



Departamento de Ciências e Tecnologias da Informação

**Data Mining no Contra-Terrorismo**  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Ana Isabel Pereira Santos Figueiredo Morais

Dissertação submetida como requisito parcial para obtenção do grau de  
Mestre em Gestão de Sistemas de Informação

Orientador:  
Doutor Raul M. S. Laureano, Professor Auxiliar  
ISCTE-Instituto Universitário de Lisboa

Setembro, 2012

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

## **Agradecimentos**

---

Este momento tão desejado marca o final de uma etapa que se iniciou há sete anos atrás, em que a vida me proporcionou simultaneamente a possibilidade de concretização de dois desejos, ser mãe e iniciar um percurso no ensino universitário. Não podendo nenhum deles ser adiado e não conseguindo prescindir de nenhum, arrisquei a sua realização conjunta. Foi de facto um caminho marcado por muito esforço, muitas cedências, renúncias e ausências, pelo que devo um agradecimento redobrado ao meu marido e ao meu filho por todo o apoio e compreensão demonstrados ao longo deste período.

Um agradecimento muito especial devo também ao meu orientador, Professor Doutor Raul Laureano, cujo apoio incondicional foi determinante para a concretização deste trabalho, com a sua motivação, conhecimentos e alento constantes, sempre presente, tantas vezes abdicando dos seus momentos em família e pausas para descanso.

Ao meu chefe, Dr. João Ventura, agradeço toda a compreensão e conhecimentos como perito de terrorismo transmitidos ao longo da execução desta investigação.

Gostaria ainda de expressar o meu agradecimento a todos que de forma geral acompanharam este percurso, colegas, professores, amigos e restante família.

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

## Resumo

---

O trabalho apresentado procurou constituir uma abordagem ao estudo do terrorismo no sentido de estabelecer mais um passo para a compreensão dos processos ou indicadores subjacentes à radicalização nas suas formas mais extremas. Definiram-se como objectivos a identificação dos factores determinantes do terrorismo e a concepção de um modelo explicativo e preditivo para a ocorrência e impacto de actos terroristas.

Entraram em análise como potenciais factores explicativos variáveis das dimensões económica, sociodemográfica, política, religiosa e de carácter multidimensional, integradas no âmbito de uma amostra de 173 países, organizada segundo modelo de dados em painel, por um período de 20 anos decorridos entre 1990 e 2010.

Através da aplicação de Árvores de Regressão com o algoritmo CART, uma das várias técnicas de data mining que são particularmente adequadas quando se têm muitos dados de origem multidimensional, foram identificados diversos factores considerados como determinantes do terrorismo, pertencentes às cinco dimensões de variáveis, tendo-se confirmado a hipótese que sustenta que os factores sociodemográficos são os que melhor explicam a actividade terrorista. Os modelos explicativos para os três indicadores de terrorismo não traduziram capacidades explicativas e preditivas muito elevadas, mas dada a complexidade e subjectividade que envolve o fenómeno do terrorismo, considera-se que os modelos gerados constituem mais um contributo para a literatura sobre terrorismo.

Não tendo sido identificado nenhum estudo anterior em que esta técnica de análise tenha sido utilizada, pôde-se confirmar a sua eficácia para a concretização deste e de futuros estudos na área do terrorismo.

**Palavras-chave:** Determinantes do Terrorismo, *Data Mining*, Árvores de Regressão, CART.

## **Abstract**

---

The presented study aimed to provide an approach to the study of terrorism in order to establish a further step towards the understanding of the processes or underlying indicators of radicalization in its most extreme forms. The main goals were the identification of the determinants of terrorism and the conception of an explanatory and predictive model for the occurrence and impact of terrorist acts.

Under analysis were, as potential explanatory factors, economic, sociodemographic, political, religious and multidimensional variables, integrated within a sample of 173 countries, organized according to panel data model, for a period of 20 years between 1990 and 2010.

Through the application of Regression Trees with the CART algorithm, one of several data mining techniques that are particularly suitable to the analysis of many multidimensional data sources, there were identified several factors considered as determinants of terrorism, belonging to the five variable dimensions, having been confirmed the hypothesis that holds that sociodemographic factors are those that best explain terrorist activity. The explanatory models for the three indicators of terrorism did not hold explanatory and predictive capacities very high, but given the complexity and subjectivity involved in the phenomenon of terrorism, it is considered that the generated models are a further contribution to the literature on terrorism.

Having not been identified any previous studies in which this analysis technique has been used, it is therefore possible to confirm its efficiency to achieve this and future studies in the area of terrorism.

**Keywords:** Determinants of Terrorism, *Data Mining*, Regression Trees, CART.

## Índice

---

Agradecimentos .....	iii
Resumo .....	v
Abstract .....	vi
Índice .....	vii
Índice de Tabelas .....	ix
Índice de Figuras .....	x
Lista de Siglas e Acrónimos .....	xi
<b>1. Introdução .....</b>	<b>1</b>
1.1 Tema, Motivação e Contribuição .....	2
1.2 Objectivos .....	4
1.3 Abordagem Metodológica .....	4
1.4 Estrutura da dissertação .....	5
<b>Enquadramento Teórico .....</b>	<b>7</b>
<b>2. Terrorismo .....</b>	<b>7</b>
2.1 Definição de Terrorismo e Evolução Histórica do Conceito .....	7
2.2 Tipos de Terrorismo .....	10
2.3 Terrorismo e Crime Organizado .....	11
2.4 Classificação do Terrorismo .....	12
2.5 Sumário .....	13
<b>3. Caracterização Económica, Sociodemográfica, Religiosa e Política de um País..</b>	<b>15</b>
3.1 Indicadores Económicos .....	16
3.2 Indicadores Sociodemográficos e Religiosos .....	17
3.3 Indicadores Políticos .....	18
3.4 Sumário .....	18
<b>4. Os Factores Económicos, Sociodemográficos, Religiosos e Políticos no Terrorismo</b>	<b>21</b>
4.1 Factores Económicos .....	21
4.2 Factores Sociodemográficos e Religiosos .....	25
4.3 Factores Políticos .....	27
4.4 Sumário .....	29
<b>Estudo Empírico .....</b>	<b>31</b>
<b>5. Metodologia .....</b>	<b>31</b>
5.1 Hipóteses de Investigação .....	32

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

<b>5.2</b>	<b>Descoberta de Conhecimento em Bases de Dados.....</b>	<b>32</b>
<b>5.3</b>	<b>Fontes de Dados e Dados .....</b>	<b>34</b>
5.3.1	Dados de Terrorismo .....	34
5.3.2	Dados de Caracterização dos Países .....	38
<b>5.4</b>	<b>Técnicas de Data Mining .....</b>	<b>47</b>
<b>5.5</b>	<b>Aplicação da Metodologia CRISP-DM .....</b>	<b>53</b>
<b>5.6</b>	<b>Sumário.....</b>	<b>56</b>
<b>6.</b>	<b>Resultados .....</b>	<b>57</b>
<b>6.1</b>	<b>Caracterização da Amostra .....</b>	<b>57</b>
<b>6.2</b>	<b>Caracterização da Actividade Terrorista .....</b>	<b>59</b>
<b>6.3</b>	<b>Explicação da Actividade Terrorista.....</b>	<b>62</b>
6.3.1	A Importância Relativa dos Atributos Explicativos .....	62
6.3.2	Modelo Explicativo do Número Total de Incidentes Terroristas.....	72
6.3.3	Modelo Explicativo do Número Total de Mortos.....	77
6.3.4	Modelo Explicativo do Número Total de Feridos .....	81
6.3.5	Comparação dos Modelos Explicativos.....	84
<b>6.4</b>	<b>Discussão de Resultados .....</b>	<b>89</b>
<b>6.5</b>	<b>Sumário.....</b>	<b>93</b>
<b>7.</b>	<b>Conclusões.....</b>	<b>95</b>
<b>7.1</b>	<b>Contributos.....</b>	<b>95</b>
<b>7.2</b>	<b>Limitações.....</b>	<b>96</b>
<b>7.3</b>	<b>Pistas Futuras.....</b>	<b>96</b>
	<b>Referências Bibliográficas .....</b>	<b>97</b>
	<b>Anexos .....</b>	<b>105</b>
	<b>Anexo A: Listagem completa das variáveis de caracterização do terrorismo.....</b>	<b>105</b>
	<b>Anexo B: Caracterização dos impactos do terrorismo .....</b>	<b>108</b>
	<b>Anexo C: Caracterização dos países.....</b>	<b>110</b>



## Índice de Tabelas

---

Tabela 1: Listagem das variáveis de caracterização do terrorismo escolhidas para análise .....	37
Tabela 2: Parametrização das árvores de regressão .....	55
Tabela 3: Número de países por região .....	57
Tabela 4: Países em estudo .....	58
Tabela 5: Área geográfica de cada região .....	59
Tabela 6: Medidas descritivas das variáveis de terrorismo .....	60
Tabela 7: Matriz de correlações entre características do terrorismo .....	61
Tabela 8: Totais dos valores das variáveis dependentes por região .....	62
Tabela 9: Importância relativa dos atributos explicativos da caracterização económica dos países para a explicação de cada um dos impactos do terrorismo .....	64
Tabela 10: Importância relativa dos atributos explicativos da caracterização sociodemográfica dos países para a explicação de cada um dos impactos do terrorismo .....	67
Tabela 11: Importância relativa dos atributos explicativos da caracterização política dos países para a explicação de cada um dos impactos do terrorismo .....	70
Tabela 12: Importância relativa dos atributos explicativos da caracterização religiosa dos países para a explicação de cada um dos impactos do terrorismo .....	71
Tabela 13: Importância relativa de outros atributos explicativos de caracterização dos países para a explicação de cada um dos impactos do terrorismo .....	72
Tabela 14: Resultados de vários modelos para o número total de incidentes terroristas .....	73
Tabela 15: Importância das variáveis explicativas para o modelo do número total de incidentes terroristas .....	76
Tabela 16: Resultados de vários modelos para o número total de mortos .....	77
Tabela 17: Importância das variáveis explicativas para o modelo do número total de mortos .....	78
Tabela 18: Resultados de vários modelos para o número total de feridos .....	81
Tabela 19: Importância das variáveis explicativas para o modelo do número total de feridos .....	82
Tabela 20: Parametrização e resultados dos três modelos e número de variáveis explicativas acima de 40% de importância relativa em cada dimensão de caracterização dos países .....	84

## Índice de Figuras

---

Figura 1: Hipóteses de investigação testadas .....	32
Figura 2: Etapas que compõem o processo de DCBD .....	33
Figura 3: Fases do modelo CRISP-DM.....	33
Figura 4: Evolução do terrorismo no período em estudo .....	59
Figura 5: Árvore de regressão do número total de incidentes terroristas .....	74
Figura 6: Árvore de regressão do número total de mortos .....	79
Figura 7: Árvore de regressão do número total de feridos.....	83

## **Lista de Siglas e Acrónimos**

---

AA - Aprendizagem Automática  
CRISP-DM - *CRoss-Industry Standard Process for Data Mining*  
DCBD - Descoberta de Conhecimento em Bases de Dados  
DIB - Despesa Interna Bruta  
DNB - Despesa Nacional Bruta  
EUA - Estados Unidos da América  
FMI - Fundo Monetário Internacional  
GTD - *Global Terrorism Database*  
ONG - Organização Não-Governamental  
NATO - Organização do Tratado do Atlântico Norte  
PIB - Produto Interno Bruto  
PNB - Produto Nacional Bruto  
PPP - Paridade do Poder de Compra  
RLEE - Renda Líquida Enviada ao Exterior  
RLRE - Renda Líquida Recebida do Exterior  
RNB - Rendimento Nacional Bruto  
UE - União Europeia  
URSS - União das Repúblicas Socialistas Soviéticas

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

## **1. Introdução**

---

A investigação criminal envolve uma série de tarefas complexas ligadas à gestão do conhecimento, tais como a recolha, o tratamento e a análise da informação. Sem a capacidade de realizar uma análise eficaz e útil, o processo de recolha de informação é reduzido a um simples armazenamento de dados efectivamente não relacionados, não revelando o conhecimento neles compreendido (Petersen *et al.*, 2011).

Dantas e Júnior (2009) definem análise criminal como um processo analítico e sistemático para a produção de conhecimento, sendo realizado a partir do estabelecimento de correlações entre os conjuntos de factos ocorridos e os padrões e tendências históricas da criminalidade de um determinado local ou região, onde a mesma possa ser contextualizada.

No processo de investigação criminal é reunida uma grande quantidade de informação, proveniente de fontes pessoais, organizacionais, documentais, de fontes abertas ou fechadas, formais ou informais, internas ou externas. Esta informação só depois de sujeita a um processo de tratamento sistemático através de sistemas centralizados que garantam a sua compilação, análise, avaliação e difusão em tempo útil, é que se transforma em conhecimento da criminalidade. A análise de informação criminal pretende a transformação de dados e notícias dispersas e em bruto em informação tratada e integrada, resultando finalmente em conhecimento fundamentado e avaliado, a chamada *intelligence* (Braz, 2009).

As tecnologias de informação e da comunicação têm conhecido relevante evolução, permitindo maior capacidade de gerar e tratar automaticamente largas quantidades de dados. A análise e processamento desse grande volume de dados vai muito além das capacidades individuais dos analistas, pelo que é de grande utilidade a criação de máquinas de aprendizagem, programas que aprendam a melhorar o desempenho a partir da experiência e sistemas que auxiliem a gestão da informação (Rocha *et al.*, 2008).

O estudo proposto configura a contextualização de técnicas de *data mining*, técnicas automáticas de análise de dados particularmente adequadas quando se têm muitos dados de origem multidimensional, na área do estudo sobre o terrorismo.

Neste capítulo serão abordados o tema e motivação para este trabalho de investigação, os objectivos delineados, a contribuição que este estudo proporciona à comunidade científica e às autoridades de segurança, bem como a estrutura que será desenvolvida ao longo da dissertação.

## **1.1 Tema, Motivação e Contribuição**

Na era da globalização é clara a noção de que o fenómeno da criminalidade também se expandiu largamente, na sua dimensão e especificidade. O crime organizado com o apoio dos facilitadores resultantes da crescente economia global, com a abertura de fronteiras e tirando proveito dos fortes avanços das tecnologias de informação e comunicação, tem assinalado rápida evolução na sociedade de hoje.

Esta globalização do crime exige das autoridades de segurança uma atenta análise criminal, que para além da vertente operacional imediata da resolução do crime, devem combatê-lo por antecipação através da formulação de programas preventivos.

Perante a enorme dimensão e volume de dados a analisar, o investigador criminal necessita de sistemas que auxiliem na avaliação, tratamento e análise. Estes geram conhecimento útil, abrangendo muito mais informação e produzindo automaticamente os cruzamentos de dados ou associações, conduzindo de forma mais célere à obtenção de resultados nas investigações em curso.

O terrorismo constitui uma das formas mais agressivas de crime organizado, pelos meios que utiliza e fins que visa alcançar (Braz, 2009). Relativamente ao crime organizado em geral, o terrorismo distingue-se por agir com base em motivações de carácter político, ideológico e religioso, não tendo como fim último a obtenção de compensações financeiras (Ventura, 2003).

O presente estudo aborda o tema do terrorismo, pretendendo compreender que factores tornam mais propício o seu desenvolvimento. A forma de luta mais eficaz contra o terrorismo passa pela prevenção da radicalização e recrutamento, a qual apenas será bem sucedida conhecendo em profundidade os indicadores que caracterizam o ambiente ideal para a sua proliferação.

A análise apresentada abrange o terrorismo numa amplitude global, ocorrido ao longo de vinte anos, procurando encontrar factores potencialmente explicativos do terrorismo entre uma série de variáveis pertencentes às dimensões económica, sociodemográfica, política, religiosa e de características diversas.

A imperativa necessidade de conhecimento na área do combate ao terrorismo, que representa uma forte ameaça global, bem como a disponibilidade em fonte aberta dos dados para análise, tanto ao nível do terrorismo como ao nível dos outros indicadores globais, foram incentivos determinantes para a execução desta investigação.

Com o objectivo de desenvolver e incentivar estudos científicos e académicos nesta área foram produzidas diversas bases de dados com origem em fontes de informação abertas (*open sources* -

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

O.S.), de que são exemplo: a base de dados do terrorismo global, *Global Terrorism Database*<sup>1</sup>; a lista dos terroristas mais procurados pelo *Federal Bureau of Investigation - FBI*<sup>2</sup>; a base de dados dos incidentes terroristas mundiais da *RAND Corporation*<sup>3</sup>, a base dados sobre terrorismo internacional: atributos de eventos terroristas - *International Terrorism: Attributes of Terrorist Events, ITERATE*<sup>4</sup> - de Edward Francis Mickolus, ou a base de dados do sistema de rastreamento de incidentes mundiais do Centro de Contra-terrorismo dos EUA - *NCTC*<sup>5</sup>.

O combate ao terrorismo passa pela prevenção dos fenómenos que estão na sua génese, através da produção de informações prospectivas sobre a evolução da ameaça e identificação das causas e factores que poderão propiciar a radicalização violenta e o recrutamento de elementos.

Perante a magnitude da ameaça que o terrorismo representa, todos os estudos que desenvolvam ou acrescentem conhecimento tendo em vista a sua compreensão, são de elevada pertinência. Torna-se cada vez mais imperativo lutar contra o terrorismo pela antecipação, não só pela prevenção directa dos ataques, mas também pelo conhecimento dos indicadores que constituam ambiente ideal para a propagação da radicalização.

O terrorismo constitui uma verdadeira ameaça para os Estados e para os povos. Para fazer face a tal ameaça, a União Europeia (UE) estabeleceu uma estratégia de luta contra o terrorismo baseada em quatro pilares: prevenir, proteger, perseguir e responder. O primeiro pilar prevê que o combate ao terrorismo passa, entre outros, por continuar a investigação científica neste domínio e partilhar análises e experiências. Este estudo contribuirá igualmente para a prossecução dos objectivos no âmbito do segundo pilar, a Protecção, já que o apuramento de factores determinantes do terrorismo promove a melhor avaliação da ameaça e grau de vulnerabilidade dos países (Síntese de Legislação da UE, 2012).

O estudo realizado fornecerá à comunidade científica mais um contributo na área da análise do terrorismo, salientando-se pelo facto de fazer uso de metodologia inserida no âmbito das técnicas de *data mining*. *Data mining* refere-se à extracção ou mineração de conhecimento útil, padrões, regras e tendências de grandes quantidades de dados, previamente desconhecidas, através do uso de algoritmos específicos (Han e Kamber, 2001). No contexto do levantamento efectuado em sede de revisão da literatura não foi detectado nenhum estudo com os presentes objectivos de investigação que fizesse uso destas técnicas, pelo que se aspira a que este trabalho possa oferecer um valor acrescentado à vertente académica e à comunidade em geral.

---

<sup>1</sup> <http://www.start.umd.edu/gtd/>

<sup>2</sup> [http://www.fbi.gov/wanted/wanted\\_terrorists](http://www.fbi.gov/wanted/wanted_terrorists)

<sup>3</sup> <http://www.rand.org/nsrd/projects/terrorism-incidents.html>

<sup>4</sup> <http://library.duke.edu/data/collections/iterate.html>

<sup>5</sup> <http://wits.nctc.gov/FederalDiscoverWITS/index.do?N=0>

## **1.2 Objectivos**

Este trabalho pretende estudar o terrorismo durante o período de 1990 a 2010, fazendo uso de vários indicadores de terrorismo, económicos, sociodemográficos, políticos e religiosos que caracterizam as diversas regiões do Mundo.

Em concreto são delineados os seguintes objectivos específicos de investigação:

- 1 - Identificar factores determinantes do terrorismo;
- 2 - Concepção de um modelo explicativo e preditivo para a ocorrência e impacto de actos terroristas.

Em cumprimento dos objectivos estabelecidos espera-se poder responder a diversas questões que torneiam o tema em causa, como por exemplo: Existe relação entre terrorismo e crescimento económico?; As circunstâncias sociodemográficas têm alguma relação com o terrorismo?; A instabilidade política está correlacionada com o aumento do terrorismo?; Baixos níveis de liberdade civil e direitos políticos geram mais terrorismo?; A globalização está directamente relacionada com o terrorismo?; O terrorismo tem alguma relação com os níveis de educação?; A religião é factor determinante do terrorismo? A democracia está de alguma forma relacionada com terrorismo?; Regimes mais repressivos devem expectar números de terrorismo mais altos?, entre outras.

A investigação a efectuar permitirá igualmente fornecer mais evidência empírica no sentido do apuramento de diversas questões, como as acima referidas, para as quais se verifica haver posições contraditórias na literatura corrente. Por outro lado, pretende-se neste estudo aplicar uma técnica de análise que constitui elemento novo neste âmbito, o que se traduz numa mais valia tendo em vista a comparação e validação de trabalhos anteriores.

## **1.3 Abordagem Metodológica**

Para a concretização dos objectivos em causa é analisado o terrorismo no período de 1990 a 2010, utilizando indicadores para 173 países das várias regiões do Mundo.

Os dados de terrorismo em análise são o número de atentados, o número de mortos e o número de feridos (variáveis dependentes), variáveis que traduzem a frequência e intensidade ou impacto do terrorismo, ficando excluídos destes os danos de propriedade. Para a identificação de factores explicativos do terrorismo foi efectuada intensa pesquisa de dados em fontes abertas para o levantamento dos indicadores, tendo-se seleccionado os económicos, sociodemográficos, políticos, religiosos e outros, que constituem as variáveis independentes.



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

A compilação e validação de dados de diversas bases de dados permitiu a obtenção de um total de 3460 registos e 85 variáveis, organizados segundo o modelo de dados em painel, por país, durante um período de 20 anos.

Na análise dos dados, para além das técnicas de análise descritiva dos dados, foi seleccionada a árvore de regressão, nomeadamente o Algoritmo CART - *Classification and Regression Trees* (Breiman *et al.*, 1984), para analisar as relações entre os factores explicativos do terrorismo e os indicadores de terrorismo consagrados.

