Figura 4-3 – Países Não Incluídos na Análise

Alguns indicadores apresentam valores omissos, nesse caso e como o software não aceita valores omissos, corrigiu-se aplicando a média do valor para a Classificação do País do IDH.

4.3. Aplicação do Modelo

- Análise Descritiva

Antes de aplicar o modelo foi necessário fazer uma análise descritiva dos Dados em SAS. Na primeira análise dos indicadores, podemos verificar que os indicadores que têm maior desvio padrão são: Crescimento da Internet, Rendimento Nacional, Número de Nascimentos com Assistência no Parto e Liberdade de Imprensa (Figura 4-4).

Variable	Mean	Variance	Minimum	Maximum	Range
IDH	0.6386935	0.0359649	0.1400792	0.9375610	0.7974818
Rank	84.2013889	2487.95	1.0000000	169.0000000	168.0000000
LifExp	69.0645347	106.2349974	44.6370000	83.1660000	38.5290000
MYSchool	7.6017424	9.1312287	1.2067000	12.6313000	11.4246000
ExpYSchool	12.2205994	11.1717764	4.3059953	20.4573853	16.1513900
GNI	13520.70	227439087	176.1652738	79426.35	79250.18
IDHNR	0.6831116	0.0326726	0.2854627	0.9890680	0.7036053
Gini	40.2227748	82.8932726	16.8000000	74.3000000	57.5000000
AdoFert	53.9320833	2127.32	3.1600000	201.4140000	198.2540000
SeatParl	18.0754051	106.9953270	0	50.9400000	50.9400000
EdGR	0.8218612	0.0598575	0.1706886	1.5526696	1.3819809
LFR	0.7346246	0.0357872	0.2514418	1.0362401	0.7847983
BirthsAten	79.6997863	675.7931327	6.0000000	100.0000000	94.0000000
Freedom	66.3964184	250.8182846	24.0000000	96.0000000	72.0000000
Democ	1.4583333	0.5576923	0	2.0000000	2.0000000
HRVio	2.6121907	1.1696884	1.0000000	5.0000000	4.0000000
PressFre	27.6933559	543.7307477	0	104.1400000	104.1400000
JourImpr	0.6527778	8.7876845	0	24.0000000	24.0000000
PoliEng	20.1498644	63.0478395	5.0000000	42.0000000	37.0000000
CO2	5.2667229	56.9138628	0.0242000	56.2425000	56.2183000
ProtAre	12.3677083	91.8975339	0.1100000	53.7500000	53.6400000
Water	14.0325292	291.7536037	0	62.0000000	62.0000000
Empl	58.6055556	117.9429060	34.8000000	84.2000000	49.4000000
Inter	2481.12	31752796.55	64.2963433	46566.67	46502.37
Urban	0.5793194	0.0507755	0.1100000	1.0000000	0.8900000

Figura 4-4 – Tabela Descritiva Sumária dos Dados

Pela Matriz de Correlação da Figura 4-5, é possível observar que a correlação dos indicadores adicionados não é muito alta. Os valores que apresentam maior correlação são os que compõem o Índice de Desenvolvimento Humano. Os indicadores que se destacam pela baixa correlação com as restantes variáveis são o Coeficiente de Gini, Lugares no Parlamento Femininos, todos os indicadores da dimensão de *Empowerment*, as Emissões de CO₂ e o Número de Hectares de Área Protegida. O Emprego e a Taxa de Crescimento do Acesso à Internet também apresentaram valores baixos.

	IDH	LifExp	MYSchool	ExpYSchool	GNI	IDHNR	Gini	AdoFert	SeatParl	EdGR	LFR	BirthsAten	Freedom	Democ	HRVio	PressFre	JourImpr	PoliEng	CO2	ProtAre	Water	Empl	Inter	Urban	
IDH	1.00																								
LifExp	0.92	1.00																							
MYSchool	0.89	0.76	1.00																						
ExpYSchool	0.90	0.80	0.85	1.00																					
GNI	0.76	0.66	0.60	0.65	1.00																				
IDHNR	0.97	0.92	0.94	0.93	0.68	1.00																			
Gini	-0.38	-0.37	-0.37	-0.36	-0.39	-0.38	1.00																		
AdoFert	-0.77	-0.76	-0.71	-0.70	-0.58	-0.78	0.50	1.00																	
SeatParl	0.19	0.13	0.19	0.24	0.22	0.19	-0.10	-0.14	1.00																
EdGR	0.73	0.66	0.71	0.69	0.48	0.74	-0.09	-0.58	0.14	1.00															
LFR	0.01	-0.10	0.17	0.12	0.05	0.06	-0.16	-0.02	0.37	0.06	1.00														
BirthsAten	0.84	0.75	0.82	0.80	0.57	0.85	-0.26	-0.67	0.09	0.75	0.00	1.00													
Freedom	0.33	0.29	0.20	0.28	0.44	0.27	0.03	-0.15	0.31	0.19	0.05	0.20	1.00												
Democ	0.31	0.29	0.38	0.37	0.12	0.37	-0.07	-0.06	0.17	0.15	0.24	0.21	0.04	1.00											
HRVio	-0.51	-0.47	-0.47	-0.49	-0.48	-0.51	0.30	0.41	-0.30	-0.28	-0.35	-0.38	-0.26	-0.30	1.00										
PressFre	-0.39	-0.30	-0.44	-0.39	-0.41	-0.40	0.10	0.10	-0.26	-0.18	-0.36	-0.30	-0.20	-0.57	0.57	1.00									
JourImpr	-0.02	0.01	-0.06	-0.03	-0.09	-0.02	-0.05	-0.13	-0.11	-0.05	-0.06	0.04	-0.06	-0.32	0.23	0.49	1.00								
PoliEng	0.04	-0.05	0.02	-0.02	0.23	-0.01	0.01	0.08	0.20	-0.10	0.24	-0.07	0.41	0.05	-0.16	-0.20	-0.08	1.00							
CO2	0.56	0.47	0.43	0.45	0.83	0.48	-0.29	-0.48	0.02	0.41	-0.09	0.48	0.30	-0.12	-0.31	-0.22	-0.03	0.08	1.00						
ProtAre	0.04	0.04	0.07	0.03	0.01	0.05	0.20	0.16	0.12	0.07	0.12	0.05	0.22	0.25	-0.05	-0.15	-0.05	0.16	-0.06	1.00					
Water	-0.84	-0.82	-0.76	-0.80	-0.58	-0.85	0.26	0.70	-0.12	-0.70	0.09	-0.82	-0.25	-0.25	0.38	0.31	0.01	0.13	-0.47	-0.08	1.00				
Empl	-0.43	-0.39	-0.42	-0.36	-0.12	-0.42	0.14	0.40	0.14	-0.28	0.41	-0.47	0.12	-0.15	0.09	0.11	0.04	0.20	-0.04	0.14	0.46	1.00			
Inter	-0.26	-0.25	-0.21	-0.33	-0.25	-0.27	0.00	0.13	-0.13	-0.10	-0.15	-0.20	-0.18	-0.25	0.25	0.28	0.05	0.13	-0.19	-0.14	0.30	-0.05	1.00		
Urban	0.79	0.68	0.65	0.72	0.71	0.74	-0.15	-0.50	0.12	0.62	-0.10	0.72	0.29	0.15	-0.37	-0.36	-0.07	0.10	0.58	0.08	-0.69	-0.35	-0.23	1.00	

Figura 4-5 – Matriz de Correlação das Variáveis

- Definição dos parâmetros da rede SOM

Neste passo vamos definir os parâmetros para o treino da rede SOM. Os parâmetros que devemos fornecer são introduzidos num formulário, como exemplificado na figura seguinte.

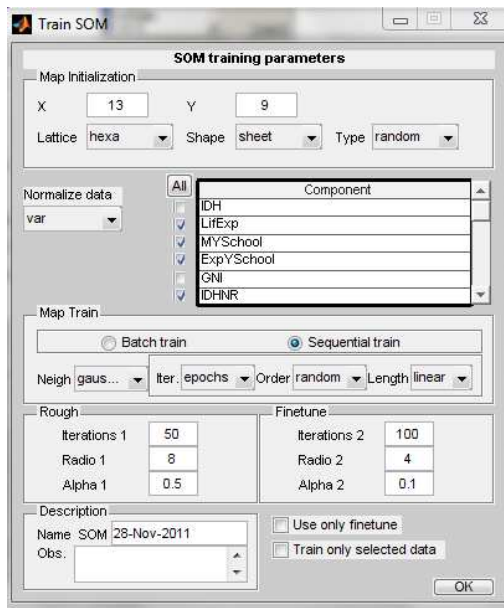


Figura 4-6 – Formulário de introdução dos parâmetros no GeoSOM Suit

De seguida ir-se-á justificar os critérios seleccionados no GeoSOM Suit.