A metodologia seguida foi o CRISP-DM, definida por Chapman *et al.* (2000). Os modelos gerados foram validados pela aplicação do coeficiente de determinação ( $R^2$ ) que fornece uma medida de qualidade do modelo em relação à sua habilidade de explicar corretamente os valores da variável dependente (capacidade explicativa do modelo). Será igualmente apresentada a medida de importância relativa das variáveis independentes, ou explicativas, por forma a hierarquizar a importância dos diversos factores explicativos na predição do terrorismo.

#### **1.4 Estrutura da dissertação**

A presente dissertação está organizada em sete capítulos. O primeiro refere-se a esta introdução, onde se apresenta o tema a estudar e seu enquadramento, os objectivos de pesquisa e a abordagem metodológica adoptada.

Os capítulos 2 a 4 visam apresentar uma breve revisão da literatura existente sobre o tema. O segundo capítulo apresenta a contextualização dos conceitos e teoria de abordagem ao terrorismo. No terceiro faz-se um levantamento dos indicadores económicos, sociodemográficos, políticos e religiosos que têm relevância na caracterização dos países. O quarto capítulo relaciona-os com o tema do terrorismo.

Os capítulos 5 e 6 enquadram-se no estudo empírico. O quinto capítulo foca o conceito de *data mining* e sua relação com o processo de descoberta de conhecimento em bases de dados (DCBD). Descreve-se brevemente a metodologia utilizada no desenvolvimento de projectos com técnicas de *data mining*, o CRISP-DM, e as técnicas de *data mining* mais usuais, incidindo com maior pormenor na técnica a ser usada nesta investigação, árvore de regressão - Algoritmo CART. Ainda neste capítulo, apresentam-se os dados que servirão de base a este estudo, e as suas fontes. No sexto capítulo são apresentados os resultados e realizada a sua discussão e confrontação com a literatura existente.

Por fim, no último capítulo (o sete) são apresentadas as conclusões deste trabalho de investigação, nomeadamente os seus contributos e limitações, e são identificadas pistas para futura investigação.

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

## **Enquadramento Teórico**

### **2. Terrorismo**

---

O terrorismo representa uma ameaça à segurança, liberdade e valores dos cidadãos. As autoridades de segurança têm como objectivo responder de forma adequada e adaptada na luta contra este fenómeno, cuja abordagem passa pela prevenção, protecção, acção judicial e resposta. Especial atenção é necessária no entendimento das causas e controlo dos meios e capacidades do terrorismo. A cooperação e coordenação internacional são essenciais para assegurar a eficácia da luta contra este fenómeno transnacional (Europol, 2012).

Os terroristas e as suas redes organizadas devem ser criteriosamente monitorizados e controlados, sendo que o melhor método de combate ao terrorismo é a acção por antecipação.

*«A ira dos terroristas raramente é descontrolada. Contrariamente à crença popular e à representação dos media, o terrorismo não é nem louco nem caprichoso. Os ataques são antes tão cuidadosamente planeados, como premeditados.»<sup>6</sup>*

Esta frase do professor e investigador Hoffman (2006: 229) é particularmente curiosa, apontando para o calculismo e detalhe no planeamento de atentados.

Assim, este capítulo incide sobre a definição e avaliação de terrorismo. Apresenta as principais diferenças que marcam os diversos tipos de terrorismo, salienta igualmente a sua distinção comparativamente à globalidade do crime organizado e por último destaca alguns dos indicadores de caracterização do terrorismo.

#### **2.1 Definição de Terrorismo e Evolução Histórica do Conceito**

O terrorismo é um sistema organizado de extrema e violenta intimidação, de forma a criar instabilidade nos regimes, democráticos ou não (Tackrah, 1987). Os terroristas cometem atentados indiscriminados e imprevisíveis, direccionados a alvos específicos, designadamente forças policiais, forças armadas, multinacionais e Estados, com o objectivo de agitar e causar efervescência no quadro político-económico mundial.

Hoffman (2006) acrescenta referindo que o fenómeno terrorista é definido por via da construção e exploração do medo. Essencialmente através da violência e da intimidação, de forma a fomentar determinadas modificações políticas, sociais, étnicas e/ou religiosas.

---

<sup>6</sup> Tradução portuguesa de «*The wrath of the terrorist is rarely uncontrolled. Contrary to both popular belief and media depiction, most terrorism is neither crazed nor capricious. Rather, terrorist attacks are generally both premeditated and carefully planned.* »

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

O historiador americano Laqueur (2009) defende que o terrorismo representa a utilização que determinado grupo faz da violência, de forma a atingir certos objectivos políticos, direccionando-a contra um Governo e, menos frequentemente, contra outros grupos, classes ou partidos, provocando a sua desestabilização através da violência e intimidação. Movidos pelo ânimo de provocar sérias perturbações aos níveis social, político e económico, os grupos terroristas recorrem a sistemáticas campanhas de atentados. No entanto, o autor acrescenta que terrorismo é sinónimo de violência mas que nem toda a violência perpetrada assume carácter terrorista.

Algumas organizações norte-americanas tentam definir o terrorismo. Por exemplo, o Departamento de Estado Norte-americano<sup>7</sup> (USDS, 1999), defende que o terrorismo é uma forma premeditada de violência política, perpetrada por grupos subversivos, ou agentes clandestinos, contra alvos civis, com a finalidade de atingir vasta audiência. Já o FBI<sup>8</sup>, serviço de inteligência e investigação criminal de âmbito federal do Departamento de Justiça dos EUA, define terrorismo de acordo com o Código Regulamentar Federal em que está inserido: “*O terrorismo é o uso ilícito de força ou violência contra pessoas ou propriedades, de forma a intimidar ou coagir um Governo, a população civil ou outra qualquer dimensão da sociedade, na prossecução de objectivos políticos ou sociais*” (FBI, 2011).

A legislação portuguesa<sup>9</sup> define como terrorismo todos os actos criminosos que visem prejudicar a integridade e a independência nacionais, impedir, alterar ou subverter o funcionamento das instituições do Estado previstas na Constituição, forçar a autoridade pública a praticar um acto, a abster-se de o praticar ou a tolerar que se pratique, ou ainda intimidar certas pessoas, grupos de pessoas ou a população em geral.

Ao longo dos anos muitos linguistas, filólogos e especialistas desta área têm desenvolvido esforços no sentido de esclarecer e alcançar uma proposta de definição de terrorismo capaz de suscitar alguma unanimidade. Porém, verifica-se ampla incerteza e polémica em redor deste conceito (Ventura *et al.*, 2001). Apesar do mediatismo desta palavra, constantemente difundida pelos meios de comunicação social, para a generalidade das pessoas não existe entendimento concreto e preciso do fenómeno terrorista (Hoffman, 2006).

A palavra terrorismo surge originalmente no período histórico da Revolução Francesa de 1789, utilizada para designar o “*regime de la terreur*” de 1793-1794, marcado pelas perseguições e execuções frequentes e arbitrárias contra os opositores do novo regime. O chamado Regime de Terror foi um instrumento legal de governação, no sentido da consolidação do poder do novo governo revolucionário, procurando causar um sentimento generalizado de intimidação de forma a inibir e dissipar eventuais tentativas contra-revolucionárias (Ventura *et al.*, 2001).

---

<sup>7</sup> US Department of State, *Patterns of Global Terrorism* - 1998, Washington, DC, 1999.

<sup>8</sup> Tradução de «*the unlawful use of force and violence against persons or property to intimidate or coerce a government, the civilian population, or any segment thereof, in furtherance of political or social objectives*».

<sup>9</sup> Lei n.º52/2003, de 22 de Agosto, com as alterações introduzidas pela Rect. n.º 16/2003, de 29 de Outubro, Lei n.º25/2008, de 5 de Junho e Lei n.º 17/2011, de 3 de Maio.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Já no séc. XIX, numa época marcada por fortes convulsões políticas, deu-se o surgimento de diversos movimentos contestatários aos regimes instituídos, onde o recurso a acções de natureza terrorista encontrava plena expressão. O terrorismo afirmou-se nessa altura como o instrumento preferido dos mais fracos na luta contra os mais fortes e poderosos. No princípio do séc. XX com a tomada de posse de regimes ditatoriais na Itália, Alemanha e Rússia, surgiu outra mudança no conceito de terrorismo, voltando a estar conotado como abuso de poder e autoridade por parte dos Estados (Hoffman, 2006).

Na segunda metade do séc. XX, posteriormente à II Guerra Mundial, verificou-se o surgimento de grupos e organizações nacionalistas/separatistas numa escala global, readquirindo o conceito de terrorismo uma conotação revolucionária, aplicado a esses movimentos que travavam lutas de libertação nacionais e punham em causa o poder das potências ocidentais (Ventura *et al.*, 2001). O terrorismo, já na década de 1980 e no âmbito dos novos adventos políticos, passou também a ser usado por alguns Estados como meio de ataque contra outros Estados sem sofrerem represálias por parte deles e da comunidade internacional, através de actos de violência perpetrados por organizações encobertas, designando-se esta nova corrente de terrorismo de Estado (Ventura *et al.*, 2001).

A criação e propagação de um clima psicológico vincado de sentimentos de insegurança e medo que condicionam e interferem no normal funcionamento das sociedades é um dos grandes objectivos do terrorismo organizado, para além da violência física e imensos prejuízos materiais e humanos (Laqueur, 2009).

No final do séc. XX reproduziram-se acções violentas por parte de grupos extremistas não direccionadas ao poder político, mas sim à população em geral, através, por exemplo, do desvio de aviões ou deflagração de engenhos explosivos em aeroportos ou estações de transportes públicos, etc. Estas ocorrências propagaram-se por todo o mundo, emergindo o fanatismo religioso de forma aparentemente predominante como causa dos actos terroristas cometidos. A partir dessa altura verificou-se que os países ou Estados deixaram de estar individualmente em causa pelas suas políticas locais, pretendendo-se ao invés por em causa a ordem global (Laqueur, 2009).

Qualquer Estado de Direito Democrático está hoje vulnerável, estando comprometidos os direitos, liberdades e garantias das populações, que vivem em ambiente de medo e insegurança, causado por indivíduos que pelo seu fanatismo religioso e ideologias políticas extremistas assumiram o controlo de um poder oculto e subversivo (Miranda, 2011).

## **2.2 Tipos de Terrorismo**

Em alusão às tipologias ou padrões de terrorismo, entende-se que o terrorismo na União Europeia (UE) é essencialmente enquadrado em circunstâncias históricas, políticas, económicas e sociais de cada um dos seus Estados Membros, com especial relevo para a questão da segurança internacional, contexto em que o terrorismo islamista representa hoje a maior ameaça (Europol, 2011).

Distinguem-se então quatro tipos de terrorismo (Ventura, 2004): i) o de extrema-esquerda versus extrema-direita; ii) o anarquismo / ecoterrorismo / movimentos antiglobalização; iii) o secessionista / separatista / nacionalista; e iv) o islamista. Estes quatro tipos caracterizam-se por:

- O terrorismo de extrema-esquerda, tem motivações político-ideológicas (contexto da guerra fria, entre os EUA e a ex-URSS) relacionadas com o conflito Leste - Oeste e a oposição entre a NATO e o Pacto de Varsóvia<sup>10</sup>, pretendendo a supressão dos regimes democráticos, do capitalismo e sistemas imperialistas, e a obtenção de poder político impondo governos de inspiração Marxista-leninista e a demonstração de oposição forte e entusiástica à NATO (Organização do Tratado do Atlântico Norte) e interesses Norte-Americanos;
- O anarquismo / ecoterrorismo e a galáxia de movimentos antimilitaristas e antiglobalização, têm motivações político-ideológicas e ecológicas, ideais originários dos antigos grupos de extrema-esquerda. Manifestam-se em oposição ao capitalismo, à globalização da economia e da actividade humana em geral.
- O terrorismo secessionista / separatista / nacionalista, com motivações político-ideológicas, religiosas, históricas, étnicas e culturais, deseja a separação/secessão ou independência de determinadas áreas geográficas (supostas nações sem Estado).
- O terrorismo islamista, com motivações político-ideológicas e religiosas, objectivando a criação de regimes monolíticos governados pelos radicais ditames da sharia, a lei islâmica, a eliminação da democracia, liberdades e direitos humanos, o estabelecimento de um Califado Pan-islâmico global, o derrube de regimes não Islâmicos e também a expulsão dos não crentes das áreas sob predominante influência Islâmica.

Apesar desta segmentação, os grupos terroristas podem estar associados a mais do que uma corrente ideológica, podendo coexistir conjugações de diferentes orientações, por exemplo, de extrema-esquerda com o separatismo nacionalista, do extremismo islâmico com o independentismo (Ventura, 2004).

---

<sup>10</sup> Pacto de Varsóvia - organização que representa os países socialistas do Leste Europeu e a (ex) União Soviética, extinta em 31 de Março de 1991.

### **2.3 Terrorismo e Crime Organizado**

Relatórios e estudos publicados recentemente por organismos como a Europol, a Organização das Nações Unidas (ONU) e o Conselho da Europa identificam as organizações criminosas como entidades que através de todos os meios ao seu alcance buscam o maior lucro e poder possíveis (Vegar, 2008).

A União Europeia (UE, 1998) define organização criminosa<sup>11</sup> como uma associação estruturada, estabelecida durante um período de tempo, de duas ou mais pessoas actuando em conjunto, com vista a cometer crimes que são puníveis com pena privativa de liberdade ou uma ordem de detenção cuja duração máxima seja de, pelo menos, quatro anos, ou com pena mais grave, quer essas infrações constituam um fim em si mesmas, quer um meio de obter benefícios materiais e, se for caso disso, de influenciar indevidamente a actuação de autoridades públicas.

O terrorismo e a criminalidade organizada distinguem-se pelas distintas motivações e objectivos. Enquanto as causas do terrorismo são essencialmente políticas, a criminalidade organizada pretende retorno financeiro. O terrorismo manifesta-se em oposição ao poder político pretendendo a sua posse, e a criminalidade organizada procura apenas influenciá-lo para a obtenção de benefícios e protecção na execução das suas actividades (Ventura, 2003). Apesar dos terroristas utilizarem muitas vezes padrões de crime e métodos operatórios idênticos aos da criminalidade organizada, pretendem apenas garantir o seu próprio financiamento e apoio logístico (Ventura, 2004).

Em ambos os casos se verifica a intenção premeditada, o uso de violência extrema e ameaça de represálias, de raptos, assassinatos e extorsões, a acção na clandestinidade e em desobediência à lei, a interdição de abandono de seus membros sob pena de o mesmo se tornar fatal, a rápida e eficaz deslocalização dos seus centros de recrutamento, a facilidade de readaptação, ou até a estrutura em modelo de hierarquia (Sageman, 2004).

Existem, contudo, diversas diferenças culturais, operacionais e práticas entre eles. Os grupos terroristas fundamentam as suas acções basicamente em princípios político-religiosos e ideológicos, optam por eliminar os líderes poderosos e corruptos enquanto os grupos criminosos precisam significativamente deles e preferem usá-los. Os terroristas vão contra o regime político e tentam a sua destruição, enquanto os grupos criminosos apenas pretendem uma posição de conforto e protecção na ilegalidade. Os terroristas estão por norma mais integrados socialmente na comunidade competindo com o governo pelo apoio popular, buscam a máxima cobertura mediática, ao contrário da criminalidade organizada em geral (Sageman, 2004).

---

<sup>11</sup> Artigo 1º da Acção Comum do Conselho, adotada em 21 de Dezembro de 1998. JO de 29 de Dezembro de 1998 L351/1

Citando Braz (2009:277): «O terrorismo constitui uma das expressões mais severas de crime organizado, pela natureza dos meios que utiliza ou que potencialmente pode utilizar, e dos fins que visa alcançar».

## **2.4 Classificação do Terrorismo**

No contexto do terrorismo as motivações e ideologias exercem forte influência na escolha dos alvos e no padrão de actividades, não sendo contudo isoladamente um factor explicativo para a selecção de objectivos. A ideologia fornece um quadro de referência ético e fundamenta a legitimidade dos seus actos, de violência premeditada e politicamente motivada perpetrada contra alvos não combatentes com a intenção de influenciar uma grande audiência (Bravo e Dias, 2005).

O medo é uma característica intrínseca ao movimento terrorista. O impacto mais visível da globalização no terrorismo foi a mundialização do medo e alarme social, visto que esta é uma das grandes armas que o terrorismo moderno tem ao seu lado, senão a principal (Cardoso, 2011).

Os ataques terroristas podem ser distinguidos quanto à sua abrangência geográfica, entre domésticos (internos) e transnacionais, quando os atentados envolvem o território ou cidadãos de mais de um país, sendo contudo na grande maioria das vezes estudados em termos globais.

A capacidade de concretização dos ataques é um facto usualmente explorado na avaliação dos actos terroristas, não só quanto à sua efectividade, mas também quanto à sua eficácia, procurando apurar se os seus objectivos globais foram alcançados na totalidade. Entre grupos e incidentes terroristas existe uma grande diversidade no que se refere a objectivos e consequente dimensão dos impactos pretendidos. O mesmo se aplica aos *modi operandi* utilizados, podendo traduzir-se por exemplo em ataques suicidas, assassinatos, assaltos armados, ataques bombistas, *hijacking*, tomada de reféns, destruição de infraestruturas, e outros. Os alvos na sua grande maioria não são escolhidos aleatoriamente, podendo estar relacionados com entidades ligadas a actividades comerciais, governamentais, militares, civis, religiosas, políticas, de transportação, bens de utilidade pública, e outros. Os impactos directos de cada ataque são avaliados de acordo com o número de mortos e feridos, bem como com o levantamento dos danos de propriedade resultantes.

A ideologia política, o radicalismo religioso, os conflitos étnicos, os movimentos de independência ou separatismo, ou a defesa dos direitos dos animais e do ambiente, várias são as motivações em causa, não sendo possível estabelecer um padrão constante para as suas manifestações.



## **2.5 Sumário**

Este capítulo abordou o conceito de terrorismo, o qual foi alvo de diversas definições e perspectivas ao longo do tempo. Foram enumerados e contextualizados as diversas tipologias de terrorismo e o seu enquadramento em relação ao crime organizado. Embora o terrorismo esteja englobado no conceito de crime organizado, ele distingue-se absolutamente dos restantes tipos de criminalidade organizada devido às diferenças ideológicas que os movem e aos fins pretendidos.

O contexto ideológico-motivacional em que se inserem os movimentos terroristas influencia as suas formas de actuação e objectivos em causa, mas apesar de instituírem o seu quadro referencial ético e de legitimidade, não podem por si só constituir um factor explicativo, ou determinar um padrão constante de comportamento.

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

### **3. Caracterização Económica, Sociodemográfica, Religiosa e Política de um País**

---

A caracterização global de um país contempla várias dimensões balizadas de acordo com os parâmetros objecto de estudo, assumindo as perspectivas económica, sociodemográfica, religiosa e política. Estimar estes conceitos de forma a que traduzam aspectos da realidade para diversos fins como pesquisa, formulação, monitorização e avaliação de programas e políticas públicas exige que estes possam ser traduzidos quantitativamente. É nesse contexto que surge o conceito de indicador, como uma estatística, um facto, uma medida, uma série quantitativa de dados - indicador quantitativo - ou uma série de evidências ou apreensões sobre uma realidade - indicador qualitativo - (Mourão, 2006).

Quivy e Campenhoudt (1997) consideram o indicador como uma manifestação observável e mensurável das componentes do conceito. No quotidiano o termo indicador é muitas vezes tratado por variável e vice-versa, mas apesar de qualquer indicador incorporar variáveis o contrário não acontece visto que cada variável é o indicador-identidade de uma dimensão (Mourão, 2006).

No que refere à selecção e valoração de indicadores, não há uma regra universal, mas existe uma série de critérios para auxiliar na sua determinação e escolha. O UNRISD (Instituto de Investigação das Nações Unidas para o Desenvolvimento Social)<sup>12</sup> recomenda os seguintes itens:

- Disponibilidade de dados;
- Possibilidade de comparação intertemporal e suprarregional, com indicadores que mensurem os mesmos aspectos em diferentes espaços e tempos, possibilitando estudos históricos comparativos e análises de tendências;
- Consistência, universalidade e confiabilidade dos dados;
- Validade dos dados, reflectindo o melhor possível o contexto conceitual teórico, medindo o que é efectivamente relevante;
- Significância conceitual, avaliando se um indicador realmente representa aspectos significativos;
- Afastar sobreposições temáticas dentro de um grupo de indicadores.

Actualmente existe uma multiplicidade e diversidade de indicadores de natureza económica, política, social, cultural, ambiental, os quais podem ser utilizados isoladamente - os indicadores simples, ou associadamente - os indicadores compostos. A utilização de indicadores deve, contudo, ser acompanhada de alguma precaução, visto que apesar de consistirem num instrumento de trabalho imprescindível e indispensável a qualquer estudo empírico contêm sempre uma margem de imperfeição ou insuficiência, não traduzindo a totalidade da realidade (Capul *et al.*, 1998). Por esse motivo se torna indispensável utilizar sempre que possível um grande número de indicadores.

---

<sup>12</sup> *United Nations Research Institute for Social Development* (<http://www.unrisd.org/>)

Siedenberg (2003:53) salienta: “*um indicador é, portanto, apenas uma unidade medida parcial, substitutiva*”.

Os pontos seguintes deste capítulo procedem a uma contextualização quanto aos indicadores económicos, sociodemográficos, políticos e religiosos que têm relevância na caracterização dos países.

### **3.1 Indicadores Económicos**

O desempenho económico de um país é um factor de influência predominante no seu destino. O crescimento económico gera bem-estar social e estabilidade política, influencia as decisões a nível internacional e o equilíbrio geopolítico, determina fluxos migratórios e produz efeitos na maior parte dos fenómenos sociais. Traduz-se no acréscimo do valor dos bens e serviços produzidos pela economia, sendo regularmente medido pela taxa de crescimento real do produto interno bruto (PIB). O rendimento *per capita*, que representa um indicador do bem-estar económico de um certo país ou região ao expressar uma medida *per capita* do valor acrescentado total de bens e serviços produzidos nessa economia, também constitui uma variável de forte relevância nesta dimensão (Capul *et al.*, 1998 e Oliveira, 2009).

Por definição o PIB refere-se “*ao valor agregado de todos os bens e serviços finais produzidos dentro do território económico de um país, independentemente da nacionalidade dos proprietários das unidades produtoras desses bens e serviços*” (Sandroni, 1987:234). Constitui um indicador-chave, fundamentalmente, porque representa um papel estratégico para as agências e políticas locais, regionais e nacionais de desenvolvimento (Siedenberg, 2003).

O PIB em termos de desempenho mede o aumento da riqueza criada numa economia, país ou região, pelas unidades institucionais residentes, segundo o critério da territorialidade, enquanto analisado em termos de produto nacional bruto (PNB), mede o aumento da riqueza criada pelas entidades residentes de uma economia, país ou região, independentemente do local onde foi obtida, segundo o critério da nacionalidade (Capul *et al.*, 1998).

Importa realçar que crescimento e desenvolvimento económico não traduzem exactamente o mesmo. Crescimento económico associa-se directamente ao aumento do PIB e/ou ao rendimento *per capita*, enquanto que o desenvolvimento económico implica, para além da melhoria de indicadores económicos, também, a melhoria de indicadores sociodemográficos tais como as habilitações literárias, esperança média de vida e rendimento (Oliveira, 2009).

Existe ainda uma panóplia de outros indicadores de medição ao nível da economia, como a produtividade do trabalho, o crescimento do emprego, a taxa de inflação, o crescimento dos custos

reais laborais, a dívida pública e o equilíbrio orçamental das contas públicas, que no seu conjunto permitem caracterizar economicamente um país.

### **3.2 Indicadores Sociodemográficos e Religiosos**

Qualquer sociedade é caracterizada pela sua dinâmica populacional e comportamentos sociais. O estudo relativo às dimensões, estatísticas, estrutura e distribuição das diversas populações humanas - Demografia - as quais não são estáticas, centra-se na análise da natalidade, mortalidade, migrações, envelhecimento, educação, nacionalidade, religião e pertença étnica (Torres, 1995).

A componente da caracterização social - sociologia - foca de forma complementar o estudo científico do comportamento humano, os factos e os fenómenos sociais, e os padrões comportamentais, ideológicos e culturais, onde se inclui a religião. No entanto, a demografia está intimamente relacionada com a sociologia no sentido em que a dimensão e estrutura populacional, as suas mudanças, o enquadramento em categorias como sexo ou idade estão estritamente relacionados com ideias e comportamentos, ou seja, com o conceito de sociologia e de sociedade.

A primeira abordagem no conhecimento de uma população concretiza-se pela análise do seu ritmo de crescimento, o qual pode ser expresso em termos de ritmo anual médio de modo a permitir comparações entre unidades populacionais ou entre diferentes períodos de uma mesma população. Na análise das suas oscilações, dois elementos devem ser tomados em consideração: a mortalidade e a fecundidade (Nazareth, 2004).

O indicador esperança de vida ao nascer mede de forma directa a esperança média de vida de uma pessoa em determinado momento, mas também, indirectamente, a situação de saúde e dos meios de vida de toda uma população (Siedenberg, 2003).

Oliveira (2009) destaca uma série de indicadores de coesão social, como a desigualdade da distribuição do rendimento, a taxa de risco de pobreza antes e depois das transferências sociais, a taxa de persistência do risco de pobreza, dispersão regional do emprego, a taxa de desistência escolar e a taxa de desemprego de longa duração, todos eles importantes na caracterização de um país.

A caracterização de um país quanto à religião é geralmente medida através de indicadores de distribuição da população pelas diversas crenças ou comunidades religiosas e nível de fraccionamento que existe em determinada sociedade.

### **3.3 Indicadores Políticos**

A caracterização de um país no que se refere à política tem um papel fundamental devido à interdisciplinariedade existente entre ela e as dimensões económica e sociodemográfica. A política diz respeito a todos os procedimentos relativos à organização e administração de uma nação ou Estado, pelo que um ambiente de instabilidade política que afecte a capacidade institucional e a eficácia da execução dos programas acaba por ter impactos no quadro económico e social de um país (Meierrieks e Gries, 2012).

Os indicadores políticos surgem como instrumentos de monitorização quanto ao nível e exercício de direitos de participação política dos cidadãos, quanto à eficácia da lei, quanto ao cumprimento dos direitos humanos, igualdade entre os sexos, liberdade, estabilidade das instituições e regimes políticos - habitualmente medida de forma directa através do indicador durabilidade de regime - e à segurança dos indivíduos. Os tipos de regime em vigor determinam o grau de repressividade do governo, que condiciona os índices efectivos de liberdade civil e direitos políticos (Krueger, 2007). Em regimes mais autoritários no cumprimento e manutenção do poder existe maior probabilidade de uso de métodos de tortura e aprisionamento por motivos políticos, indicadores comumente usados para medir a intensidade da repressividade em causa. Ter a capacidade de participar politicamente e agir e manifestar ideias livremente conduz a menores sensações de revolta social (Li, 2005). Em governos democráticos índices de regulamentação e competitividade da participação política e fraccionamento político também indicam o grau de abertura política e nível de institucionalização dos regimes.

### **3.4 Sumário**

Aquando da caracterização de um país deve-se abranger várias dimensões, como a económica, sociodemográfica, religiosa e política. Para fazer uma avaliação global torna-se necessária a existência de indicadores que traduzam os aspectos da realidade, constituindo uma medida estatística, que tanto pode ser de cariz quantitativo como qualitativo.

O desempenho económico de um país, sendo um factor de influência predominante no seu destino, assume uma papel crucial, na medida em que o crescimento da economia gera bem-estar social e estabilidade política, influencia as decisões a nível internacional e o equilíbrio geopolítico, determina fluxos migratórios e produz efeitos na maior parte dos fenómenos sociais.

Os indicadores sociais, demográficos e religiosos estão intrinsecamente interligados, já que uma sociedade é caracterizada pela sua dinâmica populacional e comportamentos sociais, em que se incluem os aspectos culturais e a religião.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

A política também detém um papel fundamental no estudo do contexto de um país, visto que demanda a organização, direcção e administração das nações ou Estados, tanto nos assuntos internos como nos externos (política externa).

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo



## **4. Os Factores Económicos, Sociodemográficos, Religiosos e Políticos no Terrorismo**

---

As causas que motivam os actos terroristas e fomentam a radicalização violenta são multifacetadas, podendo constituir um conjunto de factores que juntos originam o ambiente necessário ao surgimento e crescimento do terrorismo, ou até representarem elementos precipitantes que determinam acções de curto prazo.

Este capítulo procede a um enquadramento das dimensões económica, sociodemográfica, política e religiosa no contexto do terrorismo, tendo como base uma série de estudos empíricos desenvolvidos no âmbito desta temática.

### **4.1 Factores Económicos**

O terrorismo tem sido considerado como uma ameaça ou factor de risco para o crescimento económico. A existência de forte actividade terrorista fragiliza a economia de um país criando um conceito de insegurança nos retornos financeiros do investimento. Gera custos directos pela perda de vidas humanas e danos de propriedade e destruição de infraestruturas, mas também custos indirectos pela retracção dos agentes económicos em resultado de medidas de segurança mais rígidas, aumentando os custos de transporte e restrições ao comércio (Abadie e Gardeazabal, 2008; Crain e Crain, 2006; Mirza e Verdier, 2008 e Nitsch e Schumacher, 2004). Conduz igualmente a um ambiente de ansiedade e a uma maior percepção de risco, influenciando as decisões de investimento, bem como canalizando valores nacionais acrescidos para medidas de segurança ao invés da sua utilização em qualquer outro sector que gerasse maior crescimento económico.

Diversos estudos empíricos concluíram que o terrorismo gera prejuízos económicos afectando negativamente o comércio (Nitsch e Schumacher, 2004), o turismo (Llorca-Vivero, 2008), o investimento estrangeiro directo (Enders e Sandler, 1996, Powers e Choi, 2012) e outras dimensões económicas fundamentais como a inovação tecnológica e a migração (Dreher *et al.*, 2011; Larocque *et al.*, 2010), conduzindo a quebras no crescimento económico (Abadie e Gardeazabal, 2003; Araz-Takay *et al.*, 2009; Crain e Crain, 2006; Eckstein e Tsiddon, 2004).

Contudo, condições económicas mais desfavoráveis podem também conduzir a mais actividade terrorista, por mais facilmente encontrar suporte e ambiente propício à radicalização, pela ausência de oportunidades atractivas de integração económica. Blomberg *et al.* (2004a) apresentaram um modelo económico de terrorismo em que o acesso limitado a recursos económicos (tendo como indicador o baixo crescimento económico) aumenta a probabilidade de adesão ao terrorismo com o objectivo de lutar pela melhoria global da economia, já que o retorno dos ataques bem sucedidos é bastante positivo (redistribuição dos recursos económicos).

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Problemas económicos derivados de mudanças na actividade económica, como contracções no crescimento do PIB (Blomberg *et al.*, 2004b; Krueger e Laitin, 2008), aumento do desemprego (Feldmann e Perälä, 2004; Goldstein, 2005) ou a inflação (Feldmann e Perälä 2004) têm demonstrado ter impactos no aumento do terrorismo. Boas redes de segurança social, políticas redistributivas e ajuda internacional podem atenuar as injustiças e ressentimentos e assim conduzir ao decréscimo do terrorismo (Azam e Delacroix, 2006; Azam e Thelen, 2008; Burgoon, 2006; Crenshaw *et al.*, 2007; Neumayer e Plümper, 2009; Robison *et al.*, 2006).

Estudiosos sugerem que o terrorismo pode ser combatido pelo aumento dos estímulos à actividade económica no sentido em que estes podem criar, para os potenciais terroristas e respectivas estruturas de apoio, caminhos alternativos à violência. O terrorismo pode afectar negativamente a actividade económica, e por outro lado uma forte actividade económica pode gerar menos terrorismo (Azam e Delacroix, 2006; Azam e Thelen, 2008; Blomberg e Hess, 2008a; Frey e Luechinger, 2003; Freytag *et al.*, 2011; Krueger e Laitin, 2008; Lai, 2007).

No entanto, existe também argumentação que sustenta não haver evidência de fortes ligações causais entre a economia e o terrorismo. Vários são os autores que não encontraram indícios de que o crescimento económico possa influir de forma relevante nos padrões de terrorismo e vice-versa (Basuchoudhary e Shughart, 2010; Krieger e Meierrieks, 2011; Krueger e Maleckova, 2003; Kurrild-Klitgaard *et al.*, 2006; Meierrieks e Gries, 2012; Plümper e Neumayer, 2010). Em alguns estudos não se conseguiu provar que o terrorismo tenha algum impacto na economia, nomeadamente no investimento estrangeiro directo (Enders *et al.*, 2006) e no seu desempenho global (Gaibullov e Sandler, 2009, 2011; Gries *et al.*, 2011; Tavares, 2004). Sandler e Enders (2008) salientam que a influência dos actos terroristas na economia depende também das características específicas de cada país.

A economia de um país também se pode ajustar e antecipar aos ataques terroristas, investindo mais nos sectores mais robustos, juntamente com adaptações rápidas na alocação de recursos, ou seguindo, por exemplo, a política empresarial apostando na descentralização (Frey, 2009), o que trará mais segurança aos mercados. Os efeitos do terrorismo no crescimento também dependem da dimensão da economia de um país, já que quanto maior ela for menores são os danos causados ou sentidos. Gaibullov e Sandler (2009) concluíram que o terrorismo não teve qualquer influência no crescimento económico nas grandes potências asiáticas, revelando que níveis elevados de desenvolvimento em economias de larga escala implicam menor vulnerabilidade a danos causados pelo terrorismo. Em referência aos ataques de 11 de Setembro de 2001 o mesmo se verifica, já que apesar dos elevados custos em vidas humanas e bens materiais, os seus efeitos na grande economia norte-americana foram rapidamente recuperáveis (Sandler e Enders, 2008).

Muitos outros estudos concluem que o terrorismo poderá ter origem em fracas condições políticas e institucionais (Krueger e Maleckova, 2003; Li, 2005; Piazza, 2008b, 2009), em baixos níveis de

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

integração económica internacional (Kurrild-Klitgaard *et al.*, 2006), em condições demográficas desfavoráveis, como as tensões étnicas (Basuchoudhary e Shughart, 2010; Gassebner e Luechinger, 2011), e em actividades de política externa que criam injustiças e conduzem a uma internacionalização do conflito interno (Azam e Thelen, 2010; Savun e Phillips, 2009).

Meierrieks e Gries (2012) investigaram a ligação entre o desempenho da economia e o terrorismo para 18 países da América Latina entre 1970 e 2007. Salientam o facto de que as motivações e ideologias dos grupos terroristas por todo o mundo são diversas, em harmonia com os tipos de terrorismo acima descritos, pelo que os factores favoráveis ao seu desenvolvimento podem variar de acordo com o seu contexto. No caso da América Latina, no período estudado pelos autores o perfil ideológico dos grupos terroristas caracterizava-se substancialmente pela oposição ao *status quo* económico-social, visando uma mudança que gerasse a redistribuição da riqueza pelos mais desfavorecidos.

Como indicadores de actividade terrorista usaram o número de ataques e de vítimas, o primeiro pretendendo enquadrar a frequência e o segundo à intensidade. Estudos anteriores no âmbito das causas e consequências do terrorismo têm analisado indicadores semelhantes (Crain e Crain, 2006; Gaibullov e Sandler, 2011; Tavares, 2004).

Meierrieks e Gries (2012), como indicador da situação económica, usaram a taxa de crescimento do produto interno bruto real *per capita* (PIB), já que períodos de contracção económica, com desaceleração ou quebra no crescimento, indicam condições económicas desfavoráveis, que se presume coincidir com mais terrorismo. Épocas de expansão económica, com forte crescimento, são associadas a redução do terrorismo devido ao surgimento de melhores oportunidades de estabilização económica para a população, aumentando os custos de oportunidade da violência (Blomberg *et al.*, 2004b; Meierrieks e Gries, 2012).

O grau de abertura do comércio é uma variável de caracterização económica comumente conotada e utilizada para avaliar o domínio da globalização. O nível de integração económica de um país pode, por exemplo, conduzir a ressentimentos perante a globalização tornando mais frutífero o recrutamento para grupos terroristas. Por outro lado, quando se verifica que elevados níveis de integração económica global geram ganhos económicos, pode prever-se menor terrorismo pela redução de algumas das causas socioeconómicas a ele associadas (Krieger e Meierrieks, 2011; Meierrieks e Gries, 2012). A globalização pode acabar com estruturas existentes e, assim, proporcionar novas oportunidades para grupos anteriormente excluídos (Blomberg e Hess, 2008b; Blomberg e Rosendorff, 2009; Li e Schaub, 2004). Estudos referem haver uma forte correlação negativa entre o terrorismo e a globalização, estando os países de maior abertura menos propensos ao terrorismo, tanto como perpetradores como na posição de alvos, havendo igualmente uma relação negativa entre as políticas de bem-estar e terrorismo (Burgoon, 2006; Crenshaw *et al.*, 2007).

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Meierrieks e Gries (2012) encontraram no seu estudo uma correlação negativa entre o número de ataques terroristas e o rendimento *per capita* e abertura do comércio, o que vai ao encontro de resultados empíricos anteriores que argumentam que o progresso socioeconómico reduz a actividade terrorista elevando os custos de oportunidade do terrorismo (Blomberg *et al.*, 2004a; Blomberg *et al.*, 2004b; Blomberg e Hess, 2008b; Frey e Luechinger, 2003; Krieger e Meierrieks, 2011).

Na análise conduzida por Meierrieks e Gries (2012) não se verificaram evidências de que o terrorismo tenha uma relação causa efeito no crescimento económico durante o período em observação. Isto pode dever-se à possibilidade da actividade terrorista na América Latina não ter um nível de intensidade que cause danos na economia, ou a economia poder simplesmente ser suficientemente diversificada e flexível para absorver qualquer choque económico negativo do terrorismo. Por outro lado, nos países com menor capacidade económica verificaram existir relação com o terrorismo, contrariamente aos países com maiores proveitos financeiros e economias mais fortes. Isto pode significar que baixo progresso económico juntamente com baixos níveis de desenvolvimento, que geram quebra das condições de vida e insatisfação, podem criar ambiente favorável ao recrutamento, suporte e evolução do terrorismo.

Através de diversas séries de regressões binomiais negativas, Meierrieks e Gries (2012) concluíram que o desempenho económico exerceu um efeito causal heterogéneo no terrorismo, com o crescimento a afectar apenas o terrorismo nos países com baixo ou médio rendimento. Encontraram também evidência de que no período de 1970 a 2007 os padrões de terrorismo na América Latina foram influenciados, numa correlação positiva, pela dimensão populacional, pela instabilidade política, pelas despesas militares, pelas dinâmicas da era da Guerra Fria, por baixos níveis de desenvolvimento socioeconómico e integração económica, e graus intermédios de abertura política.

Gassebner e Luechinger (2011) conduziram estudo sobre os determinantes do terrorismo usando um vasto conjunto de indicadores de dimensão económica, social, demográfica e política. No sentido de encontrar variáveis explicativas para os números do terrorismo globalmente distribuídos, empregaram a metodologia *Extreme Bound Analysis (EBA)* num conjunto de dados em painel. Os seus resultados puderam apurar que os direitos de integridade física, o liberalismo económico, a taxa de mortalidade infantil e o nível de lei e ordem têm uma correlação negativa com o terrorismo. A relação negativa entre a taxa de mortalidade infantil e o terrorismo não suporta a visão generalizada de que o terrorismo tem como causa a pobreza. Os autores sugerem que talvez não seja a falta de bens materiais que releve para a ocorrência do terrorismo, mas sim a falta de oportunidades económicas, como fundamenta a verificada relação negativa entre a liberdade da economia e o terrorismo.

Piazza (2011) avaliou a relação entre a discriminação económica de grupos minoritários e a incidência de terrorismo doméstico, através da aplicação de uma série de regressões binomiais negativas a um conjunto de 172 países, organizados sob dados em painel, durante o período de 1970 a 2006. Pôde concluir que países que permitem a persistência deste tipo de discriminação tornam-se

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

mais vulneráveis substancialmente à incidência de terrorismo. Refere ainda que, apesar da pobreza e situação económica de um país poderem afectar os números do terrorismo, a sua influência não se torna tão determinante quanto a resultante do *status* económico dos grupos minoritários que o integram, indicador este que demonstrou ter maior capacidade preditiva do que o anterior.

Freytag *et al.* (2011) testaram a hipótese que defende que o fraco desenvolvimento socioeconómico é propício ao terrorismo. Para isso, aplicaram também uma série de regressões binomiais negativas a uma amostra de 110 países, no período de 1971 a 2007, através da utilização de diversos indicadores de cariz económico, social, demográfico e político. Destacam que as circunstâncias socioeconómicas de um país afectam o comportamento dos terroristas através dos custos de oportunidade do terrorismo, que decrescem quando a situação alternativa à não-opção não se mostra muito atractiva. Os autores encontraram relevantes evidências de que as variáveis socioeconómicas realmente influenciam o terrorismo, numa correlação negativa, concluindo que os países podem beneficiar do crescimento e desenvolvimento económico no sentido da redução do terrorismo.

## **4.2 Factores Sociodemográficos e Religiosos**

Na sequência dos atentados de 11 de setembro de 2001, a sociedade e as classes políticas ocidentais tomaram a ideia de que o fundamentalismo islâmico seria o principal factor causal dos atentados terroristas. Esse pressuposto gerou nos formuladores de política externa a crença de que o terrorismo só seria vencido se uma profunda revolução nas sociedades muçulmanas fosse realizada. Foi nessa linha de pensamento que as guerras do Afeganistão e do Iraque foram amplamente apoiadas pela população norte-americana (Magalhães *et al.*, 2011).

Intelectuais norte-americanos têm defendido a transformação radical da sociedade e da cultura política do Médio Oriente, considerando que a maior ameaça reside em regiões nas quais a modernização e a liberalização ainda não chegaram. Pletka (2003) defende que uma mudança de regime no Iraque acarretaria uma transformação em todo o Médio Oriente. Salienta a precária situação económica e social nessa região onde toda a riqueza tem como origem o petróleo e se concentra nas mãos de uma pequena elite económica e política. Atribui a intolerância, obscurantismo e ódio anti-americano à falta de escolaridade e dificuldades de subsistência, que os remete para a aprendizagem nas madraças e maior frequência de mesquitas. As madraças são escolas religiosas islâmicas onde são ensinadas a língua árabe, a interpretação do Alcorão (*tafsir*), a *Sharia* (lei islâmica), as narrações do profeta Maomé e a história do islão, sendo usual a referência às madraças islâmicas como fonte do ódio e do radicalismo islâmico (Evans, 2006).

Contudo, estudos empíricos concretizados por outros investigadores desclassificam esta perspectiva dominante, desmistificando esta correlação tão directa entre religião e terrorismo. Pape (2006) estudou um grupo de 462 terroristas suicidas em todo o mundo de 1980 a 2003, determinando a sua filiação religiosa. Estudados 83% dos elementos concluiu, ao contrário do que se esperava, que a

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

maioria dos perpetradores era laica (57%) e apenas 43% eram religiosos. Testando a inclusão dos 17% de terroristas excluídos por falta de dados, e supondo que todos eles sejam religiosos, mesmo assim os números são equivalentes: 52% de religiosos e 48% laicos.