Tamanho e Dimensão do Mapa

A dimensão da rede escolhida é de 13 por 9, a qual é igual à dimensão escolhida no estudo mencionado no (Kohonen & S.Kaski, 1996) que serviu de receita para o presente trabalho, uma vez que o problema é idêntico apesar do número de variáveis ser diferente (o estudo de Kohonen usa 39 variáveis e o presente trabalho 23).

O número de unidades deverá ser igual ao número esperado de clusters e cada cluster representa uma unidade única. Fez-se esta experiência, com uma rede 2 por 3, mas o erro era muito superior ao que se escolheu no final ($Q=3.72$). No caso do SOM emergente é utilizado um número muito grande de unidades de forma a obter uma Matriz com os clusters bem identificados. Fez-se uma experiência com uma rede de 30

por 10, mas obtiveram-se muitos clusters e num cenário como este seria difícil de definir políticas para cada um. Os dois exemplos são apresentados no Anexo C.

Topologia

O SOM usa tipicamente dois tipos de topologia, ou seja a forma como são contactados os vizinhos: rectangular ou hexagonal, como se pode ver na figura:

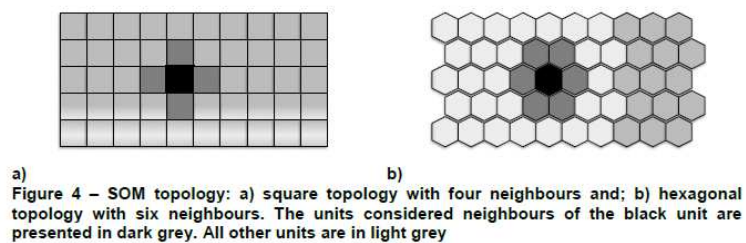


Figura 4-7– Tipo de Topologia do SOM

Fonte: (Henriques, 2010)

No primeiro exemplo, as unidades são ligadas a 4 vizinhos e no segundo a 6 (excepto as unidades que se encontram na fronteira). Normalmente a topologia hexagonal é preferida e por isso será a escolhida.

O SOM pode também utilizar diferentes formatos que estão disponíveis no GeoSOM Suit e que estão exemplificadas na figura 4-8. A forma escolhida foi a sheet, por ser a melhor na representação da distribuição dos dados.

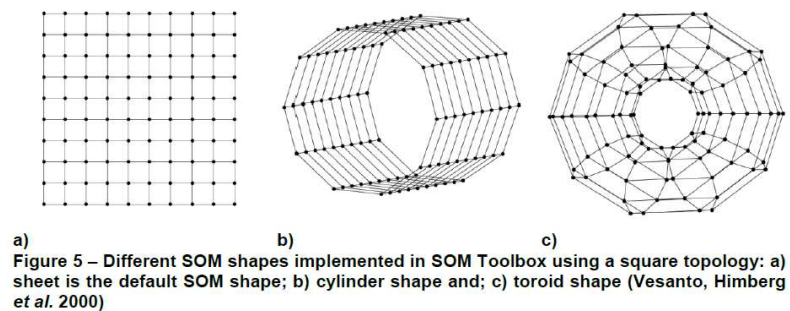


Figura 4-8 – Diferentes tipos de formato do SOM

Fonte: (Henriques, 2010)

Finalmente também foi necessário definir como cada vector será inicializado. Apesar do SOM ser robusto à iniciação, esta pode ser aleatória (*random*) ou linear. Foi escolhido o processo aleatório, que conforme o nome indica, os pesos das unidades são escolhidos e desenhados de uma forma aleatória.

Normalização dos Dados e escolha das variáveis

Como os dados têm diferentes escalas, estas foram normalizados. No Primeiro Treino escolheram-se todas as variáveis excepto o IDH, GDI e o IDH sem Rendimento, por terem correlações altas com os indicadores que compõem o índice.

Mapa de Treino

Nesta fase decide-se como será o processo de treino, isto é, como se vão movimentar os neurónios em função dos dados introduzidos e à actualização da taxa de aprendizagem (*Alpha*) e Grau de Vizinhança (*Radio*).

Para isso deve-se proceder à variação dos parâmetros de treino de forma a compreender a reacção da rede a diferentes valores e conseqüentemente que valores fornecem resultados mais válidos. Desta forma usaram-se Taxas de Aprendizagem de 0,3, 0,5 e 0,8.

O GeoSOM Suit permite fazer um treino sequencial em duas épocas. Uma época é o conjunto de interações, onde todos os padrões de treino são apresentados à rede. Esta apresentação dos padrões de entrada pode ser aleatória ou seguir a ordem do conjunto de dados, a forma escolhida foi a aleatória.

Não havendo um critério que se possa seguir, fizeram-se várias experiências alterando a Taxa de Aprendizagem da Primeira Época, mas mantendo sempre o Grau de Vizinhança da Primeira Época em 8. O número de interacções escolhido foi de 50, por se considerar suficiente para as taxas de aprendizagem de 0,5 e 0,8. No caso da Taxa de Aprendizagem de 0,3, aumentou-se este valor para 100. Na segunda época os dados

foram adaptados sempre de forma a terem um Grau de Vizinhança e uma Taxa de Aprendizagem menores, mas o dobro das interacções.

4.4. Avaliação do modelo

Com base nos resultados avaliaram-se o Erro Q (pretende-se que seja o mais próximo possível de 0), os quais estão sintetizados no quadro seguinte e as respectivas matrizes U .

	Taxa de Aprendizagem 0,3		Taxa de Aprendizagem 0,5		Taxa de Aprendizagem 0,8	
	Treino 1	Treino 2	Treino 1	Treino 2	Treino 1	Treino 2
Interacções	100	200	50	100	50	100
Grau de Vizinhança	8	4	8	4	8	4
Tx de Aprendizagem	0,3	0,1	0,5	0,1	0,8	0,1
Erro Q	0,5329		0,5470		0,5536	

Figura 4-9 – Resumo dos Modelos Feitos para diferentes taxas de aprendizagem

Os Erros obtidos não são muito diferentes e por isso teve-se em conta a Matriz U, sendo a mais intuitiva para a explicação do problema.

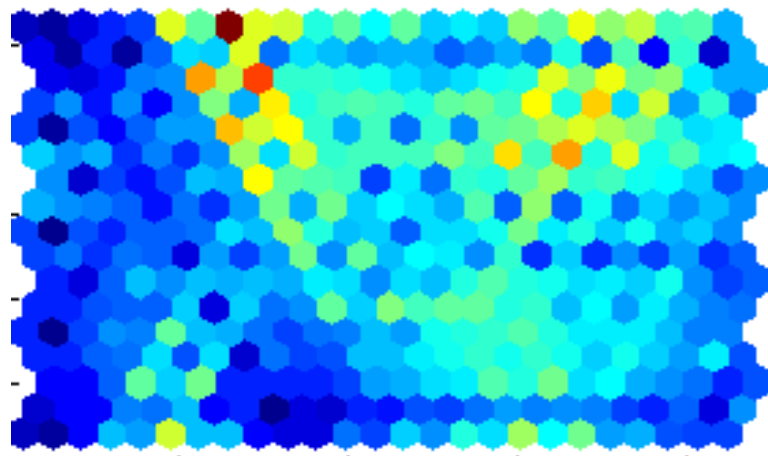


Figura 4-10 – Mapa UMAP seleccionado

Podemos ver pela Figura 4-10, que se a distância média é baixa, uma cor azul é atribuída, enquanto que pelo contrário, se os pesos ao redor são muito diferentes uma cor amarela é atribuída. Assim as manchas azuis representam os clusters e as linhas amarelas os vales que as separam.

5. Resultados

Com base na Matriz U (Figura 5-1) é possível identificar clusters através do padrão de cores. A cor azul é atribuída aos valores próximos por oposição à cor amarela e desta forma as manchas azuis representam os clusters e as linhas amarelas os vales que as separam. No software é ainda possível visualizar se cada hexágono tem muitos países. Quanto maior for o hexágono cor-de-rosa, mais países contém o hexágono.

Apenas analisando a Matriz U e sem mais nenhuma análise, é possível sugerir uma segmentação que está representada na figura 5-1 por linhas azuis escuras, separando o mapa em 6 clusters.

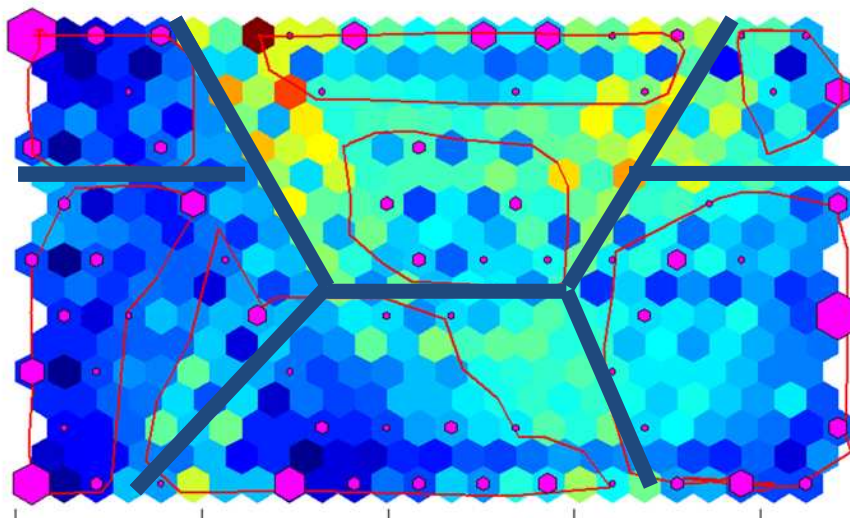


Figura 5-1 - Clusters Criados no UMAP (encarnado) e sugeridos (azul escuro)

No GeoSOM Suit é possível analisar os Planos das Componentes (CPs) dos indicadores usados e não usados. Com base nesta análise foi possível ter uma caracterização genérica dos dados e ajustar os segmentos, de forma a obter segmentos mais robustos no ponto de vista intuitivo. Na interpretação dos CPs, o azul-escuro corresponde aos valores mais baixos para a variável e o encarnado corresponde aos valores mais altos.

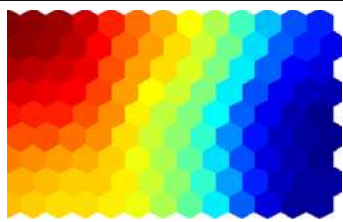


Figura 5-13 - CP Índice Desenvolvimento Humano

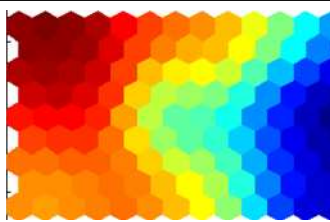


Figura 5-12 - CP Esperança de Vida

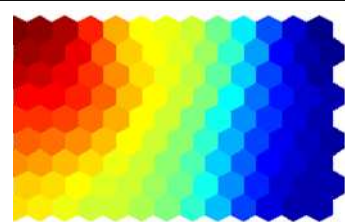


Figura 5-11 - CP Média de Anos Escolares Esperada

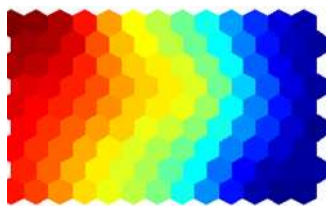


Figura 5-10 - CP Média de Anos de Ensino

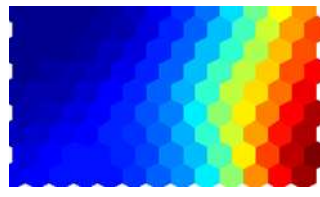


Figura 5-9 - CP Pessoas sem acesso a serviços de Água

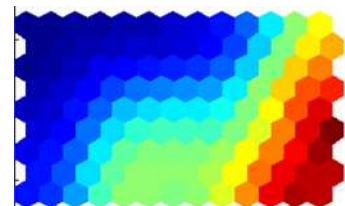


Figura 5-8 - CP Taxa de Fertilidade Adolescente

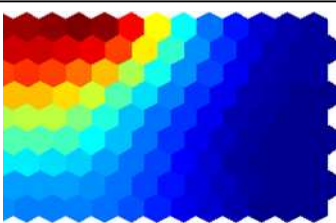


Figura 5-7 - CP Rendimento (IDH)

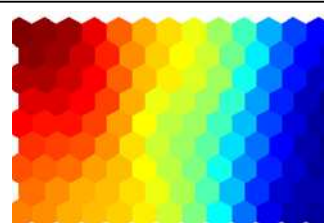


Figura 5-6 - CP IDH sem Rendimento

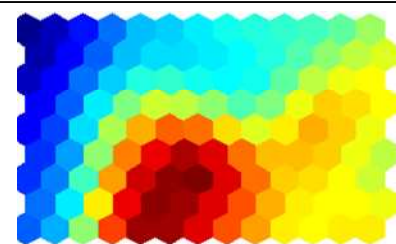


Figura 5-5 - CP Coeficiente Gini

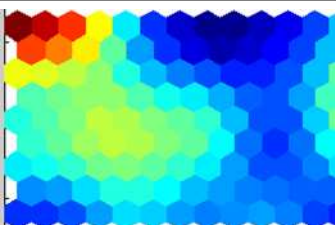


Figura 5-4 - CP % de Lugares Femininos no Parlamento

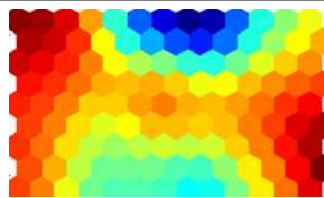


Figura 5-3 - CP Rácio na Participação no Mercado de Trabalho (F/M)

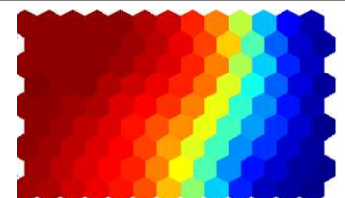


Figura 5-2 - CP Número de Partos com Assistência Médica (%)

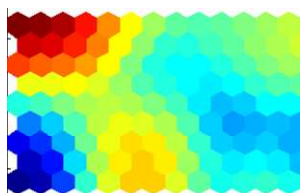


Figura 5-15 – CP Nível se Satisfação Com a Liberdade de Escolha (%)

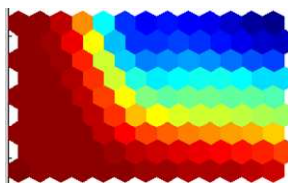


Figura 5-16 - CP Nível de Democracia (0 Não Dem; 1 Dems/Alternativa; 2

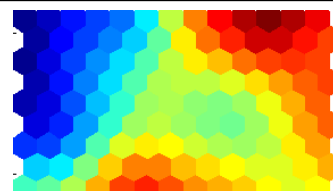


Figura 5-14- CP Nível de Violação dos Direitos Humanos (1-5)

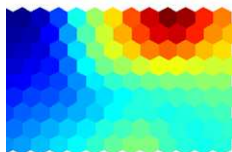


Figura 5-24 – CP Índice Liberdade de Imprensa (mais baixo mais liberdade)



Figura 5-23 – CP Jornalistas Presos (Número)

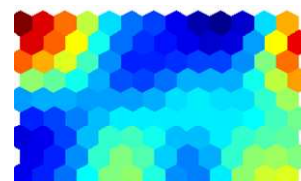


Figura 5-22- CP % de Pessoas que omite a opinião a representantes oficiais)

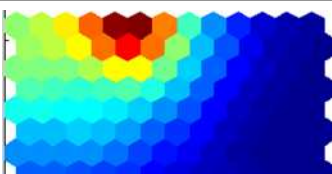


Figura 5-21 – CP Emissão de CO2 per capita



Figura 5-20 – CP Área do Território Protegida (5 do país)

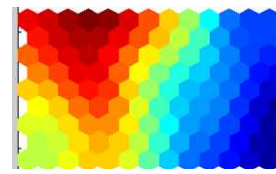


Figura 5-19 – CP % da População Urbana

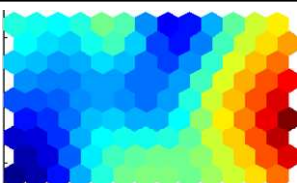


Figura 5-18 – CP % da População Empregada (15-64)

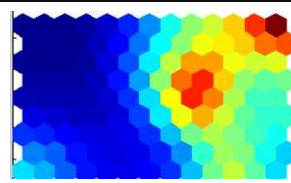


Figura 5-17 % do crescimento do número de utilizadores de internet (2000-2008)

5.1. Interpretação do Plano de Componentes

Da análise dos Planos de Componentes (Figura 5-13), podemos dizer que os países que têm um maior IDH encontram-se no canto superior esquerdo da matriz e que o IDH vai diminuindo para a direita.

Uma vez que fazem parte do IDH a Esperança de Vida, Educação e Rendimento, estas variáveis têm distribuições muito semelhantes ao IDH. No entanto o CP do rendimento (Figura 5-7) permite verificar que o Rendimento é mais concentrado no canto superior e as restantes dimensões estão mais distribuídas na vertical, reforçando a ideia que o “desenvolvimento é mais que rendimento”.

O CP do Coeficiente de Gini (Figura 5-5), permite identificar que os países com maior desigualdade se encontram no centro inferior e canto inferior direito da matriz. Estes países apresentam médio IDH (centro da matriz) ou baixo (canto inferior direito da matriz) e são muito desiguais.