Ainda em oposição à noção dominante, Pape concluiu também que em geral os terroristas possuem nível económico e educacional mais elevado do que o grupo social a que pertencem. No que se refere à comparação do *status* socioeconómico entre laicos e religiosos, apesar de haver a crença de que os segundos teriam um estatuto inferior, verificou pelo contrário que a incidência de terroristas laicos com educação abaixo do secundário demonstrou ser maior do que a de religiosos, enquanto que a percentagem destes últimos com educação superior ao secundário é de, aproximadamente, 61%, a dos suicidas laicos é de 39% (Pape, 2006).

Krueger (2007) vai ao encontro da mesma linha de pensamento de Pape, não encontrando qualquer factor causal entre a falta de desenvolvimento económico e social e o terrorismo moderno. Mas se nas sociedades originárias dos terroristas o PIB *per capita*, o crescimento da economia e a taxa de alfabetização não denotam relação com o terrorismo, no que se refere aos países alvo de atentados terroristas concluiu que esses mesmos indicadores tendem a ser relevantes quando relacionados com a incidência de ataques. Os autores, entendendo a questão religiosa como uma questão cultural, consideram que esta não constitui um factor determinante do terrorismo, apenas um agravante do conflito.

O estudo de Gassebner e Luechinger (2011) concluiu que a urbanização (população urbana), as tensões religiosas e étnicas, a dimensão populacional, o número de militares e as despesas militares demonstram uma relação positiva com o terrorismo. As cidades dispõem de uma maior abundância de alvos e proporcionam resultados com maior impacto, tendo por isso uma maior cobertura noticiosa, atingindo uma maior audiência do que no caso de uma zona rural. A relação positiva entre as tensões religiosas e étnicas e o terrorismo veio dar robustez a estudos anteriores, que evidenciam o facto de sociedades com maior fraccionamento e tensão religiosa e étnica serem solo fértil para o aumento do terrorismo. O aumento do terrorismo também demonstrou estar relacionado com as variáveis de guerra civil, guerra internacional, manifestações e outras formas de violência política. Segundo os autores isto pode significar apenas que estas são outra via de exprimir ressentimentos e injustiças, pode significar que estas variáveis criam ambiente favorável à proliferação do terrorismo, ou que o terrorismo é usado como uma estratégia de guerra.

Meierrieks e Gries (2012), assim como outros autores, destacam a dimensão populacional como uma variável que tem demonstrado ter uma correlação positiva com o terrorismo (Burgoon, 2006; Freytag *et al.*, 2008; Gassebner e Luechinger, 2011; Krieger e Meierrieks, 2011; Krueger e Maleckova, 2003; Krueger e Laitin 2008; Lai, 2007; Piazza, 2008b; Plümper e Neumayer, 2010). Através do seu estudo empírico puderam comprovar essa influência, o que pode justificar-se pelo facto de em alguns casos serem países maiores e haver maior probabilidade de ataques por disporem de mais potenciais alvos,

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

vítimas, ou perpetradores, ou até por uma grande população poder reflectir maior captação de recursos para o policiamento em detrimento de maior investimento em contra-terrorismo, ou poder ser alvo de maiores níveis de tensão demográfica. Nos estudos de Azam e Delacroix (2006) e Azam e Thelen (2008) a dimensão populacional não demonstrou ter relação com o terrorismo.

Dreher e Fischer (2010), como resultado da sua análise, reportam que o crescimento populacional tem uma relação negativa com o terrorismo, concluindo que países com população em crescimento são alvo de menos terrorismo. Nestes casos, o crescimento populacional pode estar relacionado com equilíbrio e bem-estar socioeconómico. No que se refere à caracterização etária, Tavares (2004) e Urdal (2006) referem que populações jovens estão relacionadas com mais terrorismo, enquanto Krueger e Maleckova (2009) não encontraram nenhuma relação forte entre elevados índices de juventude e o terrorismo.

Frey *et al.* (2009) salientam que o terrorismo conduz a um aumento de insatisfação com a vida, a nível individual e colectivo, o que se inclui nos impactos indirectos do terrorismo, que assumem geralmente, em conjunto com outras dimensões como a económica e política, danos mais elevados que os directamente infligidos.

A globalização e modernização acarretam mudanças económicas, novas formas de comunicação e estilos de vida, pelo avanço das novas tecnologias, urbanização, e novas ideias ou ideologias. Estes factores podem criar sentimentos de revolta e injustiça quando associados a tensões sociodemográficas e económicas. O próprio crescimento económico quando no âmbito de modernização, gera reestruturações laborais muitas vezes com dispensa de postos de trabalho. O uso das novas tecnologias pode ser um obstáculo para muitos, e para as organizações terroristas um meio favorecedor e disseminador de ideologias e incitamento ao recrutamento (Robison *et al.*, 2006).

A educação é um factor que na literatura demonstra diversas posições antagónicas quanto à sua relação com o terrorismo. Por um lado, populações com menor grau de ensino poderão eventualmente ser mais susceptíveis ao extremismo e assim constituir um alvo fácil para o recrutamento. Por outro lado, os terroristas valorizam indivíduos de melhores níveis educacionais para o recrutamento, já que isso expecta terrorismo de melhor concretização (Bueno de Mesquita, 2005). Krueger e Maleckova (2003), Berrebi (2007) e Krueger (2008) encontram uma relação positiva entre terrorismo e educação, apesar da significância ser baixa os autores concluem que o capital humano é um preditor positivo da participação no terrorismo. Freytag *et al.* (2011) detectaram um efeito significativo entre o número de anos de estudo e o terrorismo.

### **4.3 Factores Políticos**

A estabilidade dos regimes políticos é um factor que acompanha uma redução dos números do terrorismo, enquanto que os conflitos internos constituem um preditor positivo de terrorismo,

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

obrigando a um acréscimo nas despesas militares, variável que empiricamente também vem associada a maiores níveis de terrorismo (Gassebner *et al.*, 2008; Gassebner e Luechinger, 2011; Meierrieks e Gries, 2012). Clima de guerra civil constitui igualmente ambiente ideal à implementação de campos de treino para potenciais terroristas (Campos e Gassebner, 2009).

A instabilidade política de um país cria vazios políticos, afecta a capacidade institucional e do Estado, e constitui ambiente favorável para a propagação da actividade terrorista. Meierrieks e Gries (2012) e Piazza (2008b) confirmam esta relação positiva, associando as mudanças de regime político com períodos de aumento da actividade terrorista. Vários estudos evidenciam que países com conflitos internos têm e geram mais terrorismo que outros países (Campos e Gassebner, 2009; Piazza, 2007, 2008a; Testas, 2004). No que se refere a conflitos e guerras internacionais verifica-se a mesma tendência, apesar da relação ser mais ténue (Braithwaite e Li 2007; Dreher e Fischer 2010; Koch e Cranmer 2007; Piazza 2008a; Walsh e Piazza 2010).

Baixos níveis de capacidade institucional e do Estado - que está ligada ao grau de desenvolvimento económico de um país - podem significar reduzido investimento no exercício da política, como, por exemplo, nos gastos com a segurança social, que poderiam dessa forma amenizar problemas económicos, tornando mais compreensível que os factores socioeconómicos venham a afectar a actividade terrorista nestes países (Meierrieks e Gries, 2012).

Krueger (2007) e Pape (2006) consideram que o terrorismo é causado por situações políticas, o primeiro destacando como factor indiciante os regimes que determinam o seu aparecimento, e o segundo as condições políticas em que o terrorismo se desenvolve. Krueger (2007) defende que todos os Estados de origem dos perpetradores do terror possuem baixos índices de liberdade civil e direitos políticos, pelo que o terrorismo tem origem em sociedades que estão sob regimes repressores, enfatizando as causas políticas ao invés das económicas ou culturais. Pape (2006) realça igualmente as causas de cariz político, tendo verificado que em quase todos os casos analisados havia ocupação estrangeira no território que os terroristas consideram como seu. Destaca o terrorismo como uma forma de combate, maioritariamente causado por ocupações estrangeiras, que tem como finalidade um objectivo estratégico, laico e político: a libertação nacional.

O nível de democracia de um país tem sido a variável mais usada para a avaliação do grau de repressão dos regimes políticos (Feldmann e Perälä, 2004; Meierrieks e Gries, 2012). Li (2005) realça que a possibilidade de participar politicamente com relativa facilidade torna menos provável a manifestação por meio da violência, o que implica uma relação negativa entre o terrorismo e o nível de democracia. Noutra perspectiva poder-se-ia encontrar pelo contrário uma relação positiva, tendo em conta que a repressão política e acções militares contra grupos terroristas são menos prováveis quando governos têm que respeitar as liberdades civis e ter em conta uma ampla gama de interesses políticos (Frey e Luechinger, 2003; Li, 2005; Gassebner e Luechinger, 2011). Contudo, Gassebner e Luechinger (2011) no seu estudo empírico não encontraram qualquer relação entre a democracia e o



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

terrorismo. Concluíram que a ausência de abusos no que se refere ao respeito pelos direitos humanos e a força e imparcialidade de um sistema jurídico estão associados a menos terrorismo, o que fundamenta a vertente que defende que um sistema jurídico parcial e a violação dos direitos humanos criam injustiças graves e geram violência política.

Em países de baixo nível de desenvolvimento político, com regimes repressivos, existe maior susceptibilidade de controlo das notícias veiculadas na imprensa. Nesse sentido, a cobertura da informação sobre o terrorismo tende a ser restrita e, conseqüentemente, diferente do número de incidentes relatados e posteriormente registados nas bases de dados de terrorismo. Este facto poderá eventualmente explicar uma relação positiva entre o terrorismo e democracia (Li, 2005; Drakos e Gofas, 2006).

Do ponto de vista da ecologia política, tendo em conta factores de natureza geopolítica, pode-se encontrar uma relação positiva no que se refere à importância da existência de recursos naturais (minerais) estratégicos no surgimento de conflitos armados, onde se inclui o terrorismo (Le Billon, 2001; Tavares, 2004; Bravo e Dias, 2005; Sambanis, 2008). Em sintonia com o nível de dependência da sociedade relativamente a este tipo de recursos, é previsível a procura de maior capacidade e poder através do seu domínio (Ehrlich e Liu, 2002).

Bravo e Dias (2005) analisaram o terrorismo no período de 1997 a 2004, utilizando indicadores económicos, demográficos, políticos e outros indicadores geopolíticos, para 121 países das várias regiões do globo. Apresentam um modelo de natureza exploratória de explicação do terrorismo, tendo como variável dependente o número de ataques terroristas, utilizando o método dos mínimos quadrados ordinários (*Ordinary Least Squares*). Os seus resultados indicam que o maior número de ataques está associado a países com fracos níveis de desenvolvimento, com regimes não-democráticos, baixos níveis de literacia, menos dependência do comércio internacional, existência de reservas minerais e posição geográfica relevante.

#### **4.4 Sumário**

A revisão de estudos empíricos concretizada permite desde já concluir que o terrorismo é um fenómeno demasiado complexo para ser reduzido a apenas uma causa, ou mesmo a um conjunto de causas amplamente generalizáveis, visto que as dinâmicas específicas de cada país ou mesmo região, ou período temporal, podem gerar diferentes conclusões. Contudo, torna-se possível reunir um conjunto de indicadores que podem constituir um ambiente favorável ao aparecimento e fomento do terrorismo.

A grande maioria dos resultados permite antever que a ocorrência de ataques terroristas se torna mais provável em países de regimes políticos não democráticos, altamente povoados e de elevada instabilidade política e económica. A existência de conflitos internos como guerras civis ou mesmo

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

conflitos internacionais constitui ambiente favorável ao desenvolvimento, motivação e recrutamento para o terrorismo. O proeminente fraccionamento étnico e linguístico, e as conseqüentes tensões sociais, também demonstraram ter relevante significado na explicação do terrorismo. Por outro lado, os níveis de literacia, a posição geográfica ou o fraccionamento religioso foram indicadores que na maioria dos estudos não demonstraram fortes evidências de relação com o terrorismo.

O *status* económico de um país é um factor que influencia predominante o seu destino, gerando bem-estar social e estabilidade política quando em crescimento, influenciando nas decisões a nível internacional e no equilíbrio geopolítico, determinando fluxos migratórios e produzindo efeitos na maior parte dos fenómenos sociais. Da mesma forma, os regimes e actuação ao nível da política também exercem forte preponderância nas condições económicas e sociais de um país. Esta simbiose conduz à necessidade de uma abordagem generalista, não podendo este conjunto de factores ser observados de forma isolada, mas sim através de uma visão holística.

## ***Estudo Empírico***

### ***5. Metodologia***

---

A presente investigação tem como objectivos a identificação de factores determinantes do terrorismo e a concepção de um modelo explicativo e preditivo para a ocorrência e impacto de actos terroristas.

As metodologias usadas pelos autores referidos na revisão da literatura consistem maioritariamente em técnicas de regressão. De forma geral, baseiam-se em mínimos quadrados ordinários (*OLS*), variáveis instrumentais (*IV*), regressões binomiais negativas (*NBR*), processos *Markov* (*MR*), análise gravimétrica (*GAT*), teste do chi-quadrado (*CHI*), regressão de efeitos mistos *Poisson* (*MEP*), regressão logística binomial (*BLR*) e regressão logística multinomial (*MLG*). Destaca-se o facto de nem todos os estudos se referirem a amostras idênticas, examinando muitas vezes conjuntos de países de regiões diferentes, de períodos não coincidentes, através de indicadores que variam de caso para caso, o que pode, em certa medida, conduzir à existência de diferentes conclusões.

Assim, procedeu-se à análise do fenómeno do terrorismo através da utilização de outras técnicas, no âmbito do *Data Mining*, que é um processo que usa várias técnicas estatísticas e de reconhecimento de padrões para descobrir paradigmas e relações nos dados ou variáveis registadas. No caso da análise criminal, destina-se à detecção e prevenção da actividade criminal (Mena, 2003).

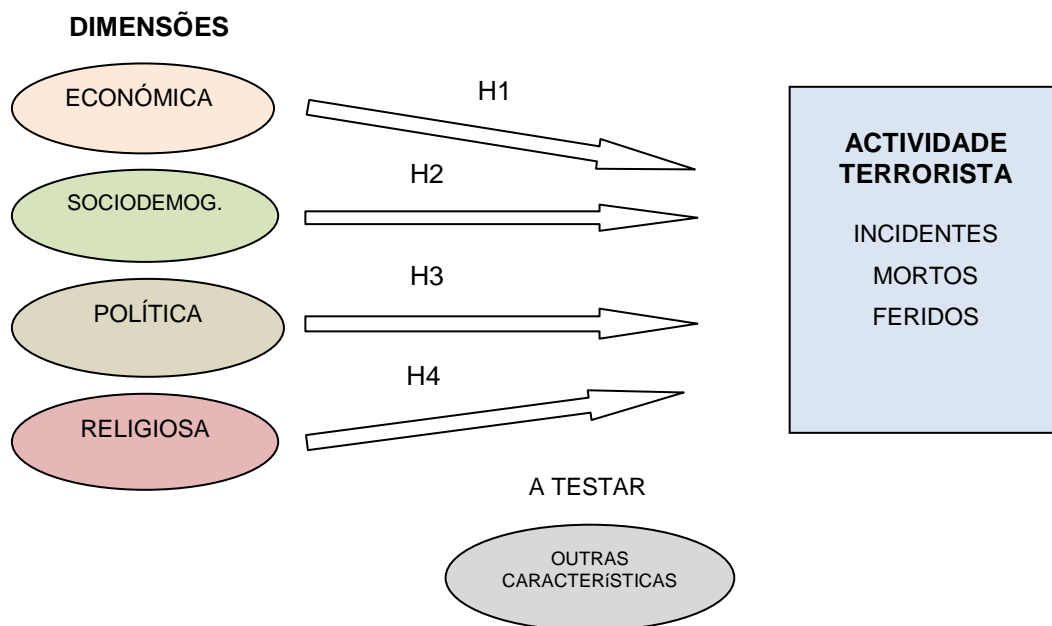
Apresentam-se em seguida as hipóteses de investigação e breve descrição do processo de descoberta de conhecimento em bases de dados, que se refere a todas as etapas que permitem transformar dados em bruto em conhecimento útil. Abordam-se igualmente a metodologia CRISP-DM, as diferentes técnicas usadas para a mineração de dados, o conjunto das variáveis, nas suas diferentes dimensões, a ser incluídas no presente estudo e a metodologia desenvolvida.

## 5.1 Hipóteses de Investigação

Na investigação desenvolvida testou-se um conjunto de quatro hipóteses relacionadas com os possíveis factores explicativos da actividade terrorista, que são:

- H1 - Os factores económicos são os que melhor explicam a actividade terrorista.
- H2 - Os factores sociodemográficos são os que melhor explicam a actividade terrorista.
- H3 - Os factores políticos são os que melhor explicam a actividade terrorista.
- H4 - Os factores religiosos são os que melhor explicam a actividade terrorista.

Figura 1: Hipóteses de investigação testadas



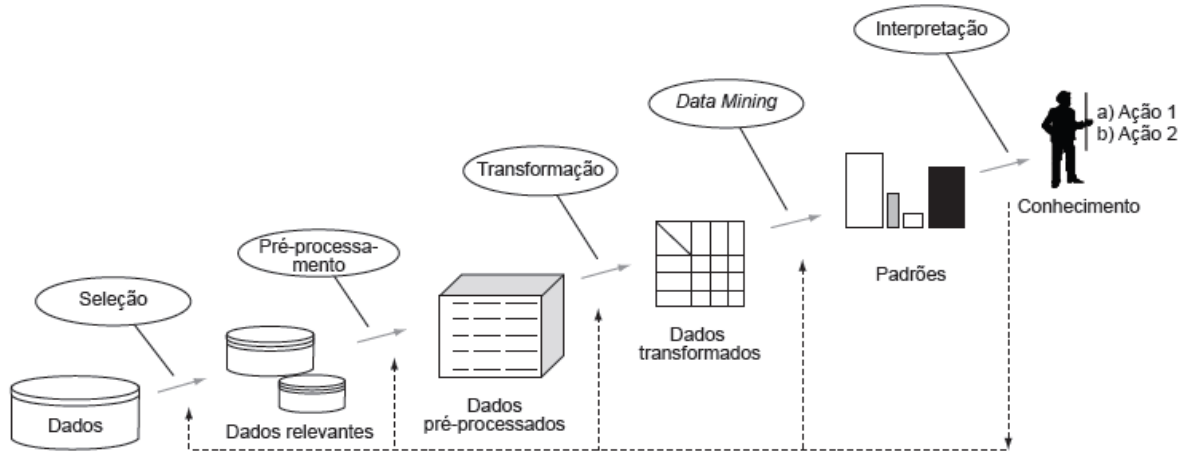
## 5.2 Descoberta de Conhecimento em Bases de Dados

A Descoberta de Conhecimento em Bases de Dados (DCBD) refere-se a todo o processo que permite transformar dados em bruto em conhecimento de alto nível (Rocha *et al.*, 2008). A DCBD refere-se ao sistema global de descoberta de conhecimento útil em bases de dados, sendo o *Data Mining*, ou mineração de dados, apenas uma etapa particular deste processo, fazendo a aplicação de algoritmos específicos para a extracção de padrões dos dados (Fayyad *et al.*, 1996).

A DCBD é então constituída pelas seguintes etapas (Fayyad *et al.*, 1996) representadas na Figura 2:

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

**Figura 2: Etapas que compõem o processo de DCBD**



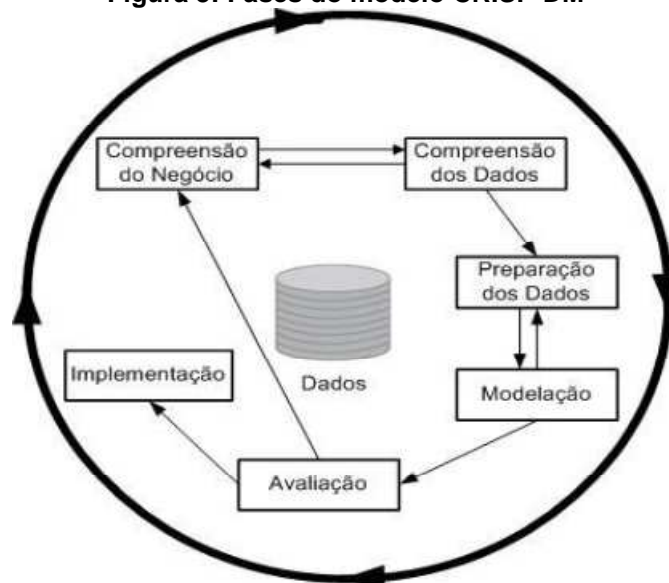
*Fonte: Adaptado de Fayyad et al. (1996: 41)*

Han e Kamber (2001) definem o *data mining* como a extracção de conhecimento de uma grande quantidade de dados, não considerando contudo esta nomenclatura adequada pois não dá o devido realce ao conceito conhecimento. Apesar do *data mining* ser uma etapa do processo de descoberta de conhecimento, habitualmente este termo tem sido mais mediatizado do que o de DCBD, tendendo a referir-se ao conjunto de todo o processo.

Uma das metodologias mais utilizadas na condução de projectos de descoberta de conhecimento é a CRISP-DM (*Cross- Industry Standard Process for Data Mining*), que é o resultado de estudo efectuado por vários especialistas no sentido da definição das melhores práticas e de um modelo processual *standard* para a sistematização da descoberta de conhecimento (Chapman et al., 2000).

Esta metodologia consiste em seis fases, como mostra a Figura 3:

**Figura 3: Fases do modelo CRISP-DM**



*Fonte: Adaptado de Chapman et al. (2000: 10)*

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

- Compreensão do problema: que visa o entendimento dos objectivos do projecto e dos requisitos do problema, definindo a questão em causa e o planeamento preliminar para o alcance dos objectivos;
- Compreensão dos dados: consiste na recolha dos dados e em actividades que visam o conhecimento dos dados, identificação dos problemas de qualidade e detecção de subconjuntos interessantes para os quais se possam formular hipóteses;
- Preparação dos dados: abarca todas as tarefas de preparação dos dados para a construção de uma base de dados final, as quais incluem selecção de tabelas, registos e atributos, bem como a transformação e limpeza dos dados para a fase de modelação;
- Modelação: aplicação de várias técnicas de modelação aos dados e ajuste dos seus parâmetros para valores óptimos, pretendendo obter um ou mais modelos que representem o conhecimento;
- Avaliação: análise dos modelos obtidos para verificar se atingem os objectivos de investigação;
- Implementação: aplicação do modelo delineado pelas equipas de investigação.