Observando os indicadores da dimensão da Desigualdade de Género, nomeadamente a participação no mercado de trabalho (Figura 5-3) e os lugares no parlamento (Figura 5-4), no centro superior da matriz os valores são baixos, sugerindo desigualdade de género nestes países. Em contrapartida, os lugares femininos no parlamento destacam-se nos países ricos e o rácio de participação feminina destaca-se nos países mais ricos, mas também nos países menos desenvolvidos (canto inferior direito da matriz). O número de partos com assistência média e a taxa de fertilidade adolescente, têm uma distribuição semelhante ao nível de desenvolvimento, por isso pode-se concluir que são mais consequência do nível de desenvolvimento humano do que reflexos de desigualdade de género.

Na dimensão de *Empowerment*, pode-se observar que os países com níveis de democracia mais baixos encontram-se no canto superior direito e centro superior (Figura 5-16). É também nesses países que há mais violações de direitos humanos (Figura 5-14), onde mais jornalistas são presos (5-23) e onde a liberdade de imprensa é mais baixa (5-24). Assim é possível identificar países não democráticos pouco desenvolvidos e países não democráticos com um desenvolvimento médio. Nesta dimensão, o indicador nível de satisfação com a liberdade de escolha (Figura 5-15) é

muito alto nos países com maior nível de desenvolvimento e atinge os valores mais baixos, nos países com um desenvolvimento médio alto.

Na dimensão de Sustentabilidade, as emissões de CO₂ (Figura 5-21) têm valores muito altos nos países do centro superior da matriz. Os países com mais áreas protegidas encontram-se no centro inferior (Figura 5-20).

Em relação aos outros indicadores, pode-se observar que os países que apresentam maior taxa de crescimento da internet são os regimes menos democráticos pobres (canto superior direito da Figura 5-18) e que a maior taxa de população empregada se encontra nos países com menor desenvolvimento (Figura 5-17).

5.2. Identificação

A decisão final decorreu depois de se correr o algoritmo várias vezes aos dados e obter-se consistência nos clusters.

Compararam-se os perfis dos países através dos vários Planos de Componentes, de forma a ganhar conhecimento sobre as suas características e verificar se estes estavam de acordo com o conhecimento obtido na fase exploratória. Como o cluster de Médio Desenvolvimento Não Democrático não apresentava a cor mais consistente, experimentou-se partir esse cluster em dois, sendo possível obter desta forma diferenças nos indicadores de Desigualdade de Género, mas não uma diferença significativa que justificasse esta partição. Por oposição, no caso dos países com baixo desenvolvimento justificou-se a separação em dois clusters, com mais ou menos democracia.

A dimensão dos Clusters e a possibilidade de identificar quais as Dimensões de Desenvolvimento prioritárias para implementação de políticas foi importante na definição dos clusters.

A Matriz U foi assim segmentada conforme é indicado na Figura 5-25:

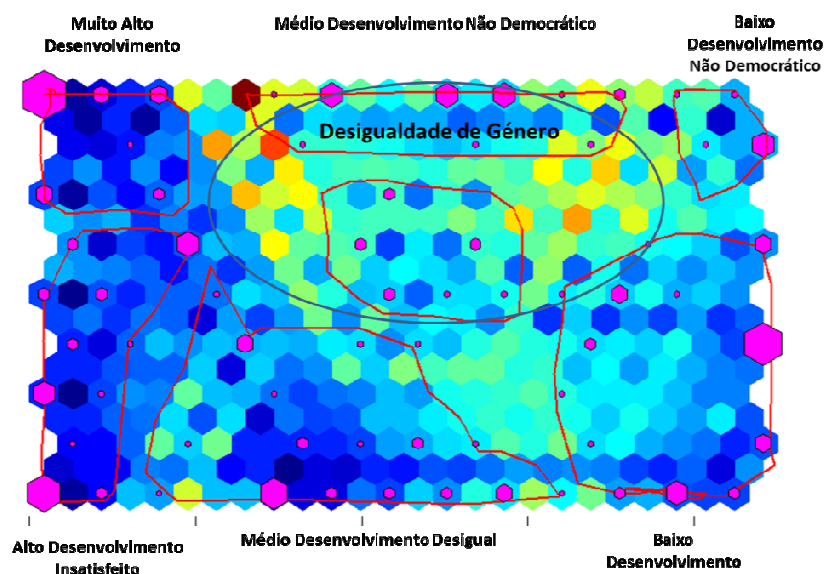


Figura 5-25 - Segmentação dos Países.

5.3. Caracterização dos Clusters

Conforme se pode ver na Figura 5-25, os critérios para a segmentação foram sobretudo o nível de desenvolvimento humano dado pelos indicadores e o nível democrático dos países. Os critérios de Desigualdade de Gênero e de Insatisfação também foram usados.

Na página seguintes encontra-se uma tabela com as médias dos indicadores para cada um dos segmentos, que foram usadas para os caracterizar (Figura 5-26).

Analisando a tabela com as médias dos clusters podemos caracterizá-los da seguinte forma:

- **Muito Alto Desenvolvimento** - Este cluster é composto por 24 países que têm em média um IDH muito alto (0,87), indicadores de igualdade de gênero altos e níveis de participação na sociedade também altos. Estes países são sobretudo democracias e os cidadãos destas sociedades estão satisfeitos com a Liberdade de Escolha (84%). Na dimensão da

sustentabilidade, estes países são dos que emitem mais CO₂ e têm um valor de áreas protegidas dentro da média. Estes países são os países mais urbanos (80% da população vive em áreas urbanas).

Dimensão a Desenvolver neste Cluster: Sustentabilidade

- **Alto Desenvolvimento Insatisfeitos** – Este cluster é composto por 27 países. Estes países têm níveis de desenvolvimento altos (média do IDH 0,75), indicadores de igualdade de género acima da média, mas são sociedades em que o nível de satisfação com a Liberdade de Escolha está abaixo da média (53%). Estes países são semelhantes aos de Muito Alto Desenvolvimento, distinguindo-se por terem um Rendimento Disponível abaixo, serem mais rurais, emitirem menos CO₂ e terem sociedades menos satisfeitas com as suas liberdades de escolha.

Dimensão a Desenvolver neste Cluster: *Empowerment* (Envolvimento da Sociedade Civil) e Rendimento

- **Médio Desenvolvimento, Não Democrático e com Desigualdade de Género** – Este cluster é composto por 27 países com um rendimento acima da média e acima da média do cluster Alto Desenvolvimento Insatisfeitos, mas com um IDH menor (o qual mesmo assim é alto (0,78)). Este é o cluster onde a participação das mulheres no mundo do trabalho é menor (Rácio na Participação no Mercado de Trabalho Feminino é 0,59) e onde há menos lugares femininos no parlamento. Nos outros indicadores de Desigualdade de Género os valores não são abaixo da média. Os indicadores da Dimensão de *Empowerment* são os mais baixos, caracterizando-se por serem em geral ditaduras (média do índice de democracia é 0,63 correspondendo a regimes não democráticos). Estes países são também os países com maiores índices de emissões de CO₂.

Dimensão a Desenvolver neste Cluster: *Empowerment* (Democracia) e Igualdade de Género Social

Media	Geral						Igualdade de Género					Empowerment					Sustentabilidade			Outros			
	Total	IDH	LifExp	MYSchool	ExpYSchool	Rendimento	Gini	AdoFert	SeatParl	LFR	BirthsAten	Freedom	Democ	HRVio	PressFre	JourImpr	PoliEng	CO2	ProtAre	Water	Empl	Inter	Urban
Total	144	0.64	69	8	12	13,521	40	54	18	0.74	80	66.39	1.46	2.63	28	0.65	20	5	12	14	59	2,481	0.58
Muito Alto Desenvolvimento	24	0.87	80	11	16	34,780	32	14	27	0.86	100	84.38	2.00	1.63	4	0.00	27	10	14	0	56	187.89	0.80
Alto Desenvolvimento Insatisfeitos	27	0.75	74	10	14	14,444	35	26	17	0.77	97	53.41	2.00	2.07	15	0.04	16	6	10	4	50	1564.99	0.65
Médio Desenvolvimento Não Democrático e com Desigualdade de Género	27	0.70	72	8	13	18,083	39	29	15	0.59	94	65.89	0.63	2.74	43	1.70	17	11	7	7	54	3,343	0.69
Médio Desenvolvimento Desigual	28	0.61	70	7	12	6,091	49	73	16.04	0.66	76	68.36	1.86	3.04	27	0.11	20	2	20	12	60	1,192	0.54
Baixo Desenvolvimento	22	0.35	54	4	8	1,128	43	119	16	0.84	44	59.64	1.50	3.05	31	0.27	21	0	14	40	70	3,647	0.30
Baixo Desenvolvimento Não Democrático	16	0.44	59	4	8	2,416	43	80	19	0.73	52	68.00	0.38	3.56	54	2.38	24	1	9	32	65	6,664	0.39

Figura 5-26 – Média dos Indicadores dos segmentos encontrados

- **Médio Desenvolvimento Desigual** – este cluster é composto por 28 países que se caracterizam por terem IDH médio (0,61), mas com um coeficiente de Gini alto (0,49), reflectindo uma grande desigualdade de rendimento. Nestes países o Rácio da Participação no Mercado de Trabalho Feminino é abaixo da média (0,66). O índice de violação de direitos humanos é acima da média (3,04), mas os restantes indicadores desta dimensão estão dentro da média. Na dimensão da sustentabilidade, é o cluster com maior área protegida por país (20%).

Dimensão a Desenvolver neste Cluster: Igualdade (Rendimento e de Género Social) e *Empowerment* (Direitos do Homem).

- **Baixo Desenvolvimento** – este cluster é composto por 22 países com IDH baixo (0,35) e que se caracterizam por ter valores baixos nas dimensões de desenvolvimento básicas que compõem o IDH (Educação e Esperança de Vida). Estes países também se caracterizam por ter um nível de desigualdade de rendimento alto (coeficiente de Gini 43). Na dimensão de desigualdade de género, apresentam valores altos da taxa de fertilidade infantil e um valor baixo no número de partos assistidos, mas em contrapartida a participação das mulheres no mundo do trabalho é acima da média (0,84). Neste cluster a percentagem de pessoas sem acesso a água potável é de 40% e é uma população sobretudo rural (apenas 30% vive em centros urbanos. Este cluster apresenta a maior taxa de população empregada (70%). Na dimensão *Empowerment* apresenta valores acima da média no indicador de violação de direito humanos, embora na média sejam democracias.

Dimensão a Desenvolver neste Cluster: Básicas que compõem o IDH, *Empowerment* (Direitos do Homem) e Igualdade de Género(Saúde)

- **Baixo Desenvolvimento Não Democrático** – este cluster é composto por 16 países e apresenta valores de IDH um pouco acima do cluster “Baixo Desenvolvimento” devido a terem um rendimento superior, mas mesmo assim a média do IDH é baixa (0,44). Quase todas as dimensões de desenvolvimento têm valores semelhantes ao cluster Baixo Desenvolvimento, mas apresenta um valor próximo de 0 no indicador regimes democráticos e o maior valor no indicador de violação de direitos humanos. A satisfação com a Liberdade de Escolha e a participação na vida política são acima da média (68%). Estes países são mais urbanos que o cluster Baixo Desenvolvimento, mas a população urbana só representa 48% da população.

Dimensão a Desenvolver neste Cluster: Básicas que compõem o IDH, Empowerment (Direitos do Homem) e Igualdade de Género Saúde

5.4. Comparação dos segmentos com a classificação do IDH

Comparou-se a segmentação encontrada com a classificação do IDH e que pode ser consultada com mais detalhe no Anexo D. Como se pode ver na figura 5-27, que faz o cruzamento dos clusters encontrados com a classificação do IDH, os clusters do SOM seguem o padrão do IDH. O SOM não usou o indicador IDH, mas sim os indicadores que o compõem. Por esta razão os pesos que são dados a cada uma das dimensões que compõem o índice no SOM têm um peso calculado pelo próprio algoritmo e se só tivéssemos usados estes indicadores iríamos certamente obter uma distribuição dos países diferente da do IDH.

SOM\IDH	Muito Alto	Alto	Médio	Baixo	Total
Muito Alto	23	1			24
Alto Insatisfeito	10	13	4		27
Médio Não Democrático	4	15	8		27
Médio Desigual		9	14	5	28
Baixo			1	21	22
Baixo Não Democrático			4	12	16
Total	37	38	31	38	144

Figura 5-27 – Quadro que cruza os clusters do SOM com o IDH

Na Figura 5-28, pode-se ver a distribuição dos indicadores usados na segmentação do IDH, os quais quando comparada com os clusters obtidos no SOM (Figura 5-26) permite concluir que através da segmentação do SOM é possível identificar de uma forma mais eficaz as dimensões de desenvolvimento que é necessário desenvolver para cada conjunto de países.

Assim, por exemplo, se analisarmos as dimensões por Índice de Desenvolvimento, concluiríamos que os países com mais actos de violação de Direitos Humanos são países de Baixo Desenvolvimento, mas como o SOM demonstra, há também países com indicadores nas dimensões básicas do desenvolvimento humano alto e médio que têm regimes pouco democráticos e onde há violação dos direitos humanos.

Country	Geral							Desigualdade de Género				Empowerment						Sustentabilidade			Outros	
	Total	IDH	LifExp	MYSchool	ExpYSchool	Rendimento	Gini	AdoFert	SeatParl	LFR	BirthsAten	Freedom	Democ	HRVio	PressFrt	JourImpr	PoliEng	CO2	ProtAre	Water	Empl	Urban
Muito Alto	37	0.86	80	11	16	34.048	33	12	23	0.80	100	76.89	1.81	1.68	8	0.00	23	12	13	0	56	0.78
Alto	38	0.73	74	9	14	12.669	41	40	16	0.67	97	60.08	1.47	2.74	30	0.95	17	6	12	4	53	0.71
Medio	31	0.60	68	7	11	4.971	44	55	17	0.68	78	69.00	1.35	2.87	40	1.29	18	2	13	14	58	0.47
Baixo	38	0.37	55	4	8	1.359	43	107	17	0.78	45	60.34	1.18	3.24	34	0.47	22	0	11	37	67	0.34

Figura 5-28 - Média dos Indicadores do IDH

6. Conclusões

O Índice de Desenvolvimento Humano foi lançado em 1990 e desde a sua publicação que tem originado um intenso debate. A maioria das críticas relaciona-se com: (a) fórmula de cálculo; (b) dimensões de desenvolvimento usadas (apenas usa as mais básicas); (c) peso de cada uma dessas dimensões e (d) os indicadores usados para representar essas dimensões.

Sendo o SOM uma rede neuronal não supervisionada, testou-se este algoritmo num conjunto de indicadores, de forma a tentar resolver algumas das questões levantadas pelos críticos do IDH. O SOM permite usar um grande número de indicadores e o peso de cada indicador para a explicação do problema é determinado pelo próprio algoritmo. O SOM tem também uma grande vantagem na forma de representação dos resultados, uma rede que pode ser bidimensional (foi a escolhida) e onde é possível identificar conjuntos de dados (países) que têm características semelhantes entre si e distintas dos restantes (clusters).

Os indicadores seleccionados tentaram englobar não só as dimensões do IDH, mas também as dimensões sugeridas por ul Haq: Igualdade (Género e de Distribuição), *Empowerment* e Sustentabilidade. Por se considerar importante na caracterização dos países considerou-se mais alguns indicadores.

Através do SOM foi possível identificar os seguintes clusters:

- Muito Alto Desenvolvimento – países que têm valores altos em quase todas as dimensões de desenvolvimento humano, excepto na sustentabilidade;
- Alto Desenvolvimento Insatisfeitos – países que têm valores de desenvolvimento altos, mas inferiores ao segmento muito alto e onde não há um envolvimento tão grande com a actividade política e onde as pessoas apresentam valores de satisfação com as suas liberdades de escolha abaixo da média;
- Médio Desenvolvimento Não Democrático e com Desigualdade de Género – países que têm IDH alto ou médio, mas onde os indicadores de

democracia e de igualdade de género, na perspectiva social e não física, têm valores abaixo da média;

- Médio Desenvolvimento Desigual – nestes países o IDH é médio, mas há uma grande desigualdade de rendimento e alguma desigualdade de género. Neste países o indicador de violação dos direitos humanos é alto;
- Desenvolvimento Baixo – países pobres que apresentam índices baixos em quase todos os indicadores usados;
- Desenvolvimento Baixo Não Democrático – países que em termos de IDH estão um pouco acima dos países de desenvolvimento baixo, mas mesmo assim apresentam indicadores baixos em quase todos os indicadores de desenvolvimento, apresentando valores muito baixos nos indicadores da democracia e da violação dos direitos humanos.

A vantagem dos clusters do SOM versus a classificação do IDH é permitir identificar de uma forma mais rápida dimensões de desenvolvimento que podem estar em perigo e que no IDH não são identificadas, como por exemplo os Direitos Humanos e a Sustentabilidade. Com esta segmentação é possível agrupar os países em várias dimensões de desenvolvimento e identificar prioridades para as políticas de desenvolvimento.