No decurso dos processos de investigação verifica-se muitas vezes a necessidade de recuar para fases anteriores, sendo que o resultado de cada fase determina a próxima fase a executar. Existem várias técnicas para o mesmo tipo de problema de mineração de dados, tendo algumas requisitos específicos na formação de dados, pelo que retornar à fase de preparação de dados é frequentemente necessário (Chapman *et al.*, 2000).

### **5.3 Fontes de Dados e Dados**

Os pontos seguintes introduzem as variáveis que entram no presente estudo e suas fontes. O primeiro ponto refere-se aos dados de terrorismo, dos quais são extraídas três variáveis, que constituem as variáveis dependentes: o número de ataques, o número de mortos e o número de feridos, e que medem a frequência e intensidade do terrorismo.

Os restantes cinco pontos focam as variáveis independentes, de cariz multidimensional, que constituem, respectivamente, indicadores de caracterização económica, sociodemográfica, política, religiosa, e outros. Para evitar os inconvenientes associados à retrocausalidade - permitindo que um efeito preceda a sua causa, ou exista inversão de causalidade - as variáveis explicativas entraram em análise desfasadas num período de um ano ( $n-1$ ).

#### **5.3.1 Dados de Terrorismo**

No universo das bases de dados de incidentes terroristas disponíveis, a escolha para este estudo recaiu sobre a *Global Terrorism Database (GTD)*<sup>13</sup> disponibilizada pelo Consórcio Nacional para o

---

<sup>13</sup> <http://www.start.umd.edu/gtd/>

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Estudo do Terrorismo e Resposta ao Terrorismo, conhecido por START<sup>14</sup>. Trata-se de um centro de pesquisa sediado na Universidade de Maryland - EUA - ligado ao estudo científico das causas e consequências humanas do terrorismo nos EUA e resto do mundo.

Para além de apoiar directamente um conjunto de investigadores em mais de 50 instituições académicas e de pesquisa responsáveis por trabalhos de investigação em várias áreas fundamentais relativas ao terrorismo, o START tem também um papel primordial na disseminação do conhecimento sobre terrorismo e promoção de diverso material e programas especificamente criados para formadores e estudantes, pretendendo o aprofundamento do conhecimento do terrorismo e sua dinâmica. Esta instituição encontra-se inserida num conjunto de centros de excelência apoiados pela Direcção de Ciência e Tecnologia do Departamento de Segurança Interna dos EUA.

A GTD é uma base de dados de fonte aberta que disponibiliza informação sobre incidentes terroristas de todo o mundo no período de 1970 a 2010, concebida em ambiente favorável às actuais técnicas de análise quantitativas usadas pelas ciências sociais e computacionais. Destaca-se por incluir dados de terrorismo nacional e internacional abrangendo atualmente mais de 98.000 casos. Para cada ocorrência a GTD contém informação sobre a data, local, armas usadas, táctica usada, natureza e identificação do alvo, sucesso da concretização, reivindicação, número de feridos e mortos, nível de danos e, sempre que identificados, os perpetradores.

A sua construção teve início em 2001 quando investigadores da Universidade de Maryland tiveram acesso a uma ampla base de dados originária do *Pinkerton Global Intelligence Services (PGIS)*, a qual tinha na altura a característica única de contemplar dados de terrorismo doméstico mas também transnacional. Durante o processo de transferência de todo o material e cópias físicas do PGIS para a Universidade de Maryland foi descoberto que os dados referentes ao ano de 1993 tinham sido perdidos numa mudança anterior de instalações. Por não se ter mostrado possível a sua recuperação, avançaram para a tentativa de reconstrução dos mesmos através dos relatórios originais da PGIS e de novas pesquisas em fontes abertas. Contudo, não se tornou exequível restaurar os dados na sua totalidade apesar dos intensos esforços nesse sentido, o que se pressupõe dever ao facto de muitas das notícias que serviram de base terem sido veiculadas em jornais locais e regionais que depois de algum tempo decorrido não estão mais disponíveis (LaFree, 2010).

Estando a informação relativa ao ano de 1993 incompleta, e pelo facto de na análise descritiva dos dados se verificar uma relevante inconsistência relativamente aos anos anteriores e seguintes, optou-se por não contemplar este ano na investigação a concretizar.

Na produção e manutenção da GTD a selecção de incidentes a incluir segue a definição de três requisitos cumulativos, os actos em causa têm que ser de carácter intencional e premeditado, devem implicar algum nível de violência ou ameaça de violência a pessoas e bens, e por último não podem

---

<sup>14</sup> <http://www.start.umd.edu/start/>

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

constituir demonstrações de terrorismo de Estado. Para além destes, outros critérios são tidos em linha de conta: os actos devem visar a obtenção de um objectivo político, económico, religioso ou social, não se referindo contudo a proveitos financeiros, mas sim à reivindicação de mudanças profundas e sistémicas na economia do país, devem igualmente haver evidências da intenção de coagir, intimidar ou transmitir alguma outra mensagem a um público maior do que as vítimas imediatas, e a área de ocorrência deverá estar fora do contexto de actividades legítimas de guerra.

A GTD encontra-se disponível na sua totalidade via *interface online* (START, 2012), de fácil e imediato acesso, com o objectivo de constituir elemento facilitador conducente ao aumento do estudo, entendimento e conseqüente combate ao terrorismo. Este factor, conjuntamente com todas as suas características específicas e peculiares, foi relevante para a escolha desta base de dados.

Toda a informação constante desta base de dados foi recolhida de materiais públicos e disponíveis em fontes abertas, como notícias digitais em arquivo, bases de dados já existentes, livros, jornais e em documentos de carácter legal. Esta recolha e permanente actualização é supervisionada por um painel consultivo de diversos peritos de investigação de terrorismo, os quais intentam pela constante sustentação da fiabilidade dos dados.

A base de dados em análise, que constitui a origem dos dados de terrorismo do presente estudo, serviu de base à criação de um novo ficheiro que inclui, para além destes, dados relativos a diversos indicadores de cariz económico, sociodemográfico, político e religioso. O presente estudo irá desenvolver-se no âmbito de uma análise em série temporal, avaliando diversas variáveis ao longo do tempo. Nesse sentido construiu-se um novo ficheiro organizado segundo o modelo de dados em painel, por país, durante um período de vinte anos decorrido entre 1990 a 2010.

Estando a GTD organizada por incidente, foi necessário o processamento da informação de forma a que focasse os atentados terroristas inseridos num contexto anual, contínuo e em referência ao país de ocorrência. A base de dados nascente foi desenhada por país, nos quais se inseriu o período acima referido em ordem sequencial anual ascendente. A informação antes definida para cada atentado passa, assim, a entrar para um global anual conjunto de cada país.

Os dados recolhidos do ficheiro original da GTD tiveram que sofrer algumas adaptações, nomeadamente no que se refere aos países nela constantes. Ao longo dos anos, visto que contém dados de incidentes desde 1970, as circunstâncias políticas dos diversos países mudaram substancialmente. Em muitos casos alguns países deixaram de existir e outros foram criados, tendo os autores da base decidido manter as designações originais dos países aquando da ocorrência dos incidentes terroristas. Temos, por exemplo, o caso da Eritreia (independente desde 24 de Maio de 1993), da Alemanha (a qual se unificou em 03 de Outubro de 1990), o desmembramento da Checoslováquia (em 01 de Janeiro de 1993), a dissolução da União das Repúblicas Socialistas



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Soviéticas (entre 21 de Agosto de 1991 e 05 de Janeiro de 1992), bem como a fragmentação da Iugoslávia (que decorreu também faseadamente desde 1991 até 2008).

Contudo, estas diferentes designações para um mesmo local impossibilitariam o estudo a desenvolver, visto que se pretende a sua contextualização no quadro geográfico actual e também o seu ajustamento com os restantes dados recolhidos que se pretendem auxiliar à análise explicativa. Assim, foi necessário examinar todos os países constantes da base e proceder à sua harmonização, a qual foi minuciosamente determinada através do estudo da evolução político-geográfica de cada país. Diversos países tiveram que ser eliminados por se tornar impossível estabelecer a sua localização à luz actual após terem ocorrido decomposições.

Seguidamente foi necessário adaptar os dados da GTD aos dados reunidos ao nível dos indicadores económicos, sociodemográficos, políticos e religiosos, os quais foram recolhidos de diversas fontes. Vários procedimentos tiveram que ser tidos em linha de conta como a confrontação ao nível da tradução e nomenclatura dos países, que variava com frequência nos diversos ficheiros, a normalização de formato, bem como a junção e acerto de todos os dados por país. Após todas estas tarefas que geraram a base de dados final, foram eliminados mais alguns países, tendo-se optado por manter os países para os quais se conseguiu reunir maior quantidade de dados disponíveis. Resulta, desta forma, um total de 173 países que serão pormenorizadamente descritos no ponto 6.1, referente à caracterização da amostra (Tabela 4).

As variáveis de terrorismo apuradas perfazem um total de 45<sup>15</sup>. Para a concretização dos presentes objectivos de investigação, deste conjunto foram escolhidas três variáveis dependentes: o número total de incidentes terroristas, o número total de mortos e o número total de feridos, conforme discriminado na Tabela 1.

**Tabela 1: Listagem das variáveis de caracterização do terrorismo escolhidas para análise**

	CÓDIGO	DESIGNAÇÃO	INGLÊS ORIGINAL	NATUREZA	UNIDADE DE MEDIDA	DESCRIÇÃO
2	IncTerror	Nº total incidentes terroristas	-	Quantitativa	Número soma total var. discreta	Engloba o número total de incidentes terroristas ocorridos no país nesse ano
4 3	NrMortos	Nº total de mortos	<i>Total Number of Fatalities</i>	Quantitativa	Número soma total var. discreta	Este campo armazena o número total de mortes confirmadas para todos os incidentes ocorridos nesse ano nesse país. O número inclui todas as vítimas e atacantes que morreram como resultado directo do incidente.
4 4	NrFeridos	Nº total de feridos	<i>Total Number of Injured</i>	Quantitativa	Número soma total var. discreta	Refere-se ao número total de feridos não fatais confirmados.

<sup>15</sup> No Anexo A são apresentadas as outras variáveis que constituem a GTD (2012), mas que não foram incluídas neste estudo.

### 5.3.2 Dados de Caracterização dos Países

Os países podem ser caracterizados segundo diferentes perspectivas, nomeadamente económica, social, política e religiosa<sup>16</sup>.

- **Caracterização Económica**

A análise da situação económica de um país possibilita conhecer a relação existente entre o terrorismo e a economia. Avaliará se efectivamente condições socioeconómicas desfavoráveis, resultantes de períodos de contracção económica, propiciam o aumento do terrorismo e se o crescimento económico poderá ser associado a uma redução do terrorismo. Com o objectivo de mensurar a actividade económica é utilizada a variável crescimento do PIB (proposta por: Blomberg *et al.*, 2004b; Campos e Gassebner, 2009; Drakos e Gofas, 2006; Dreher e Fischer, 2010; Dreher e Gassebner, 2008; Feldmann e Perälä, 2004; Freytag *et al.*, 2011; Gassebner e Luechinger, 2011; Krueger e Laitin, 2008; Kurrild-Klitgaard *et al.*, 2006; Li, 2005; Meierrieks e Gries, 2012; Piazza, 2006; Powers e Choi, 2012; Tavares, 2004; Urdal, 2006).

Atendendo ao facto de um PIB elevado poder significar apenas um país de grande dimensão ou dinâmica económica, sendo contudo o seu PIB *per capita* baixo devido à existência de uma grande população, para analisar a qualidade de vida de um país são aplicados os indicadores PIB *per capita* (proposto por: Abadie, 2006; Azam e Delacroix, 2006; Azam e Thelen, 2008, 2010; Basuchoudhary e Shughart, 2010; Berman e Laitin, 2008; Blomberg e Hess, 2008a, 2008b; Blomberg e Rosendorff, 2009; Blomberg *et al.*, 2004a, 2004b; Braithwaite e Li, 2007; Campos e Gassebner, 2009; Crenshaw *et al.*, 2007; Dreher e Fischer, 2010, 2011; Dreher e Gassebner, 2008; Freytag *et al.*, 2011; Gassebner e Luechinger, 2011; Goldstein, 2005; Kis-Katos *et al.*, 2011; Krueger e Laitin, 2008; Krueger e Maleckova, 2003; Kurrild-Klitgaard *et al.*, 2006; Lai, 2007; Li, 2005; Li e Schaub, 2004; Meierrieks e Gries, 2012; Neumayer e Plümper, 2009; Piazza, 2007; Plümper e Neumayer, 2010; Powers e Choi, 2012; Robison *et al.*, 2006; Sambanis, 2008; Savun e Phillips, 2009; Tavares, 2004; Testas, 2004; Urdal, 2006; Walsh e Piazza, 2010) e o crescimento do PIB *per capita* (proposto por: Basuchoudhary e Shughart, 2010; Gries *et al.*, 2011; Kis-Katos *et al.*, 2011; Meierrieks e Gries, 2012). Um valor de rendimento *per capita* elevado poderá significar menos terrorismo, devido a uma maior integração económica dos indivíduos e estabilidade social, o que, desta forma, gera um aumento dos custos de oportunidade do terrorismo, tornando-se numa opção menos atractiva.

Outro indicador vulgarmente usado para o levantamento do nível de desenvolvimento económico de um país é o Rendimento Nacional Bruto (RNB). Por muitos anos designado de Produto Nacional Bruto (PNB), o actual RNB consiste numa expressão monetária dos bens e serviços produzidos por factores de produção nacionais, independentemente do território económico. Distingue-se do PIB pelos valores da Renda Líquida Enviada ao, ou Recebida do, Exterior - RLEE ou RLRE, cujos valores

---

<sup>16</sup> No anexo C são apresentadas tabelas com todas as variáveis explicativas recolhidas para estudo.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

não são considerados nos cálculos do PIB. Esta Renda representa a diferença entre recursos enviados ao exterior (pagamento de factores de produção internacionais alocados em território nacional) e os recursos recebidos do exterior a partir de factores de produção que, sendo nacionais, se encontram em actividade noutros países. Em geral, os países desenvolvidos possuem um RNB maior que o PIB, ao contrário do que acontece com os países em desenvolvimento. No presente estudo são considerados o crescimento do RNB em valores absolutos e *per capita*, anteriormente abordados por Piazza (2011).

A inflação é outra variável a ter em conta, tanto ao nível dos preços de consumo como ao nível do deflator do PIB. Este indicador assume elevada importância devido à sua forte relação com a economia de um país, podendo causar impactos crescentemente negativos numa economia. A inflação medida pelo índice de preços ao consumidor reflecte a variação percentual anual no custo médio para o consumidor de aquisição de um cabaz de bens e serviços que podem ser fixados ou alterados regularmente. Consiste na queda do valor de mercado ou poder de compra do dinheiro, referindo-se popularmente ao aumento geral dos preços. Elevadas taxas de inflação podem estar positivamente correlacionadas com o terrorismo, porquanto se verificam significativas evidências de que a hiperinflação está directamente relacionada com mudanças de regime e instabilidade política, estando esta última acompanhada na maioria das vezes de violência política (Piazza, J., 2006). A inflação medida pela taxa de crescimento anual do deflator implícito do PIB mostra a taxa de variação dos preços na economia como um todo. O deflator do PIB é uma estatística simples calculada pela divisão do PIB nominal pelo PIB real multiplicados por cem. Reflecte as mudanças que ocorrem nos preços do mercado e, portanto, é usado para controlar o nível médio de preços em dada economia. O cálculo da taxa de inflação de um determinado ano leva em consideração, geralmente, o deflator do PIB deste ano em relação à mesma estatística referente ao ano anterior (proposta por: Feldmann e Perälä, 2004; Gassebner e Luechinger, 2011; Piazza, 2006).

A crescente incerteza gerada pelas constantes variações dos preços e elevada inflação pode desestimular o investimento e a poupança. No comércio com o exterior se a taxa de inflação for maior do que a praticada em outros países, uma tarifa fixa de comércio será encoberta pelo enfraquecimento da posição do país na balança comercial. A variável comércio abordada neste estudo refere-se ao grau de abertura comercial de um país, consistindo na soma das exportações e importações de bens e serviços medidos como proporção do PIB. Traduzirá parcialmente a situação socioeconómica existente, estando positivamente relacionada com elevado crescimento e desenvolvimento. Largamente usado por muitos autores, este indicador tem sido relevante na avaliação do grau de globalização e integração económica de um país (Blomberg e Hess, 2008a, 2008b; Blomberg e Rosendorff, 2009; Braithwaite e Li, 2007; Bravo e Dias, 2005; Burgoon, 2006; Campos e Gassebner, 2009; Crenshaw *et al.*, 2007; Drakos e Gofas, 2006; Freytag *et al.*, 2011; Gassebner e Luechinger, 2011; Gries *et al.*, 2011; Kis-Katos *et al.*, 2011; Kurrild-Klitgaard *et al.*, 2006; Li e Schaub, 2004; Li, 2005; Meierrieks e Gries, 2012; Powers e Choi, 2012; Robison *et al.*, 2006).

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

O investimento directo estrangeiro também é um indicador a incluir neste estudo, consistindo na aquisição de interesses efectivos em empresas que operem fora da economia do investidor. Compreende a existência de uma empresa matriz e uma filial estrangeira, as quais, em conjunto, formam uma empresa multinacional, devendo à matriz ter o controlo sobre a sua filial. Em economias mais instáveis é uma variável que tem denotado uma relação negativa com o terrorismo. Anteriormente já foi abordada por: Gassebner e Luechinger (2011), Li e Schaub (2004), Powers e Choi (2012) e Robison *et al.* (2006).

A variável despesa geral de consumo final do Governo, anteriormente estudada por: Burgoon (2006), Crenshaw *et al.* (2007), Gassebner e Luechinger (2011) e Robison *et al.* (2006), inclui todas as despesas correntes do governo para a compra de bens e serviços (incluindo a remuneração dos empregados). Também inclui a maioria dos gastos com defesa e segurança nacional, mas exclui os gastos militares que fazem parte da formação de capital do governo.

O indicador referente à utilização de crédito do Fundo Monetário Internacional (FMI), reporta o uso de assistência técnica e financeira do FMI, que se trata de uma organização internacional que pretende assegurar o bom funcionamento do sistema financeiro mundial pela monitorização das taxas de câmbio e da balança de pagamentos. Constituída actualmente por 188 países<sup>17</sup>, tem como objectivo básico zelar pela estabilidade do sistema monetário internacional, através da promoção da cooperação e da consulta de assuntos monetários entre os seus membros. Azam e Delacroix (2006), Azam e Thelen (2008, 2010), Campos e Gassebner (2009), Gassebner e Luechinger (2011) e Neumayer e Plümpfer (2009) abordaram empiricamente o tema da ajuda externa internacional, que ao fomentar o equilíbrio financeiro de um país proporcionaria maior estabilidade económica e daí contribuiria para a diminuição do terrorismo. A redução da pobreza é amplamente considerada como o principal objectivo da ajuda externa, favorecendo igualmente a harmonia política e uma maior capacidade do governo.

A existência e o acesso a recursos minerais e energéticos têm sido considerados factores relevantes no posicionamento geopolítico e geoestratégico de um país. O nível de dependência da sociedade relativamente a este tipo de recursos, torna previsível a busca de maior capacidade e poder através do seu domínio, constituindo um alvo de elevado interesse para ataque, do qual poderão advir avultados danos económicos. Nesse sentido, e em sintonia com Sambanis (2008) e Gassebner e Luechinger (2011), serão incluídos neste estudo duas variáveis relativas à exportação e importação de combustíveis<sup>18</sup> - em percentagem das mercadorias comercializadas.

A variável despesa de consumo final, também abordada por Freytag *et al.* (2011), consiste na despesa efectuada pelas unidades institucionais residentes com bens ou serviços utilizados para a satisfação directa de necessidades individuais ou no âmbito da administração pública. Tem como

<sup>17</sup> <http://www.imf.org/external/about/whatwedo.htm>

<sup>18</sup> Por combustíveis entende os incluídos na secção 3 da *Standard International Trade Classification (SITC)* <http://comtrade.un.org/pb/CommodityPagesNew.aspx?y=2008&v=true>

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

referência a percentagem do PIB, enquadrando uma estimativa da discrepância estatística na utilização dos recursos em relação ao fornecimento de recursos. Também se aplica como indicador das condições socioeconómicas de um país, sendo altos níveis de consumo final vulgarmente conotados com menor intervenção estatal na economia, denotando eventualmente um mais elevado grau de satisfação.

A Despesa Nacional Bruta (DNB) ou Despesa Interna Bruta (DIB) entende-se pelo total dos gastos da renda pela economia num dado período. Geralmente compreende o total dos gastos em bens e serviços de consumo, em bens de capital (investimento), os gastos do governo e as exportações menos as importações. Estes investimentos em bens de capital são abordados de forma individual com a inclusão da variável Formação Bruta de Capital, e sua taxa de crescimento, que é medida pelo valor total da formação bruta de capital fixo, variação de existências e aquisições líquidas de cessões de objectos de valor para uma unidade ou sector. O indicador poupanças internas brutas também será considerado, consistindo na diferença entre o PIB e a despesa de consumo final.

As variáveis acima descritas foram extraídas do catálogo de dados do *World Bank* (2012), *World Development Indicators*.

Outros três indicadores foram ainda recolhidos da *PENN World Table 7.0*, originária do *Center for International Comparisons of Production, Income and Prices* da Universidade da Pennsylvania. O referente à quota de consumo do governo sobre o PIB *per capita*, à quota de consumo do PIB *per capita* e à quota de investimento do PIB *per capita*.