7. Limitações e recomendações para trabalhos futuros

O presente trabalho seguiu uma linha de desenvolvimento que se considerou adequada para dar resposta ao objectivo definido, que era aplicar o SOM a um conjunto de dados que representasse mais dimensões de desenvolvimento do que as apresentadas actualmente pelo IDH. No entanto o estudo tem algumas limitações que não devem ser ignoradas e que devem ser tidas em conta para futuros projectos.

Num futuro projecto dever-se-á incluir mais dimensões de desenvolvimento e mais indicadores (de preferência todos os disponíveis). Alguns indicadores das dimensões utilizadas não foram usados por terem muitos *missing values*, mas poder-se-ia usar técnicas de *Datamining* para resolver esta questão.

Na determinação dos clusters, as experiências que foram feitas foi em termos de taxa de aprendizagem mudando muito pouco os indicadores. Seria interessante fazer testes alterando os indicadores e desta forma determinar os indicadores mais relevantes para segmentar os países.

ANEXOS

Anexo A: Definição do Conceito de Desenvolvimento Humano do 1º

Relatório

BOX 1.1

Human development defined

Human development is a process of enlarging people's choices. In principle, these choices can be infinite and change over time. But at all levels of development, the three essential ones are for people to lead a long and healthy life, to acquire knowledge and to have access to resources needed for a decent standard of living. If these essential choices are not available, many other opportunities remain inaccessible.

But human development does not end there. Additional choices, highly valued by many people, range from political, economic and social freedom to opportunities for being creative and productive, and enjoying personal self-respect and guaranteed human rights.

Human development has two sides: the formation of human capabilities — such as improved health, knowledge and skills — and the use people make of their acquired capabilities — for leisure, productive purposes or being active in cultural, social and political affairs. If the scales of human development do not finely balance the two sides, considerable human frustration may result.

According to this concept of human development, income is clearly only one option that people would like to have, albeit an important one. But it is not the sum total of their lives. Development must, therefore, be more than just the expansion of income and wealth. Its focus must be people.

Fonte: (UNDP, 1990)

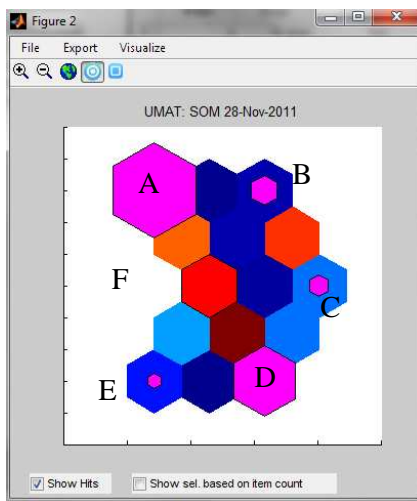
Anexo B: Definição dos Indicadores Usadas

Dimension	Indicador	Cod	Definition	
HDI	Human Development Index (HDI) value	IDH	A composite index measuring average achievement in three basic dimensions of human development—a long and healthy life, knowledge and a decent standard of living. See Technical note 1 for details on how the HDI is calculated.	
	Rank	Rank	HDI country ranking	
	Life expectancy at birth (years)	LifeExp	Number of years a newborn infant could expect to live if prevailing patterns of age-specific mortality rates at the time of birth stay the same throughout the infant's life.	
	Mean years of schooling (of adults) (years)	MYSchool	Average number of years of education received by people ages 25 and older, converted from education attainment levels using official durations of each level.	
	Expected years of schooling (of children under 7) (years)	ExpYSchool	Number of years of schooling that a child of school entrance age can expect to receive if prevailing patterns of age-specific enrolment rates persist throughout the child's life.	
	GNI per capita in PPP terms (constant 2005 international \$)	GNI	Aggregate income of an economy generated by its production and its ownership of factors of production, less the incomes paid for the use of factors of production owned by the rest of the world, converted to international dollars using purchasing power parity (PPP) rates, divided by midyear population.	
	Non Yield IDH	IDHNR	Non Yield IDH	
Income Gini coefficient	Gini	Measure of the deviation of the distribution of income (or consumption) among individuals or households within a country from a perfectly equal distribution. A value of 0 represents absolute equality; a value of 100 absolute inequality.		
Inequality	Adolescent fertility rate (births per 1,000 women aged 15-19)	AdoTert	Number of births to women ages 15 - 19 per 1,000 women ages 15 - 19.	
	Shares in parliament, female-male ratio	SeatParl	Ratio of seats held by a respective gender in a lower or single house or an upper house or senate, where relevant.	
	Population with at least secondary education (female/male ratio)	EdGR	Percentage of the population ages 25 and older that has attained a secondary or higher level of education, Woman versus Man.	
	Labour force participation rate (female-male ratio)	LFR	Ratio of female to male of the working-age population (ages 15-64) that actively engages in the labour market, by either working or actively looking for work.	
	Births attended by skilled health personnel	BirthsAten	Percentage of deliveries attended by personnel (including doctors, nurses and midwives) trained to give the necessary care to women during pregnancy, labour and the postpartum period. Excludes traditional birth attendants, whether trained or not.	
	Empowerment	Political freedom, democracy	Demncr	Score on the Democracy and Dictatorship measure of political regimes, which distinguishes between regimes in which executive and legislative offices are filled through contested elections and those in which they are not.
		Human rights violations	HRVlo	Score published by the Database of Political Institutions which calls it the Political Terror Scale) measuring human rights violations, as classified in Gibney, Cornett, and Wood (2010) and based on sanctioned killing, torture, disappearance and political imprisonment. The score is based on expert coding of the scope (type), intensity (frequency) and range of violence.
Press freedom		PressFre	Data refer to verified cases of journalists having been imprisoned as of December 1, 2009. Countries with a value of 0 did not have any verified cases as of that date.	
Journalists imprisoned		JourImpr	Data refer to verified cases of journalists having been imprisoned as of December 1, 2009. Countries with a value of 0 did not have any verified cases as of that date.	
Political engagement		PolEng	% of people who voiced opinion to public officials in 2008	
Sustainability	Carbon Dioxide Emissions per capita (tonnes)	CO2	Human-originated carbon dioxide emissions stemming from the burning of fossil fuels, gas flaring and the production of cement, divided by midyear population.	
	Protected area	ProtAre	% terrestrial area	
	Population Without Access to Improved Services of Water	Water	% of Population Without Access to Improved Services of Water in 2008	
Others	Employment to population ratio	Empl	% of population in ages 15-64 that works	
	Internet Users Growth	Inter	% growth, population-based internet users in 2008	
	Urban Population	Urban	% of Population that lives in cities	

Anexo C: Experiências com o SOM: Rede usando método K-means

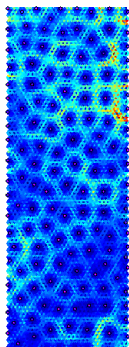
Figura Anexo C - 1 – UMAT de SOM para rede 2X3

São visíveis 5 clusters compostos por:



- A: 58 países com Desenvolvimento Alto
- B: 18 países com Desenvolvimento Médio, mas mais alto
- C: 13 países com Desenvolvimento Médio Baixo
- D: 43 países Desenvolvimento Médio Baixo
- E: 9 países com o *Empowerment* muito baixo
- F: 3 países que não se enquadram em nenhum dos restantes clusters (Tailândia, Guyana e Africa do Sul)

Figura Anexo C - 2 – UMAT de SOM para 30X100



Foram feitas 1000 interações no primeiro treino e 2000 no segundo. O erro é muito baixo ($Q=0,26$), mas obtiveram-se demasiados *clusters*

Figura Anexo C - 3 – UMAT de SOM com erro 0,3

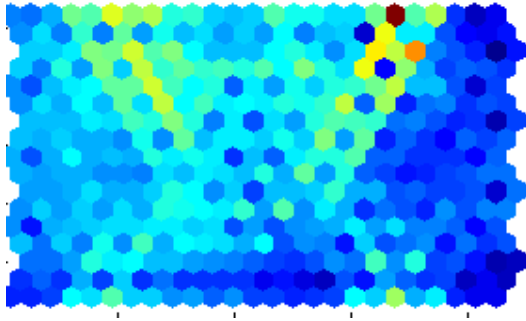


Figura Anexo C - 4 – UMAT de SOM com erro 0,5

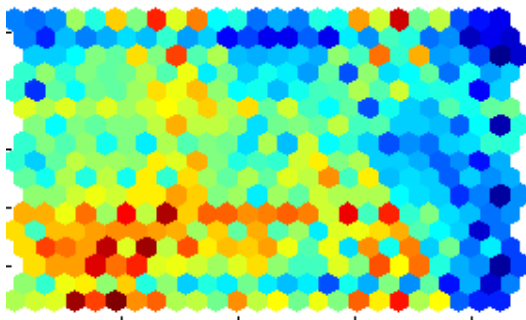
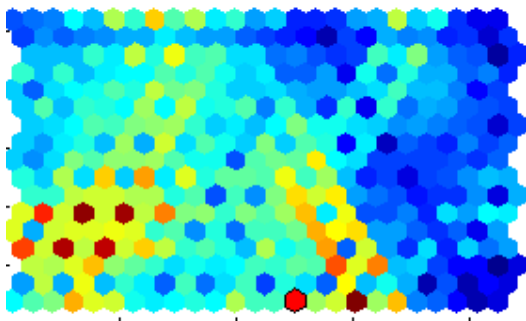


Figura Anexo C 5 – UMAT de SOM com erro 0,8



Anexo D: Comparação dos Clusters com a Classificação do IDH

País	Clusters Estudo	Cluster IDH	IDH	Ranking RNB
Norway	Muito Alto	Muito Alto	0.938	2
Australia	Muito Alto	Muito Alto	0.937	10
New Zealand	Muito Alto	Muito Alto	0.907	29
United States	Muito Alto	Muito Alto	0.902	7
Ireland	Muito Alto	Muito Alto	0.895	21
Netherlands	Muito Alto	Muito Alto	0.89	8
Canada	Muito Alto	Muito Alto	0.888	11
Sweden	Muito Alto	Muito Alto	0.885	13
Germany	Muito Alto	Muito Alto	0.885	15
Switzerland	Muito Alto	Muito Alto	0.874	9
France	Muito Alto	Muito Alto	0.872	19
Finland	Muito Alto	Muito Alto	0.871	20
Iceland	Muito Alto	Muito Alto	0.869	32
Belgium	Muito Alto	Muito Alto	0.867	17
Denmark	Muito Alto	Muito Alto	0.866	14
Luxembourg	Muito Alto	Muito Alto	0.852	5
Austria	Muito Alto	Muito Alto	0.851	12
UK	Muito Alto	Muito Alto	0.849	16
CzRepublic	Muito Alto	Muito Alto	0.841	33
Slovenia	Muito Alto	Muito Alto	0.828	28
Japan	Muito Alto	Muito Alto	0.884	18
Israel	Muito Alto	Muito Alto	0.872	25
Spain	Muito Alto	Muito Alto	0.863	22
Uruguay	Muito Alto	Alto	0.765	48
Korea	Alto Insatisfeito	Muito Alto	0.877	24
Greece	Alto Insatisfeito	Muito Alto	0.855	26
Italy	Alto Insatisfeito	Muito Alto	0.854	23
Slovakia	Alto Insatisfeito	Muito Alto	0.818	36
Malta	Alto Insatisfeito	Muito Alto	0.815	37
Estonia	Alto Insatisfeito	Muito Alto	0.812	40
Cyprus	Alto Insatisfeito	Muito Alto	0.81	35
Hungary	Alto Insatisfeito	Muito Alto	0.805	39
Portugal	Alto Insatisfeito	Muito Alto	0.795	34
Poland	Alto Insatisfeito	Muito Alto	0.795	38
Lithuania	Alto Insatisfeito	Alto	0.783	44
Latvia	Alto Insatisfeito	Alto	0.769	53
Romania	Alto Insatisfeito	Alto	0.767	55
Croatia	Alto Insatisfeito	Alto	0.767	42
Bulgaria	Alto Insatisfeito	Alto	0.743	59
Trinidad	Alto Insatisfeito	Alto	0.736	31
Peru	Alto Insatisfeito	Alto	0.723	67
Albania	Alto Insatisfeito	Alto	0.719	73
Ukraine	Alto Insatisfeito	Alto	0.71	77
Macedonia	Alto Insatisfeito	Alto	0.701	64
Georgia	Alto Insatisfeito	Alto	0.698	85
Armenia	Alto Insatisfeito	Alto	0.695	83
Turkey	Alto Insatisfeito	Alto	0.679	50
El Salvador	Alto Insatisfeito	Médio	0.659	78
Moldova	Alto Insatisfeito	Médio	0.623	99
Mongolia	Alto Insatisfeito	Médio	0.622	95
Kyrgyzstan	Alto Insatisfeito	Médio	0.598	106

Anexo D: Comparação dos Clusters com a Classificação do IDH (cont)

País	Clusters Estudo	Cluster IDH	IDH	Ranking RNB
Singapore	Médio Não Democrático	Muito Alto	0.846	6
UAE	Médio Não Democrático	Muito Alto	0.815	3
Qatar	Médio Não Democrático	Muito Alto	0.803	1
Bahrain	Médio Não Democrático	Muito Alto	0.801	27
Chile	Médio Não Democrático	Alto	0.783	49
Argentina	Médio Não Democrático	Alto	0.775	45
Kuwait	Médio Não Democrático	Alto	0.771	4
Libyan	Médio Não Democrático	Alto	0.755	41
Saudi Arabia	Médio Não Democrático	Alto	0.752	30
Malaysia	Médio Não Democrático	Alto	0.744	47
Belarus	Médio Não Democrático	Alto	0.732	54
Russian Federation	Médio Não Democrático	Alto	0.719	43
Kazakhstan	Médio Não Democrático	Alto	0.714	62
Azerbaijan	Médio Não Democrático	Alto	0.713	65
Bosnia	Médio Não Democrático	Alto	0.71	70
Iran	Médio Não Democrático	Alto	0.702	58
Tunisia	Médio Não Democrático	Alto	0.683	72
Jordan	Médio Não Democrático	Alto	0.681	80
Algeria	Médio Não Democrático	Alto	0.677	68
Gabon	Médio Não Democrático	Médio	0.648	56
Egypt	Médio Não Democrático	Médio	0.62	81
Uzbekistan	Médio Não Democrático	Médio	0.617	100
Guyana	Médio Não Democrático	Médio	0.611	97
South Africa	Médio Não Democrático	Médio	0.597	63
Syrian Arab Republic	Médio Não Democrático	Médio	0.589	87
Tajikistan	Médio Não Democrático	Médio	0.58	112
Morocco	Médio Não Democrático	Médio	0.567	89
Panama	Médio Desigual	Alto	0.755	51
Mexico	Médio Desigual	Alto	0.75	46
Costa Rica	Médio Desigual	Alto	0.725	60
Brazil	Médio Desigual	Alto	0.699	61
Venezuela	Médio Desigual	Alto	0.696	57
Ecuador	Médio Desigual	Alto	0.695	74
Belize	Médio Desigual	Alto	0.694	82
Colombia	Médio Desigual	Alto	0.689	66
Jamaica	Médio Desigual	Alto	0.688	76
Dominican Republic	Médio Desigual	Médio	0.663	69
Sri Lanka	Médio Desigual	Médio	0.658	86
Thailand	Médio Desigual	Médio	0.654	71
Bolivia 2	Médio Desigual	Médio	0.643	91
Paraguay	Médio Desigual	Médio	0.64	90
Philippines	Médio Desigual	Médio	0.638	92
Botswana	Médio Desigual	Médio	0.633	52
Namibia	Médio Desigual	Médio	0.606	79
Honduras	Médio Desigual	Médio	0.604	94
Indonesia	Médio Desigual	Médio	0.6	93
Nicaragua	Médio Desigual	Médio	0.565	103
Guatemala	Médio Desigual	Médio	0.56	88
India	Médio Desigual	Médio	0.519	96
Pakistan	Médio Desigual	Médio	0.49	102
Ghana	Médio Desigual	Baixo	0.467	121
Benin	Médio Desigual	Baixo	0.435	119
Nepal	Médio Desigual	Baixo	0.428	127
Togo	Médio Desigual	Baixo	0.428	137
Haiti	Médio Desigual	Baixo	0.404	134

Anexo D: Comparação dos Clusters com a Classificação do IDH

País	Clusters Estudo	Cluster IDH	IDH	Ranking RNB
Cambodia	Baixo	Médio	0.494	113
Kenya	Baixo	Baixo	0.47	115
Madagascar	Baixo	Baixo	0.435	132
Papua New Guinea	Baixo	Baixo	0.431	107
Nigeria	Baixo	Baixo	0.423	109
Uganda	Baixo	Baixo	0.422	125
Senegal	Baixo	Baixo	0.411	114
Tanzania	Baixo	Baixo	0.398	124
Zambia	Baixo	Baixo	0.395	122
Gambia	Baixo	Baixo	0.39	123
Malawi	Baixo	Baixo	0.385	135
Ethiopia	Baixo	Baixo	0.328	131
Sierra Leone	Baixo	Baixo	0.317	138
Mali	Baixo	Baixo	0.309	129
Burkina Faso	Baixo	Baixo	0.305	126
Liberia	Baixo	Baixo	0.3	142
Chad	Baixo	Baixo	0.295	130
Mozambique	Baixo	Baixo	0.284	136
Burundi	Baixo	Baixo	0.282	141
Niger	Baixo	Baixo	0.261	140
Congo (Democratic Republic of the)	Baixo	Baixo	0.239	143
Zimbabwe	Baixo	Baixo	0.14	144
China	Baixo Não Democrático	Médio	0.663	75
Viet Nam	Baixo Não Democrático	Médio	0.572	101
Lao	Baixo Não Democrático	Médio	0.497	105
Congo	Baixo Não Democrático	Médio	0.489	98
Bangladesh	Baixo Não Democrático	Baixo	0.469	118
Cameroon	Baixo Não Democrático	Baixo	0.46	108
Myanmar	Baixo Não Democrático	Baixo	0.451	117
Yemen	Baixo Não Democrático	Baixo	0.439	104
Mauritania	Baixo Não Democrático	Baixo	0.433	110
Angola	Baixo Não Democrático	Baixo	0.403	84
Rwanda	Baixo Não Democrático	Baixo	0.385	128
Sudan	Baixo Não Democrático	Baixo	0.379	111
Afghanistan	Baixo Não Democrático	Baixo	0.349	120
Guinea	Baixo Não Democrático	Baixo	0.34	132
Central African Republic	Baixo Não Democrático	Baixo	0.315	139
Côte d'Ivoire	Baixo Não Democrático	Baixo	0.397	116