A totalidade destas variáveis permitirá mensurar a actividade económica dos países e analisar a importância do fenómeno do crescimento. Qualquer um destes índices isoladamente apenas nos dá uma visão parcial do crescimento económico, já que nenhum deles cobre todos os aspectos desse crescimento, expectando-se através da sua utilização conjunta obter uma visão real do contexto económico.

#### • **Caracterização Sociodemográfica**

No que se refere à caracterização sociodemográfica são consideradas variáveis ao nível da estrutura da população, dimensão, distribuição etária, mortalidade, escolarização, direitos, fraccionamento, entre outros. A dimensão populacional (fonte: *World Bank*) consiste numa variável de ampla utilização empírica - anteriormente abordada por: Azam e Delacroix (2006), Azam e Thelen (2008, 2010), Braithwaite e Li (2007), Burgoon (2006), Campos e Gassebner (2009), Crenshaw *et al.* (2007), Dreher e Fischer (2010, 2011), Dreher e Gassebner (2008), Eyerman (1998), Freytag *et al.* (2011), Gassebner e Luechinger (2011), Kis-Katos *et al.* (2011), Koch e Cranmer (2007), Krueger e Laitin (2008), Krueger e Maleckova (2003), Lai (2007), Li e Schaub (2004), Li (2005), Meierrieks e Gries (2012), Neumayer e Plümper (2009), Piazza (2006, 2007, 2008a, 2008b, 2011), Plümper e Neumayer (2010), Robison *et al.* (2006), Sambanis (2008), Savun e Phillips (2009), Urdal (2006) e Walsh e

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Piazza (2010) - dada a relevante importância como indicador demográfico. A este nível serão incluídas igualmente as variáveis crescimento populacional, anteriormente considerada por: Dreher e Fischer (2010), Dreher e Gassebner (2008), Gassebner e Luechinger (2011) e Piazza (2006, 2009), a densidade populacional - Newman (2006) e a população urbana (fonte: *World Bank*), por vários autores designada apenas por urbanização - proposta por: Campos e Gassebner (2009), Gassebner e Luechinger (2011), Kis-Katos *et al.* (2011), Newman (2006), Robison *et al.* (2006) e Tavares (2004).

Diversos estudos reforçaram a relação positiva entre a dimensão populacional e o terrorismo. Dado o pressuposto de que o terrorismo se propaga em ambientes de maior tensão social e também pelo facto dos ataques pretenderem o maior impacto possível, espera-se que as variáveis relativas à densidade populacional e população urbana estejam positivamente correlacionadas com o terrorismo.

Os dados de caracterização etária foram recolhidos do catálogo do *World Bank* e distribuem-se por três escalões: população de 0-14 (anteriormente abordada por: Gassebner e Luechinger, 2011; Newman, 2006 e Tavares, 2004), população idade de 15-64 e população com 65 anos e acima. Igualmente se irá abordar a variável expectativa de vida ao nascer - com origem na mesma fonte - empiricamente considerada por: Drakos e Gofas (2006) e Gassebner e Luechinger (2011).

A avaliação do grau de escolarização (fonte: *World Bank*), será ponderada ao nível do ensino primário, anteriormente estudada por: Gassebner e Luechinger (2011) e Kurrild-Klitgaard *et al.* (2006), do ensino secundário, Azam e Thelen (2008, 2010), Drakos e Gofas (2006), Gassebner e Luechinger (2011) e Kurrild-Klitgaard *et al.* (2006), e do ensino terciário (ensino superior), Gassebner e Luechinger (2011), Kurrild-Klitgaard *et al.* (2006), e Testas (2004). Da mesma forma, serão tidas em conta as variáveis média de anos de educação (homens) e média de anos educação (mulheres), recolhidas na base de dados do *Institute for Health Metrics and Evaluation - University of Washington* (Gakidou *et al.*, 2010). Ao longo da literatura estes dois itens foram abordados de forma global por Blomberg e Hess (2008a), Bravo e Dias (2005), Gassebner e Luechinger (2011), Krueger e Maleckova (2003), Kurrild-Klitgaard *et al.* (2006) e Tavares (2004), tendo sido estudados conjuntamente no indicador taxa de alfabetização, total de adultos (% de pessoas com idades entre 15 e acima). No presente estudo este indicador não será incluído visto que os dados são demasiado escassos, com muitos valores em falta. Por outro lado, uma abordagem diferenciada entre homens e mulheres poderá conduzir a novos e discriminados resultados.

A inclusão da taxa de participação laboral das mulheres (% da população do sexo feminino 15+), proposta por Gassebner e Luechinger (2011) e Robison *et al.* (2006), pretende essencialmente abordar a questão cultural (fonte: *World Bank*). Embora a participação das mulheres na força de trabalho tenha vindo a aumentar de forma sustentada em todo o mundo, as desigualdades existentes com base no sexo persistem. É também incluída a variável direitos económicos da mulher, que avalia sobre igualdade de salários, independência de consentimento do marido na escolha e exercício de uma profissão, equidade nas práticas de contratação e promoção, segurança no emprego, não

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

discriminação pelos padrões, assédio sexual no trabalho, trabalho nocturno, integração em profissões perigosas ou em forças militares e policiais. De uma forma geral será considerada também a variável relativa aos direitos dos trabalhadores (Fonte: *The Cingranelli-Richards Human Rights Dataset*). Estudos anteriores abordam a caracterização dos direitos através de índices globais, como nos estudos de Feldmann e Perälä (2004) e Gassebner e Luechinger (2011).

Os dados relativos à utilização de Internet e linhas telefónicas, por cada 100 pessoas, têm origem na base de indicadores de desenvolvimento mundial do *World Bank*, e estão relacionados com o apuramento do nível de desenvolvimento e qualidade de infraestruturas de um país (incluídas por: Bravo e Dias, 2005; Lai, 2007; Gassebner e Luechinger, 2011 e Kis-Katos *et al.*, 2011).

A despesa total em saúde (fonte: *World Bank*) refere-se à soma dos números registados entre consumo público e privado, incluindo cuidados preventivos e curativos, planeamento familiar, nutrição e assistência em casos de urgência. Este constitui outro indicador relativo à avaliação do nível de desenvolvimento humano, uma vez que o investimento na saúde é crucial para evitar óbitos, incapacidades e criar um melhor futuro socioeconómico para as populações.

São igualmente introduzidos alguns indicadores relativos à mortalidade infantil e materna, geralmente usados para conhecer as condições de vida das populações. Serão eles a taxa de mortalidade pós-neonatal (por 1.000 nascimentos), que se refere a bebés dos 28 dias de vida até a 1 ano de idade - anteriormente abordada por: Gassebner e Luechinger (2011) e Urdal (2006), o rácio de mortalidade materna (por 100.000 nascidos vivos), a taxa de mortalidade sub-5 anos (por 1.000 nascidos vivos), proposta por Kurrild-Klitgaard *et al.* (2006), e a taxa de mortalidade neonatal (por 1.000 nascimentos), que inclui bebés de 0 a 28 dias de idade (Fonte: *Institute for Health Metrics and Evaluation - University of Washington*, Gakidou *et al.*, 2010).

No sentido de conhecer a caracterização da população ao nível da distribuição étnica e linguística, são consideradas duas variáveis que determinam a probabilidade de duas pessoas seleccionadas aleatoriamente não pertencerem ao mesmo grupo étnico e linguístico, estando avaliadas de 0 (mais homogéneo) a 1 (mais heterogéneo). Elevados níveis poderão indiciar agravamento das tensões sociais, expectando-se uma correlação positiva com o terrorismo. A fragmentação étnica foi considerada nos estudos de: Abadie (2006), Azam e Thelen (2010), Basuchoudhary e Shughart (2010), Bravo e Dias (2005), Drakos e Gofas (2006), Dreher e Fischer (2010), Gassebner e Luechinger (2011), Goldstein (2005), Kurrild-Klitgaard *et al.* (2006), Lai (2007), Piazza (2006), Sambanis (2008), Savun e Phillips (2009) e Tavares (2004), e a fragmentação linguística estudada por: Abadie (2006), Blomberg e Hess (2008a), Dreher e Fischer (2010), Gassebner e Luechinger (2011), Goldstein (2005), Kurrild-Klitgaard *et al.* (2006) e Piazza (2008a) - Fonte: Alesina *et al.* (2003).

- **Caracterização Política**

A caracterização ao nível da política tem como item de maior relevância o apuramento do grau de repressão dos regimes políticos. Para isso são considerados uma série de indicadores que pretendem estudar os países no que se refere ao tipo de regime, exercício do poder político, capacidade do Estado, estabilidade política e à liberdade de forma geral ou em campos específicos.

Nesta medida a variável de maior impacto, e mais utilizada empiricamente - Basuchoudhary e Shughart (2010), Blomberg e Hess (2008a, 2008b), Blomberg e Rosendorff (2009), Blomberg *et al.* (2004b), Braithwaite e Li (2007), Bravo e Dias (2005), Burgoon (2006), Campos e Gassebner (2009), Crenshaw *et al.* (2007), Drakos e Gofas (2006), Dreher e Fischer (2010, 2011), Dreher e Gassebner (2008), Freytag *et al.* (2011); Gassebner e Luechinger (2011), Kis-Katos *et al.* (2011), Koch e Cranmer (2007), Kurrild-Klitgaard *et al.* (2006), Lai (2007), Li e Schaub (2004), Li (2005), Meierrieks e Gries (2012); Neumayer e Plümper (2009), Piazza (2006, 2007, 2008b), Plümper e Neumayer (2010), Powers e Choi (2012); Sambanis (2008), Savun e Phillips (2009), Testas (2004), Urdal (2006), Walsh e Piazza (2010) - é a que aborda a democracia. Serão utilizados três indicadores: democracia - revelando se se trata de um regime democrático ou não (Fonte: Cheibub *et al.*, 2009); democracia institucionalizada - que revela sobre o grau de implantação e institucionalização da democracia (Fonte: *Polity IV Project*. Marshall e Jagers, 2009) e o nível de democracia - que avalia dentro dos limites de um intervalo contínuo, e por isso mais minucioso, o cariz democrático de cada país (Fonte: *Freedom House - Freedom in the World*).

Também entrará em análise o atributo fraccionamento político total, que em estudos anteriores tem sido aplicado em dois níveis, o fraccionamento do governo e o fraccionamento da oposição. Este item mede a probabilidade de dois deputados escolhidos aleatoriamente na Assembleia pertencerem a partidos diferentes, tendo já sido abordado por Dreher e Fischer (2010, 2011), Dreher e Gassebner (2008), Gassebner e Luechinger (2011), Koch e Cranmer (2007) e Piazza (2006) - Fonte: *Database of Political Institutions* (Beck *et al.*, 2001).

Para apuramento do tipo de regime é ainda incluída a variável instituições em regime (Fonte: Cheibub *et al.*, 2009), que enquadra o tipo de regime político de acordo com seis categorias, democracia parlamentar, semi-presidencial ou presidencial, ou ditadura civil, militar ou real (monárquica) - anteriormente analisada por: Bravo e Dias (2005). Aplica-se igualmente o índice autocracia institucionalizada (Fonte: *Polity IV Project*. Marshall e Jagers, 2009), anteriormente abordado por: Li (2005) e Li e Schaub (2004), que de acordo com a valoração por aplicação de vários indicadores singulares atribui o grau de repressividade do governo no exercício da autoridade. Ainda assim, entrarão em análise parte desses indicadores singulares, os quais têm empiricamente demonstrado diferentes relações com os números do terrorismo: regulamentação da participação política - Piazza (2008b, 2011), competitividade da participação política - Kis-Katos *et al.* (2011) e Piazza (2011), que indicam o nível de liberdade de participação política entre partidos permitida pelos regimes, que se prevê ser um preditor com efeito negativo do terrorismo, e por último a variável restrições executivas -



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Li (2005) e Piazza (2008b, 2011), que se gradua de acordo com a intensidade de influências diversas dirigidas ao governo limitando o seu poder de decisão, que poderá estar associada ao crescimento do terrorismo.

Dada a importância da estabilidade política demonstrada nos efeitos sobre o terrorismo, considerar-se-á também o atributo durabilidade de regime (Fonte: *Polity IV Project*: Marshall e Jagers, 2009), medido através do apuramento do número de anos passados desde a mais recente mudança de regime - anteriormente proposto por: Campos e Gassebner (2009), Eyerman (1998), Freytag *et al.* (2011), Gassebner e Luechinger (2011), Li (2005), Kis-Katos *et al.* (2011), Meierrieks e Gries (2012), Piazza (2007, 2008a, 2008b, 2011) e Savun e Phillips (2009).

Finalmente, entrarão para análise diversos indicadores que pretendem determinar a situação dos países relativamente à experiência de liberdade por parte da população. Será introduzido o atributo autodeterminação eleitoral (Fonte: *The Cingranelli-Richards Human Rights Dataset*), que traduz até que ponto os cidadãos gozam de liberdade de escolha política, bem como de direitos e capacidade prática de causar mudanças na alteração de leis e nos representantes do governo através de eleições livres e justas. As variáveis direitos políticos (Fonte: *Freedom House - Freedom in the World*), abordada anteriormente por: Abadie (2006), Basuchoudhary e Shughart (2010), Feldmann e Perälä (2004), Gassebner e Luechinger (2011), Goldstein (2005), Kurrild-Klitgaard *et al.* (2006), Newman (2006), Robison *et al.* (2006) e Tavares (2004), e os direitos políticos da mulher (Fonte: *The Cingranelli-Richards Human Rights Dataset*) igualmente entrarão como atributos explicativos, expressando a capacidade dos cidadãos, e no segundo em particular das mulheres, de participar livremente no processo político, incluindo o direito de votar por alternativas distintas em eleições legítimas, competir por cargo público, ou fazer parte de organizações ou partidos políticos.

A estes se juntam a variável liberdades civis, indicador global que abrange a liberdade de expressão e crença, associação e autonomia pessoal sem influência do Estado - abordada por: Basuchoudhary e Shughart (2010), Feldmann e Perälä (2004), Gassebner e Luechinger (2011), Krueger e Laitin (2008), Krueger e Maleckova (2003), Kurrild-Klitgaard *et al.* (2006), Li (2005), Newman (2006) e Robison *et al.* (2006) - e o índice *status* de liberdade, que combina classificações médias entre os níveis de liberdades civis e direitos políticos, dados recolhidos da base da *Freedom House: Freedom in the World*.

Outro indicador composto em análise consiste nos direitos de integridade física (Fonte: *The Cingranelli-Richards Human Rights Dataset*), proposto por Feldmann e Perälä (2004), Gassebner e Luechinger (2011) e Walsh e Piazza (2010), construído com base nas variáveis tortura, execução extrajudicial, prisão por motivos políticos e rapto ou desaparecimento.

Da mesma fonte foram ainda compilados atributos simples, como os referentes à liberdade de assembleia e associação, liberdade de expressão, prisão por motivos políticos e tortura, pois não se

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

pode excluir a possibilidade de poderem de forma isolada ter boa capacidade explicativa para as variáveis do terrorismo. A liberdade de assembleia e associação significa o direito dos cidadãos de livremente se reunirem com outras pessoas em partidos políticos, sindicatos, organizações culturais, ou outros grupos de interesses especiais. A liberdade de expressão indica o grau de liberdade de discurso ou imprensa que são afectados pela censura do governo, incluindo a propriedade dos meios de comunicação. Baixos índices de liberdade, característicos de sociedades que estão sob regimes repressores, sugerem uma menor intensidade de terrorismo, pelo que se espera encontrar resultados que apoiem a vertente que defende que a ausência de abusos no respeito pelos direitos humanos e a força e imparcialidade de um sistema jurídico estão associados a menos terrorismo, evitando injustiças graves geradoras de violência política.

- **Caracterização Religiosa**

As variáveis introduzidas em análise para estabelecer de forma geral a caracterização religiosa consistem no estudo da fragmentação religiosa (Fonte: Alesina *et al.*, 2003), proposta por: Abadie (2006), Blomberg e Hess (2008a, 2008b), Dreher e Fischer (2010), Freytag *et al.* (2011), Gassebner e Luechinger (2011), Goldstein (2005), Piazza (2006, 2008a) e Tavares (2004), no atributo religião católica (Kurrild-Klitgaard *et al.*, 2006), religião muçulmana (Bravo e Dias, 2005; Freytag *et al.*, 2011 e Kurrild-Klitgaard *et al.*, 2006) e outras religiões (Kurrild-Klitgaard *et al.*, 2006). O indicador referente à fragmentação religiosa reflecte a probabilidade de duas pessoas seleccionadas aleatoriamente de um dado país não pertencerem ao mesmo grupo religioso. Quanto mais alto o número, mais heterogénea e diversificada é a população. Os atributos relativos à religião medem a percentagem de população relativamente ao total e têm origem em La Porta *et al.* (1999).

Pretende-se aferir se a religião tem algum valor na explicação do terrorismo. Pelo senso comum, devido à mediatização do terrorismo de cariz islamista, existe uma previsível expectativa de encontrar uma relação entre religião e terrorismo. Alguns estudos abordam a questão da fragmentação religiosa em indicadores compostos com a inclusão em comum da fragmentação étnica e linguística, resultando num atributo que pretende estimar o grau de tensões sociais, o qual tem demonstrado ter significância no estudo do terrorismo através de uma relação positiva.

- **Outros dados de caracterização**

Por se esperar constituírem atributos com alguma importância explicativa, outras variáveis de carácter multidimensional foram ainda consideradas. O estudo do terrorismo no sentido de conhecer a sua relação com o avanço da globalização foi abordado por diversos autores, mas na sua maioria através da aplicação de atributos parciais que não atingem a totalidade da sua dimensão. Nesta investigação será também considerado o Índice de Globalização KOF, que define globalização como um processo que gera redes de conexões entre actores a distâncias multicontinentais, mediado por uma variedade de fluxos incluindo os de pessoas, informações e ideias, capitais e bens. Este atributo é um indicador composto que abrange o estudo da globalização ao longo de três dimensões: económica, política e social, utilizando em cada uma delas um vasto e abrangente conjunto de

variáveis<sup>19</sup>. O Índice de Globalização KOF tem origem em Dreher (2006) e Dreher *et al.* (2008), tendo sido abordado anteriormente por Gassebner e Luechinger (2011).

Para avaliar se existe relação entre a dimensão geográfica dos territórios e o terrorismo serão incluídas duas variáveis referentes a área: área de superfície (km<sup>2</sup>) - área total geográfica do país incluindo áreas de corpos de água e alguns canais costeiros - e área terrestre (km<sup>2</sup>), que exclui a área sob corpos de água interiores como grandes rios e lagos. A área foi anteriormente estudada como atributo explicativo do terrorismo por Eyerman (1998) e Piazza (2008b, 2011).

O presente estudo aplica igualmente a variável referente a despesas militares, empiricamente considerada por: Braithwaite e Li (2007), Burgoon (2006), Crenshaw *et al.* (2007), Freytag *et al.* (2011), Gassebner e Luechinger (2011), Koch e Cranmer (2007), Li (2005), Li e Schaub (2004), Meierrieks e Gries (2012), e que apesar das expectativas tem demonstrado ter uma relação positiva com o terrorismo.

#### 5.4 Técnicas de Data Mining

O campo do *data mining* surge da necessidade de extrair automaticamente informação potencialmente útil a partir de bases de dados de elevada dimensão e complexidade, consistindo na aplicação de técnicas estatísticas e de inteligência artificial para encontrar tendências ou padrões a fim de apoiar decisões (Furtado, 2002). Essas tendências ou padrões resultam na construção de modelos, os quais consoante a representação que se adopte podem tomar formas diversas tais como equações, regras, árvores ou grafos. Um modelo constitui um sumário global do conjunto de dados, devendo ter informação disponível para qualquer ponto no espaço de representação considerado (Rocha *et al.*, 2008).

Podem identificar-se diversas categorias de tarefas usadas para a mineração de dados (Fayyad *et al.*, 1996):

- **Exploração dos dados** - apenas se pretende realizar uma análise aos dados, usando ferramentas interactivas e visuais;
- **Descrição** - procuram-se modelos que sumariem os dados ou a forma como foram gerados;
- **Predição** - o objectivo é o de construir modelos que permitam prever o valor de uma variável a partir dos valores conhecidos de um conjunto de outras variáveis. O seu desenvolvimento pode ocorrer a dois níveis:
  - **Classificação**, onde se procura gerar conceitos genéricos sobre algo com o intuito de permitir classificações futuras. É aplicada a uma variável de saída discreta<sup>20</sup>, executando a correcta atribuição de uma classe a um conjunto de valores para as

<sup>19</sup> <http://www.globalization.kof.ethz.ch/>

<sup>20</sup> Variável discreta pertence a um domínio composto por um conjunto finito de valores possíveis. Na estatística aparecem como variáveis qualitativas (Rocha *et al.*, 2008).

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

variáveis de entrada. Abrangem tarefas de diagnóstico e de reconhecimento de padrões;

- **Regressão**, pretendendo-se neste caso aproximar-se uma função desconhecida com um contradomínio real. As saídas são contínuas<sup>21</sup> e os modelos construídos tipicamente com base em expressões matemáticas.
- **Descoberta de padrões e regras** - procura detectar padrões nos dados, recorrendo a regras de associação;
- **Clustering ou segmentação** - pretende a descoberta de grupos de itens com características semelhantes, tratando-se de aprendizagem não supervisionada. Estes grupos são analisados quanto ao seu significado e conceito particular de afinidade.

Associado aos sistemas computacionais aparece o conceito de Aprendizagem Automática (AA), em que programas são capazes de aprender ou adquirir conhecimento através de estudo, ensino ou experiências anteriores (Rocha *et al.*, 2008). Mitchell (1997) define AA como o estudo de algoritmos computacionais que produzem conhecimento através da experiência.

A Aprendizagem Automática caracteriza-se por três paradigmas fundamentais (Patterson, 1996; Rocha, 2008):

- A **aprendizagem supervisionada**, em que é fornecida uma resposta correcta para cada situação, sendo a aprendizagem realizada a partir de exemplos ou casos de treino compostos por um vector de entradas e por um vector de saídas desejadas (assume-se a presença de um supervisor);
- A **aprendizagem por reforço**, em que também se assume a presença de um supervisor, embora não seja dada a resposta correcta. Apenas se fornecem recompensas ou punições mediante as respostas dadas pelo modelo em determinadas situações, usadas pelo algoritmo de aprendizagem para melhorar a sua eficácia, podendo ser fornecidas logo a seguir à acção ou de forma mais indirecta, após um conjunto de acções. Mais utilizada na aprendizagem de estratégias para jogos, ou em problemas de controlo (robôs ou processos químicos/biotecnológicos).
- A **aprendizagem não supervisionada**, em que não é fornecida nenhuma indicação externa, decorrendo a aprendizagem pela descoberta de regularidades nos dados de entrada.

Na escolha da técnica de mineração de dados a utilizar verifica-se que não existe uma técnica que resolva todos os problemas, pelo que diferentes métodos existem para diferentes propósitos. Dias (2002) distingue um resumo das técnicas mais frequentes e as tarefas a que se aplicam:

- **Descoberta de Regras de Associação** - estabelece uma correlação estatística entre atributos de dados e conjuntos de dados. Relacionada com tarefas de associação;

---

<sup>21</sup> Variável numérica pertence a um domínio definido como sendo um subconjunto do conjunto de números reais, usualmente verificando-se a delimitação dos valores de um atributo num dado intervalo definido com limites inferiores e superiores. Na estatística aparecem como variáveis quantitativas (Rocha *et al.*, 2008).

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

- Árvores de Decisão - ligada à hierarquização dos dados, com base em estágios de decisão (nós) e na separação de classes e subconjuntos. Relacionada com tarefas de classificação (árvores de classificação) ou regressão (árvores de regressão);
- Raciocínio Baseado em Casos ou MBR (*Model-based Reasoning*) - baseado no método do vizinho mais próximo, combinando e comparando atributos para estabelecer hierarquia de semelhança. Relacionada com tarefas de classificação e segmentação;
- Algoritmos genéticos - baseado em métodos gerais de busca e optimização, constituindo uma classe particular de algoritmos evolutivos que usam técnicas inspiradas pela biologia evolutiva como a hereditariedade, mutação, selecção natural e recombinação (ou *crossing over*). Relacionada com tarefas de classificação e segmentação;
- Redes Neurais Artificiais: Modelos inspirados na fisiologia do cérebro, onde o conhecimento é fruto do mapa das conexões neurais e dos pesos dessas conexões. Relacionada com tarefas de classificação e segmentação.

No processo de descoberta de conhecimento, a determinação das características dos dados em análise, tem como meta seleccionar a técnica de mineração de dados que minimiza o número e dificuldades de transformação de dados para, a partir destes, obter bons resultados. Cada atributo tem associado um tipo de dados, que determina o seu domínio de valores possíveis. Importa distinguir dois tipos básicos de atributos (Rocha *et al.*, 2008):

- Catoriais ou discretos: são campos que apresentam valores de um conjunto de possibilidades limitado e predeterminado. Também conhecidos como enumerados, qualitativos ou categóricos. A este tipo de dados estão associadas as técnicas de descoberta de regras de associação e árvores de decisão. Este tipo de atributo pode ainda ser diferenciado em duas classes, ordenados, quando os valores estão compreendidos numa escala, e nominais, quando não existe nenhuma ordem especial nos dados.
- Numéricos ou contínuos: são aqueles que podem ser somados e ordenados. O domínio pode ser definido como um subconjunto do conjunto de números reais, sendo comum a delimitação dos valores de um atributo num dado intervalo, definindo-se limites inferiores e superiores.

O *data mining* no contra-terrorismo é entendido como uma abordagem que usa algoritmos para a determinação de padrões em bases de dados. A análise automática de dados aplica modelos aos dados para a previsão do seu comportamento, determinação de associações, riscos, entre outros. Os modelos usados podem ser baseados em padrões (aferidos por *data mining*) ou em assuntos, que direccionam a pesquisa incidindo sobre um tema específico (DeRosa, 2004).

O presente estudo irá ser explorado com base em Árvores de Regressão, especialmente indicadas para a predição de valores numéricos. A globalidade das árvores de decisão, sejam de classificação ou regressão, tem a mesma estrutura, baseando-se a sua construção num processo recursivo que faz uma pesquisa *top-down* (do mais geral para o particular) até atingir progressivamente as especificidades do conjunto de dados em análise. Quando gerada a árvore, cada registo percorre um

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

determinado caminho para uma série de testes, até que o registo alcança uma folha ou nó puro da árvore, sendo-lhes atribuída uma classe. Pela avaliação do conjunto de treino, em problemas de classificação são classificados os registos numa classe, e no caso das árvores de regressão, um valor baseado numa função matemática, como a média, mediana ou moda, dos valores que alcançaram essa folha no conjunto de treino (Santos *et al.*, 2005).

As árvores de decisão são largamente utilizadas e é um método adequado para problemas com muitas variáveis, visto tratar-se de um método de predição não linear rápido, e com fácil representação e visualização do conhecimento. Tem a vantagem de ser bem sucedido em situações em que as variáveis explicativas são uma mistura de variáveis nominais, ordinais e contínuas. É um método que se adapta facilmente a dados omissos, tendo outras vantagens como ser inalterável após aplicação de transformações nas variáveis (logaritmização); não necessitar de satisfazer nenhum pressuposto como condição para a aplicabilidade do modelo (modelo não-paramétrico); construir modelos para qualquer função desde que o número de exemplos seja suficiente; modelar relações não lineares sem necessidade de transformações nos dados, que tornam os modelos difíceis de interpretar; ser bastante eficiente na construção de modelos e possuir robustez à presença de pontos extremos e atributos redundantes (Cabete e Cardoso, 2006).

De entre os vários tipos de algoritmos de construção de árvores de decisão destacam-se o *CHAID* (Kass, 1980), o *CART* (Breiman *et al.*, 1984) e o *C5* (Quinlan, 1993), como os mais conhecidos. Os dois primeiros adaptam-se à construção de árvores de classificação e regressão, enquanto que o último destina-se apenas para a classificação. Dos três distingue-se o *CART* (*Classification and Regression Trees*), apresentado por Breiman *et al.* (1984), que proporciona árvores binárias que são de fácil leitura e interpretação, gerando também medidas de importância relativa das variáveis explicativas.

Na construção de uma árvore importa definir um critério para decidir qual a melhor ramificação de um nó e regras para decidir a paragem do processo de partição a partir de um nó, passando a considerá-lo terminal. É necessário também determinar um método de cálculo a partir de cada nó-folha e medidas adequadas do desempenho do modelo na previsão. Esta construção é feita a partir dos dados de treino, de uma forma repetitiva, subdividindo este conjunto de dados, para os quais é inferida uma ou mais regras, até que seja formada apenas por nós que correspondam à satisfação de um critério (Santos *et al.*, 2005).

Seguidamente à fase de construção das árvores obtém-se uma estrutura ajustada aos dados, normalmente demasiado longa. Nesta fase deve-se aperfeiçoar o modelo proposto através da sua poda, eliminando ramos da árvore depois de uma análise ponderada acerca da sua inclusão ou não. A poda desempenha um papel importante, produzindo árvores menores com melhor potencial e precisão quando aplicadas a novos casos (Han e Kamber, 2001; Santos *et al.*, 2005), ou seja, com maior capacidade de generalização. A poda pode ser feita durante ou após a aprendizagem, mas a

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

maioria dos algoritmos utiliza a técnica de poda após a aprendizagem, altura em que já se pode fazer uma análise global do contexto. As regras de paragem antecipada da construção da árvore favorecem a obtenção de um modelo com maior potencial de generalização e interpretação, efectivando-se pela imposição de um número máximo de níveis à árvore de decisão ou um número mínimo de observações num nó interno e/ou nos seus nós descendentes.

A escolha do algoritmo a utilizar neste trabalho recaiu sobre o CART, metodologia de regressão não paramétrica através da qual se pretende explicar a actividade terrorista e identificar os seus determinantes.

No que se refere ao crescimento da árvore, esta é obtida a partir de sucessivas divisões binárias do conjunto de dados (amostra de treino) através de uma medida de homogeneidade, que é usada para decidir qual a melhor variável de corte e valor de corte associados a cada nó, partindo este processo de construção do geral para o particular (Cabete e Cardoso, 2006).

As componentes básicas da árvore são os nós e as regras de divisão, estando os nós associados aos subconjuntos resultantes da aplicação de uma regra de divisão a um determinado conjunto de dados. O primeiro nó de uma árvore é chamado de nó raiz e equivale ao conjunto de dados completo. Os nós gerados pela divisão de um nó já existente recebem o nome de descendentes e o nó que os originou é chamado de ascendente ou pai.

Cada nó é dividido em dois nós descendentes, sendo a diversidade dos valores da variável dependente nestes nós mais reduzida do que no nó ascendente. Para cada divisão, com o objectivo de definir a melhor variável de corte, é avaliada a redução da variância respeitante à variável alvo. Este processo é feito iterativamente, visto que cada nó obtido é considerado como um nó pai, ao qual será aplicado um novo critério de ramificação, que assumirá menor variabilidade do que a ramificação que a antecedia. Cabete e Cardoso (2006) referem ainda que no caso de sobreajustamento no crescimento da árvore aos valores da amostra de treino, que dificulta a generalização do modelo obtido, é usual definirem-se regras de paragem de crescimento da árvore, gerando uma melhor interpretação da mesma.

Aquando da análise à árvore construída, verifica-se que a previsão associada a um elemento que foi encaminhado para determinado nó folha será dada pela média no nó-folha onde esse elemento se enquadra, sendo essa previsão igual para todos os elementos que pertençam ao mesmo nó-folha. Um indicador da precisão do modelo global será a soma ponderada das variâncias intra nós-folha na amostra de treino, o que não dispensa a verificação quanto a sobreajustamento do modelo gerado (Cabete e Cardoso, 2006).

Para a validação do modelo podem ser utilizadas duas técnicas, a primeira com base no resultado da aplicação do modelo a um conjunto separado de dados (chamado conjunto de teste), que é distinto

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

do conjunto de treino, tendo sido extraído de forma aleatória da amostra original. Na altura de construção da árvore este conjunto de dados - normalmente de dimensão inferior em relação aos restantes - não foi tido em conta, pelo que as estimativas obtidas a partir do mesmo não são tendenciosas e se existirem dados suficientes são confiáveis. Quando o número de observações disponíveis não for suficiente para a constituição de amostras de treino e teste, sendo a dimensão da amostra reduzida, opta-se pela utilização de uma segunda técnica que consiste na validação cruzada (Santos, F. e Azevedo, C., 2005; Cabete e Cardoso, 2006).

Quando os dados são escassos, a validação cruzada permite calcular um erro mais realista para a árvore apresentada. Neste caso, os dados de treino são divididos em  $n$  conjuntos de igual tamanho, de preferência com a mesma distribuição de classes e seguidamente são construídas  $n$  árvores diferentes usando casos de todos os blocos  $[(n-1)/n]$  e testadas com um bloco reservado  $[1/n]$  para avaliar o erro. Ou seja, o treino é efectuado em várias iterações sobre o conjunto de dados, sendo que, a cada iteração são definidos dois subconjuntos diferentes de treino e testes. A partir da construção das  $n$  árvores consegue-se calcular o erro associado a todas as observações, obtendo resultados que reflectem todo o conjunto de dados inicial, tirando proveito da diversidade do mesmo na sua totalidade. O método mais utilizado é a aplicação de  $n=10$  no processo de validação cruzada, sendo a proporção de variação explicada resultante capaz de gerar uma estimativa da capacidade preditiva do modelo proposto.

Após a validação do modelo em árvore torna-se necessária a sua interpretação. Breiman *et al.* (1984) sugerem como apoio à interpretação da árvore de decisão CART uma medida  $M_j$  de importância das variáveis explicativas  $X_j$  usadas na construção da árvore. Devido à possibilidade de as variáveis explicativas poderem aparecer na previsão mascaradas por outras, não aparecendo como responsáveis por ramificações mas proporcionando boas substitutas nessa tarefa, a sua importância deve ser medida atendendo a uma potencial contribuição para a previsão.

A medida de importância das variáveis explicativas ( $M_j$ ) baseia-se na redução da diversidade desencadeada pelo uso da variável  $X_j$  em cada ramificação ou pelo seu potencial uso traduzido no conceito de ramificação substituta. Desta forma, a medida de importância de uma variável explicativa é dada pelo somatório das diversas reduções de diversidade associadas a essa variável em cada uma das ramificações, ou em possíveis ramificações substitutas. Esta medida pode ser normalizada, entre um valor máximo de 100 e um mínimo de zero. Quanto maior o valor da medida de importância das variáveis explicativas, maior é o seu poder explicativo da variável dependente (Cabete e Cardoso, 2006).



## **5.5 Aplicação da Metodologia CRISP-DM**

A primeira etapa desenvolvida consistiu em intensas pesquisas ao nível da revisão da literatura sobre o tema em causa, quais os factores determinantes do terrorismo. No sentido do apuramento sobre a pertinência do tema e qual a direcção de investigação a tomar foram analisados diversos estudos publicados ao longo da última década e vários *sites* com informação sobre terrorismo. As pesquisas efectuadas permitiram determinar e escolher as variáveis de terrorismo e as variáveis explicativas mais pertinentes a incluir na análise, as questões de investigação a examinar e quais as fontes de dados disponíveis. Sendo pretensão retirar conclusões generalizáveis decidiu-se analisar indicadores de terrorismo no seu conjunto total, o número total de incidentes e o número total de mortos e de feridos, incluindo o terrorismo doméstico e transnacional.

Seguidamente foram consultadas todas as bases de dados *online* e os respectivos *codebooks*, para as perceber e descrever, analisando a sua capacidade e interesse para inclusão no estudo e possível compatibilização com os restantes dados. Algumas variáveis foram logo excluídas por terem muitos valores omissos, contendo dados de poucos países ou poucos anos.

Tendo-se optado pela concretização de uma análise em série temporal, avaliando diversas variáveis ao longo do tempo, construiu-se um ficheiro de dados organizado segundo o modelo de dados em painel, por país, durante um período de vinte anos decorrido entre 1990 a 2010.

Os dados de terrorismo vinham organizados por incidente terrorista, pelo que foi necessário adaptar a base ao formato de dados em painel acima referido. Visto que a base de dados de terrorismo manteve para cada atentado a nomenclatura original dos países aquando da sua ocorrência, tornou-se necessário uniformizar e reunir as observações para uma única designação por país, com o nome actual. Dado que as circunstâncias políticas dos diversos países mudaram substancialmente, em que alguns países deixaram de existir e outros foram constituídos, foi necessário examinar todos os países constantes da base, tendo a sua harmonização sido minuciosamente determinada através do estudo da evolução político-geográfica de cada país. Diversos países tiveram que ser eliminados por se tornar impossível estabelecer a localização dos incidentes à luz actual, após terem ocorrido decomposições e reorganizações de território.

Ainda quanto aos dados de terrorismo foi decidido não incluir no estudo o ano de 1993, cujos dados estavam bastante incompletos devido ao extravio do seu suporte físico pela empresa responsável, conforme acima mencionado na descrição das variáveis de terrorismo, os quais não foram possíveis de reconstituir.

Aos dados de terrorismo ora constantes do novo ficheiro juntaram-se as variáveis de caracterização dos países, o que devido ao seu elevado número e diferente disposição, gerou um considerável esforço de uniformização entre todos os registos. Da análise aos dados disponíveis referentes às

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

diferentes variáveis de caracterização dos países, que tiveram origem em diversas fontes, resultou a eliminação de mais alguns países, que totalizaram o número de 173 na versão final.

Os dados recolhidos foram minuciosamente analisados quanto à sua consistência, valores omissos, *outliers* e sua distribuição ao longo da amostra. A escolha da técnica de modelação a utilizar recaiu sobre as árvores de regressão com o algoritmo CART, que é um método adequado para problemas multidimensionais e de fácil representação e interpretação, adaptando-se com sucesso a conjuntos que incluam simultaneamente variáveis explicativas nominais, ordinais e contínuas.

Brown e Kross (2003) recomendam uma série de soluções para o tratamento de valores omissos ou desconhecidos que possam surgir nas bases de dados:

- Ignorar os registos com valores omissos;
- Substituir (imputação) cada valor omissos por:
  - Valor dado por um perito (*case substitution*);
  - Valor médio, mediana ou mais comum (moda) do atributo;
  - Valor retirado de outra base de dados (*cold deck*);
  - Valor do exemplo mais semelhante ou próximo (*hot deck*);
  - Valor estimado por regressão linear;
  - Combinação dos métodos anteriores (*multiple imputation*).

Desta forma, os valores omissos e desconhecidos foram ignorados, apesar de o CART ter a vantagem de se adaptar bem à presença de valores omissos nas variáveis explicativas pois podem ser estimados a partir de outras variáveis explicativas para que todos os dados, sempre que possível, sejam usados numa árvore. Nesse sentido, o algoritmo encontra variáveis de corte alternativas para cada divisão da árvore. Estas variáveis alternativas tendem a ser similares às variáveis seleccionadas em primeiro lugar, obtendo-se, assim, resultados idênticos. Outra das vantagens é a capacidade de lidar de uma forma intuitiva com os valores extremos (*outliers*), evitando que estes afectem a ramificação da árvore.

Para evitar os inconvenientes associados à retrocausalidade, procedeu-se ao desfasamento das variáveis explicativas num período de um ano ( $n-1$ ) relativamente aos dados de terrorismo. Os dados foram introduzidos e analisados no software aplicativo IBM SPSS, tendo sido inseridas as escalas no programa e convertidos os atributos não numéricos.

Posteriormente foram analisadas as correlações entre as variáveis de cada dimensão, de forma a, com base nas correlações com coeficientes superiores a 90%, decidir que atributos eliminar, visto que a sua utilização conjunta não iria acrescentar valia ao trabalho, acabando por dificultar a explicação e interpretação dos modelos.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

No sentido de ter como entradas apenas os atributos de maior importância de cada uma das dimensões explicativas, foram geradas árvores de regressão parciais, pretendendo identificar e seleccionar apenas as variáveis com importância relativa acima de 40% na explicação de cada variável de terrorismo em análise.

Aquando da construção das árvores foram consideradas algumas regras de paragem específicas. A parametrização das árvores de regressão referidas foi a constante da Tabela 2.

**Tabela 2: Parametrização das árvores de regressão**

Regra de paragem	Parametrização fixa
N.º máximo de níveis da árvore	5 níveis
N.º mínimo de observações em nó <i>pai</i>	2 observações
N.º mínimo de observações em nó <i>filho</i>	1 observação
Decremento mínimo de variância numa ramificação	0,0001

A escolha do número de níveis (profundidade da árvore) limitando a construção de uma árvore completa, pretende evitar um sobreajustamento do modelo aos dados, facilitando a sua interpretação e capacidade de generalização. Quanto maior for o número de níveis, maior será a capacidade preditiva sobre a amostra de treino da árvore, devido à redução da dispersão em cada nó, aumentando a probabilidade de acertar numa previsão com base nas regras dos nós. Contudo, desta forma pode-se estar a dificultar a generalização do modelo, por estar demasiado ajustado aos dados de treino, bem como a interpretação das árvores devido à existência de muitos níveis. Deve-se, assim, encontrar o equilíbrio entre a simplicidade do modelo e a sua capacidade preditiva, de forma a que a sua qualidade global, medida pelo  $R^2$ , não seja posta em causa. A proporção de variação explicada permite obter uma estimativa da capacidade preditiva dos modelos propostos, sendo obtida através do seguinte cálculo:

$$1 - \sum_{v=1}^V \frac{n_v}{n} \left( \frac{\sum_{i=1}^{n_v} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n_v} (y_i - \bar{y}_v)^2} \right) \quad (1)$$

As estimativas da capacidade preditiva dos modelos foram obtidas sobre dados novos e não sobre dados empregues na sua construção mediante um processo de validação cruzada «10-fold».

Com as restantes variáveis foram gerados os modelos finais, para os três indicadores de terrorismo, em que foram avaliados diversos modelos para cada problema, com diferentes parametrizações quanto à profundidade e número de casos nos nós *pai* e *filho*. Tendo-se procedido à avaliação de todos os modelos relativamente à sua capacidade explicativa para a amostra e capacidade preditiva medida através da validação cruzada, escolheram-se, para cada variável de terrorismo, os modelos que consistissem no melhor equilíbrio entre capacidade explicativa e preditiva.

Na avaliação dos resultados foram interpretados e comparados os modelos das três variáveis dependentes, para identificar variáveis e dimensões explicativas comuns a todos, delineando de

forma global os factores determinantes para o terrorismo e qual a dimensão que melhor explica a actividade terrorista.

## **5.6 Sumário**

Neste capítulo foi abordado o conceito de *data mining*, que se caracteriza como um processo que recorre a várias técnicas estatísticas e de reconhecimento de padrões para a descoberta de paradigmas e relações nos dados ou variáveis registadas. As técnicas de *data mining* são adequadas quando se têm muitos dados de origem multidimensional e não foi identificado nenhum estudo em que tenham sido utilizadas. Foi descrito o processo de DCBD e a metodologia CRISP-DM, destacando-se esta última em resultado da progressiva inclusão e generalização de todo o procedimento de descoberta de conhecimento no processo de *data mining*.

As tarefas e técnicas de *data mining* foram sucintamente enumeradas, tendo sido alvo de maior enfoque as Árvores de Regressão CART, que foram aplicadas no presente estudo. Foram apresentadas as variáveis de caracterização económica, sociodemográfica, política, religiosa e outras, que constituem as variáveis independentes ou explicativas do terrorismo. Os dados de terrorismo que se pretendem explicar são o número de ataques, o número de mortos e o número de feridos. Na recolha de elementos para a definição dos indicadores a aplicar ao estudo do terrorismo procurou-se abranger o máximo de variáveis possíveis de forma a testar e incluir maiores possibilidades explicativas dentro de uma visão globalizante.