Figura Anexo D 1 Comparação da Segmentação encontrada com o IDH

Referências Bibliográficas

- Ahn, Jae-Wook, e Sue Yeon Syn. **Self Organizing Map Tutorial System**. 2005.
<http://www.sis.pitt.edu/~ssyn/som/som.html> (acedido em 15 de Maio de 2011).
- Alkire, Sabina. **Human Development: Definitions, Critiques, and Related Concepts**. 2010. University of Oxford: United Nations Development Programme.
- Anonimo. **Kohonen Network Self Organizing Map for Color Organization**. 2010.
<http://www.youtube.com/watch?v=-6a7LATC-9g&feature=related> (acedido em Agosto de 15 de 2011).
- Barro, R.J., e J.W.Lee. **A New Data Set of educational Attainment in the World, 1995-2010**, National Bureau of Economic Research (National Bureau of Economic Research) 15902
- Chakravarty, Satya R. **A Generalized Human Development Index**. Review of Development Economics 7 (2003): 99-114.
- Dasgupta, Partha, e Martin Weale. **On measuring the quality of life**. World Development 20 (1992): 119-131.
- Desai, M.J. "Human Development: Concepts and Measurement." European Economic 35 (1991): 350-357.
- Despotis, DK. **A reassessment of the human development index via data envelopment analysis**. Journal of the Operational Research Society 56 (2005): 969-980.
- Explorer, Google Public Data. Google Public Data Explorer. 2011.
<http://www.google.com/publicdata/home> (acedido em 23 de August de 2011).

- Fayyad, Usama, Gregory Piatetsky-Shapiro, e Padhraic Smyth. **The KDD Process for Extracting Useful Knowledge from Volumes of Data.** Communications of the ACM 39 (1996): 34.
- Foster, James E., Luis F. Lopez-Calva, e Miguel Szekely. **Measuring the Distribution of Human Development: methodology and an application to Mexico.** Journal of Human Development 1 (2005): 5-29.
- Gaye, Amie, Jeni Klugman, Milorad Kovacevic, Sarah Twigg, e Eduardo Zambrano. **Measuring Key Disparities in Human Development: The Gender Inequality Index.** New York: PNUD, 2010.
- Gertner, J. **The rise and fall of GDP.” 13 de May de 2010.** <http://www.nytimes.com/2010/05/16/magazine/16GDP-t.html> (acedido em 3 de Setembro de 2011).
- Haq, Mahbub ul. **Reflections on Human Development.** In Chapter 4, Oxford University Press, 1995.
- Henriques, Roberto André Pereira. **Artificial Intelligence in Geospatial Analysis:applications of Self-Organizing Maps in the context of Geographic Information Science.** Lisbon: ISEGI, 2010.
- Henriques, Roberto, e Fernando Bação. **SOMGIS: Uma Ferramenta para construir Regiões. Lisboa: Instituto Superior de Estatística e Gestão de Informação, Universidade Nova de Lisboa,, 2004.**
- Herrero, Carmen, Ricardo Martinez, e Antonio Villar. **Multidimensional Social Evaluation: An Application to the Measurement of Human Development.** Review of Income and Wealth 3 (2010): 483-497.
- Hicks, Douglas A. **The Inequality-Adjusted Human Development Index:A Constructive Proposal.** World Development 25 (1997): 1283-1298.
- ISEGI. GeoSOM Suit. 2009. [www.isegi.unl.pt/labnt/GeoSOM Suit](http://www.isegi.unl.pt/labnt/GeoSOM%20Suit) (acedido em 24 de Novembro de 2011).

- Klasen, S., G. Nguefack, e W. Zucchini. **On Weighting the Components of the Human Development Index: A Statistical Justification.** Journal of Human Development and Capabilities 2 (2011): 183-202.
- Klugman, Jeni, Francisco Rodriguez, e Hyung Jin Choi. **The HDI 2010: New Controversies, Old Critiques.** New York: UNDP-HDRO, 2011.
- Kohonen, T. **Self-organizing formation of topologically correct feature maps.** RecMap: rectangular map approximations 43, n.º 1 (1982): 59-69.
- Kohonen, T., e S.Kaski. **Exploratory Data Analysis by the Self Organizing Map: Structures of Welfare and Poverty in the World.** World Scientific Singapore, 1996: 498-507.
- Kohonen, T., J. Hynninen, J. Kangas, e J. Laaksonen. SOM_PAK. 1995. www.cochlea.hut.fi (acedido em 14 de Dezembro de 2010).
- Lind, N. C. **Values Reflected in the Human Development Index.** Social Indicators Research 66 (2004): 283-293.
- Lind, Niels C. **Some Thoughts on the Human Development Index.** Social Indicators Research 27, n.º 1 (1992): 89-101.
- McGillivray, Mark. **The human development index: Yet another redundant composite development indicator?** World Development 19 (1991): 1461-1468.
- Nathan, H. S. K., S. Mishra, e B. S. Reddy. **An Alternative Approach to Measure HDI.** January de 2008. <http://www.igidr.ac.in/pdf/publication/WP-2008-001.pdf> (acedido em 3 de Setembro de 2011).
- Noorbakhsh, Farhad. **The Human Development Index: Some Technical Issues and Alternative Indices.** Journal of international Development, 1998: 589-605.
- Nussbaum, Martha. **Women and Human Development - The Capabilities Approach.** 1. United States of America: Cambridge University Press, 2000.
- O'Neill, Helen. **Ireland's Foreign Aid.** Irish Studies in International Affairs in 2004 16, n.º Irish Development Aid and Irish Foreign Relations (2005): 279-316.

- Palazzi, P., e A. Lauri. **The Human Development Index: Suggested Corrections.** Banca Nazionale del Lavoro Quarterly 51 (1998): 193-221.
- PNUD. HDR Chapters and Downloads. 2010.
<http://hdr.undp.org/en/reports/global/hdr2010/chapters/> (acedido em 23 de July de 2011).
- Ravallion, M. **Good and Bad Growth: The Human Development Reports.** World Development 25 (1997): 631-638.
- Sagar, Ambuj D., e Adil Najam. **Survey, The human development index: a critical review.** Ecological Economics 25, n.º 1 (1998): 249-264.
- Sen, A.K. **The Income Component of the Human Development Index.** Journal of Human Development and Capabilitie 1, n.º 1 (2000): 83-106.
- Sen, Amartya. **A 20th Anniversary Human Development Discussion with Amartya Sen** (4 de November de 2010).
- Sen, Amartya K. **Commodities and Capabilities.** 3. New York: Oxford University Press, 1999.
- Seth, Suman. **Inequality, Interactions, and Human Development.** Journal of Human Development and Capabilities 10 (2009): 375-396.
- Som Pak. 7 de April de 1995. www.cis.hut.fi/research/som_pak (acedido em 1 de 5 de 2011).
- Srinivasan, T. N. **Human Development: A New Paradigm or Reinvention of the Wheel?** The American Economic Review 84 (1994): 238-243.
- UNDP. **Human Development Report 1990: Concept and Measurement of Human Development.** New York: Oxford University Press, 1990.
- UNDP. **Human Development Report: New Dimensions of human security.** New York: Oxford University Press, 1994.

UNDP. **Human Development Report: Gender and human development.** New York: Oxford University Press, 1995.

UNDP. **Human Development to Eradicate Poverty.** New York: Oxford Publication, 1997.

UNDP. **Human Development Report 2007/2008:Fighting climate change:human solidarity in a divided world.** New York: Oxford University Press, 2008.

UNDP. **Human Development Report: The Real Wealth of Nations.** New York: Oxford University Press, 2010.