Para o presente estudo reuniram-se os grupos de variáveis usados em estudos anteriores, acrescentando alguns atributos explicativos que poderão ter significância na relação com o terrorismo, como taxas de mortalidade neonatal e materna, que ajudam a caracterizar as condições de vida das populações, escolaridade diferenciada por géneros, estrutura etária, introdução do índice *status* de liberdade, que combina classificações médias entre os níveis de liberdades civis e direitos políticos e também outros atributos relativos a direitos e liberdades, que em alguns casos estão inseridos em indicadores globais, flexibilizando a determinação da sua importância parcial. A utilização de atributos ainda pouco abordados empiricamente proporcionará um maior aprofundamento no seu estudo e poderá conceder maior robustez a anteriores resultados.

Por último, apresenta-se a abordagem metodológica utilizada descrevendo-se as etapas referentes ao desenvolvimento e aplicação da metodologia CRISP-DM deste trabalho de investigação.

## 6. Resultados

---

No presente capítulo é apresentada a análise descritiva da amostra e dos dados de terrorismo que se pretendem explicar. Antes da fase de modelação é determinada a importância relativa dos atributos explicativos através de uma análise de correlações e da construção de árvores regressão parciais, por dimensões - económica, sociodemográfica, política, religiosa e indiferenciada, no sentido de eliminar as variáveis irrelevantes, visto de que a sua presença pode afectar a capacidade de aprendizagem devido a um aumento de ruído nos dados tornando mais difícil a explicação dos modelos.

Seguidamente são gerados e interpretados os modelos explicativos para as três variáveis dependentes, sendo ainda enquadrados e discutidos os resultados em harmonia com estudos empíricos anteriores.

### 6.1 Caracterização da Amostra

A investigação tem por base uma amostra constituída por 173 países, distribuídos por um conjunto de 13 regiões, geograficamente dispersas por todo o globo. A análise incide num período de vinte anos, decorrido de 1990 a 1992 e de 1994 a 2010.

Na distribuição dos países pelas diversas regiões (Tabela 3) constata-se que é na África Sub-Sahariana que se concentram 25,43% dos países, perfazendo cerca de um quarto dos países. Seguem-se as regiões da Europa Ocidental, Médio Oriente e Norte de África e da América Central e Caraíbas com valores a rondar entre os 11,56% e os 10,40% do total de países em estudo.

**Tabela 3: Número de países por região**

Região	Países	%
África Sub-Sahariana	44	25,43
Europa Ocidental	20	11,56
Médio Oriente e Norte da África	19	10,98
América Central e Caraíbas	18	10,40
Europa de Leste	13	7,51
América do Sul	12	6,94
Sudeste Asiático	10	5,78
Ásia do Sul	10	5,78
Rússia e os Novos Estados Independentes (NEI)	9	5,20
Australásia e Oceania	7	4,05
Ásia Central	4	2,31
Leste Asiático	4	2,31
América do Norte	3	1,73
<b>Total</b>	<b>173</b>	<b>100,00</b>

Na tabela 4 pode-se observar o conjunto total de países constantes da amostra, que se encontram mais concentrados em algumas regiões. Contudo, um elevado número de países não significa mais

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

território. De acordo com os resultados apresentados na Tabela 5, a região África Sub-Sahariana apesar de manter o valor mais elevado, apenas representa 18,62% da área geográfica de toda a amostra, seguida da América do Norte, com 16,55%, que na distribuição dos países detinha o valor mais baixo, 1,73%. A Europa Ocidental que ocupa o segundo lugar no ranking dos países, com 11,56%, no cálculo da área correspondente figura apenas com o valor de 2,83%. Igualmente díspar é a relação revelada pela América Central e Caraíbas que detém 10,40% dos países da amostra e apenas 0,57% da totalidade da área em estudo.

**Tabela 4: Países em estudo**

África Sub-Sahariana		América do Norte	Europa de Leste	Médio Oriente e Norte da África	Sudeste Asiático
África do Sul	Rep. África Central	Canadá	Albânia	Arábia Saudita	Brunei
Angola	Ruanda	EUA	Bósnia-	Argélia	Camboja
Benin	Senegal	México	Bulgária	Bahrain	Filipinas
Botswana	Serra Leoa	<b>América do Sul</b>	Croácia	Chipre	Indonésia
Burkina Faso	Somália	Argentina	Eslovénia	Egito	Laos - R.D.P.
Burundi	Suazilândia	Bolívia	Hungria	Emi. Árab. Unidos	Malásia
Camarões	Sudão	Brasil	Macedónia	Irão	Singapura
Chade	Tanzânia	Chile	Moldávia	Iraque	Tailândia
Comoros	Togo	Colômbia	Polónia	Israel	Timor-Leste
Congo, Rep. Democ.	Uganda	Equador	República Checa	Jordânia	Vietnam
Congo, Rep.	Zâmbia	Guiana	República	Kuwait	<b>Ásia do Sul</b>
Costa do Marfim	Zimbabwe	Paraguai	Roménia	Líbano	Afganistão
Djibouti	<b>América Central e</b>	Perú	Sérvia	Líbia	Bangladesh
Eritreia	<b>Caraíbas</b>	Suriname	<b>Europa Ocidental</b>	Marrocos	Butão
Etiópia	Antígua e Barbuda	Uruguai	Alemanha	Qatar	Índia
Gabão	Bahamas	Venezuela	Andorra	Síria	Maldivas
Gâmbia	Barbados	<b>Australásia e Oceania</b>	Áustria	Tunísia	Maurícia
Gana	Belize	Austrália	Bélgica	Turquia	Nepal
Guiné	Costa Rica	Fiji	Dinamarca	Yemen	Paquistão
Guiné Equatorial	Cuba	Ilhas Salomão	Espanha	<b>Rússia e os Novos</b>	Seychelles
Guiné-Bissau	Dominica	Nova Zelândia	Finlândia	<b>Estados Indep.</b>	Sri Lanka
Lesoto	Granada	Papua-Nova Guiné	França	Arménia	
Libéria	Guatemala	Samoa (Ocidental)	Grécia	Azerbaijão	
Madagascar	Haiti	Vanuatu	Holanda	Bielorrússia	
Malawi	Honduras	<b>Ásia Central</b>	Irlanda	Estónia	
Mali	Jamaica	Cazaquistão	Islândia	Geórgia	
Mauritânia	Nicarágua	Quirguistão	Itália	Letónia	
Moçambique	Panamá	Taijquistão	Luxemburgo	Lituânia	
Namíbia	República	Uzbequistão	Malta	Rússia	
Níger	Salvador	<b>Leste Asiático</b>	Noruega	Ucrânia	
Nigéria	São Cristóvão e Nevis	China	Portugal		
Quénia	Trinidad e Tobago	Coreia do Norte	Reino Unido		
		Coreia do Sul	Suécia		
		Japão	Suíça		

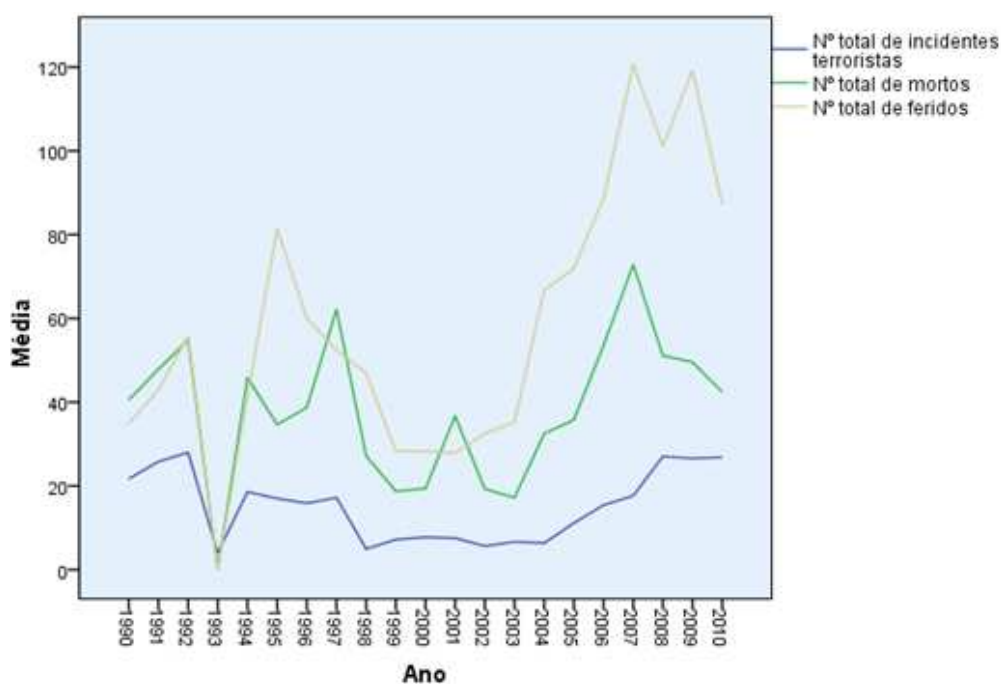
**Tabela 5: Área geográfica de cada região**

Região	Área geográfica (km <sup>2</sup> )	
	Total	%
África Sub-Sahariana	486.073.400	18,62
América do Norte	432.193.580	16,55
Rússia e os Novos Estados Independentes (NEI)	365.410.250	13,99
América do Sul	354.332.330	13,57
Médio Oriente e Norte da África	236.535.760	9,06
Leste Asiático	203.957.260	7,81
Australásia e Oceania	170.679.400	6,54
Ásia do Sul	102.714.030	3,93
Sudeste Asiático	76.366.190	2,92
Europa Ocidental	73.904.800	2,83
Ásia Central	70.295.980	2,69
Europa de Leste	23.753.980	0,91
América Central e Caraíbas	14.816.200	0,57
<b>Total</b>	<b>2.611.033.160</b>	<b>100,00</b>

## 6.2 Caracterização da Actividade Terrorista

Na investigação em curso a caracterização do terrorismo é abordada ao nível do número total de incidentes ocorridos, número total de mortos e número total de feridos. Pela observação da evolução destas três variáveis, em termos médios, (Figura 4) verifica-se um comportamento semelhante entre os três indicadores.

**Figura 4: Evolução do terrorismo no período em estudo**



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

A evolução do número de incidentes surge acompanhada por uma aproximada movimentação dos valores médios de mortos e feridos. No ano de 1993 existe uma quebra acentuada em todas as variáveis, justificada pelo extravio de uma parte da base de dados GTD, os quais se afiguraram de difícil reconstrução, causando um lapso de informação relativa a esse ano, conforme reportado no ponto 5.3.1, o que justifica a exclusão deste ano da análise.

No ano de 1995 constata-se uma redução do número de incidentes terroristas e mortos, mas por outro lado uma quase duplicação na contagem dos feridos relativamente ao ano anterior. Em 2001 também se observa uma discrepância quanto ao número de mortos que apresenta praticamente o dobro do ano anterior, tendo-se mantido a média de incidentes e feridos. Estes desvios justificam-se com a intensidade e impactos de cada incidente. No ano de 2001 o diferencial relativamente ao número de mortos<sup>22</sup>, que perfaz aproximadamente 3000, encontra explicação nos atentados de 11 de Setembro nos EUA em que apenas um incidente causou um valor muito elevado de mortos.

De 2003 para 2004 deu-se uma descida do número de incidentes, enquanto que o número de mortos e feridos praticamente atingiu o dobro. No ano de 2006 para 2007 constata-se uma subida de valores nos três impactos do terrorismo, mas comparativamente com o ano seguinte existe um aumento de incidentes em cerca de metade, ao passo que os valores de mortos e feridos têm uma queda acentuada.

O enquadramento geral dos dados apresentado pela Tabela 6 permite aferir que na amostra em estudo existe uma elevada concentração de incidentes, mortos e feridos, particularmente em 25% das observações (países/ano). Em metade das observações o valor máximo de incidentes é de 1, e de mortos e feridos 0. A análise a 75% dos casos conclui o valor máximo de 5 incidentes, 3 mortos e 4 feridos, enquanto que no total da amostra o número mais alto de incidentes é de 1176 (país/ano), de mortos 6534 (país/ano) e de feridos 11965 (país/ano). Desta forma, as médias de 16 incidentes, 40 mortos e 61 feridos por ano em cada país, constituem uma medida bastante desfasada da realidade e neste caso inadequada para caracterizar a distribuição, como aliás se verifica pela análise dos valores do desvio padrão. No total dos incidentes o peso do desvio padrão em relação à média é de 419%, o que se traduz num valor muito elevado, reflectindo uma grande dispersão nos dados. Quanto ao número de mortos e feridos constatam-se igualmente coeficientes de variação muito altos, nomeadamente de 573% e 685%.

**Tabela 6: Medidas descritivas das variáveis de terrorismo**

	Observ.	Média	Desv. Padrão	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
Nº total de incidentes terroristas	3460	15,74	67,27	0	0	1	5	1176
Nº total de mortos	3460	40,04	228,62	0	0	0	3	6534
Nº total de feridos	3460	61,18	417,71	0	0	0	4	11965

<sup>22</sup> No anexo B podem observar-se os valores referentes às três variáveis por ano.



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Numa perspectiva global parece evidente a existência de uma relação positiva entre os três indicadores. Pela análise da Tabela 7 pode confirmar-se uma correlação positiva intensa. O número total de incidentes quando associado ao número total de mortos apresenta um coeficiente de correlação *Pearson* de 0,773 ( $p < 0,001$ ) e em relação ao número total de feridos um coeficiente de correlação *Pearson* de 0,753 ( $p < 0,001$ ), valores estes correspondentes a uma correlação forte e positiva, tal como o valor resultante da relação entre o número de mortos e o número de feridos (*Pearson* 0,810;  $p < 0,001$ ).

Assim, verifica-se a existência de uma tendência forte para quando aumenta o número de incidentes aumentar o número de mortos e feridos e também que quanto maior o número de feridos, maior tende a ser o número de mortos.

**Tabela 7: Matriz de correlações entre características do terrorismo**

	Nº total de incidentes terroristas	Nº total de mortos	Nº total de feridos
Nº total de incidentes terroristas	1,000	0,773***	0,753***
Nº total de mortos		1,000	0,810***
Nº total de feridos			1,000

\*\*\* Significativa para 0,001

Relativamente à ponderação sobre o número e impacto dos ataques terroristas por região, exposta na Tabela 8, verifica-se que nesta avaliação a África Sub-Sahariana - que detém 25,43% dos países da amostra - é largamente ultrapassada pelas regiões da Ásia do Sul, Médio Oriente e Norte de África e América do Sul. A Ásia do Sul, que no apuramento dos valores ao nível dos países e área deteve 5,78% e 3,93% respectivamente, na avaliação quanto ao número de ataques e impactos do terrorismo posicionou-se em ponto de destaque, com 26,49% do número total dos incidentes terroristas, 29,19% do total de mortos e 30,45% do total de feridos.

A região do Médio Oriente e Norte de África, representativa de 10,98% dos países e de 9,06% da área total da amostra, encontra-se posicionada no segundo lugar relativamente à contagem do número total de incidentes com 24,79%, mas detém os valores mais elevados em referência ao número total de mortos, com 30,97%, e ao número total de feridos, com 36,89%.

As duas variáveis dependentes relativas ao número de mortos e feridos dão a conhecer o impacto ao nível dos danos humanos causados pelos incidentes terroristas ocorridos nessa região. Valores percentuais mais elevados nos totais de mortos e feridos em referência aos incidentes terroristas denotam que essa região foi alvo de terrorismo de cariz mais violento. A Ásia do Sul e com maior evidência o Médio Oriente e Norte de África destacam-se pelo elevado número de atentados, que entre as duas consiste em cerca de metade do valor total da amostra (51,28%), bem como pela dimensão dos impactos resultantes, perfazendo 60,16% do total de mortos e 67,34% no total de feridos relativamente ao conjunto global de valores.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Os baixos índices no número de mortos e feridos na Europa Ocidental, que representa 8,33% dos incidentes terroristas em análise, mostram que esta região é caracterizada por terrorismo de menor intensidade, diferenças associadas em parte às diversas motivações ideológicas que movem as várias organizações terroristas em actividade por todo o mundo.

**Tabela 8: Totais dos valores das variáveis dependentes por região**

Região	Nº total de incidentes terroristas		Nº total de mortos		Nº total de feridos	
	Total	%	Total	%	Total	%
Ásia do Sul	14428	<b>26,49</b>	40442	<b>29,19</b>	64446	<b>30,45</b>
Médio Oriente e Norte da África	13503	<b>24,79</b>	42895	<b>30,97</b>	78095	<b>36,89</b>
América do Sul	6865	<b>12,60</b>	11264	8,13	8994	4,25
África Sub-Sahariana	4569	8,39	25229	<b>18,21</b>	18495	<b>8,74</b>
Sudeste Asiático	4541	8,34	6408	4,63	11883	5,61
Europa Ocidental	4537	8,33	1245	0,90	7134	3,37
Rússia e os Novos Estados Independentes	1917	3,52	4129	2,98	6995	3,30
América Central e Caraíbas	1887	3,46	1713	1,24	2353	1,11
América do Norte	906	1,66	3809	2,75	1575	0,74
Europa de Leste	562	1,03	485	0,35	569	0,27
Leste Asiático	404	0,74	457	0,33	9715	4,59
Ásia Central	219	0,40	359	0,26	1266	0,60
Australásia e Oceania	136	0,25	89	0,06	159	0,08
<b>Total</b>	<b>54474</b>	<b>100,00</b>	<b>138524</b>	<b>100,00</b>	<b>211679</b>	<b>100,00</b>

### 6.3 Explicação da Actividade Terrorista

O presente tópico inicia-se com a determinação da capacidade explicativa das variáveis de cada dimensão sobre cada uma das variáveis dependentes. Para cada grupo de atributos são medidos os graus de associação, pretendendo-se ponderar a eliminação de atributos redundantes, isto é, quando os coeficientes de associação são superiores a 0,9. Através da aplicação de árvores de regressão a cada conjunto de atributos determina-se para cada um quais têm importância relativa superior. As variáveis com valores iguais ou superiores a 40% mantêm-se em estudo e são usadas na construção dos modelos finais, sendo as restantes eliminadas.

Seguidamente são construídos os três modelos finais, relativos ao número total de incidentes, número total de mortos e número total de feridos. Para cada uma destas variáveis são gerados vários modelos com regras de parametrização diferentes, sendo por fim escolhido o que confira uma melhor capacidade explicativa. Finaliza este item com a interpretação e discussão sobre os modelos explicativos de cada variável.

#### 6.3.1 A Importância Relativa dos Atributos Explicativos

Devido à existência de elevado número de variáveis em cada dimensão e com o objectivo de reduzir o número de variáveis a usar na construção dos modelos finais, são construídos CART's para as três variáveis dependentes no âmbito de cada conjunto de dados de caracterização. Atributos com baixa

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

importância relativa na explicação dos modelos podem afectar a capacidade de aprendizagem devido a um aumento de ruído nos dados, tornando mais difícil a sua interpretação. Pretendendo seleccionar apenas os atributos de maior importância relativa, para cada árvore foram escolhidas as variáveis com valor igual ou superior a 40%.

No sentido de caracterizar a relação entre as variáveis económicas foram determinados os coeficientes de correlação (*Pearson*) para todo o conjunto. Os valores resultantes variam entre zero, por exemplo entre a variável formação bruta de capital (crescimento anual em%) e combustíveis, % das importações de mercadorias (n-1) e 0,968 ( $p < 0,001$ ) para a relação entre o crescimento do PIB e o crescimento do PIB *per capita*. Os indicadores poupanças internas brutas (% do PIB) e despesa de consumo final (% do PIB) detêm uma relação perfeita negativa, visto que o primeiro indicador reflecte a diferença entre o PIB e a despesa de consumo final, o que justifica esta correlação. Pelo facto de resultarem numa medição idêntica, opta-se por eliminar o primeiro já que o segundo constitui um indicador mais popular e conhecido.

As variáveis RNB, crescimento *per capita* (% anual) e o crescimento do PIB *per capita* (% anual), bem como o RNB, crescimento (% anual) e o crescimento do PIB (% anual) detêm um coeficiente de correlação elevado, de 0,908 e 0,905 ( $p < 0,001$ ), respectivamente. Na revisão da literatura apenas no estudo de Piazza (2011) se detectou o uso de indicador referente ao RNB, enquanto que o crescimento do PIB e o crescimento do PIB *per capita* são atributos de referência em diversos estudos, conforme referido no ponto 5.3.2. Por esse motivo mantêm-se em análise estas duas últimas em detrimento das primeiras.

Avaliando a correlação entre o crescimento do PIB e o crescimento do PIB *per capita* constata-se um coeficiente *Pearson* de 0,968 ( $p < 0,001$ ), indicando uma relação muito forte. Perante a mais comum utilização teórica do crescimento do PIB, e por representar a economia de um país de forma global, como um todo, escolhe-se pela manutenção deste atributo.

Em seguida, foram geradas três árvores, uma para o número total de incidentes, outra para o número total de mortos e outra para o número total de feridos. Em todas elas se verificou a existência de duas variáveis referentes à inflação com níveis explicativos muito altos. Nesse sentido, como as duas estão fortemente correlacionadas (*Pearson* 0,743;  $p < 0,001$ ) e por ambas consistirem em conceitos aproximados, não trazendo assim a sua utilização conjunta mais-valia à investigação, optou-se pela eliminação da variável inflação, deflator do PIB (% anual), permanecendo a inflação, preços de consumo (% anual) por ser comumente mais usada e de melhor interpretação.

Após a construção de novas árvores para cada variável dependente com os atributos restantes, puderam apurar-se os valores mais relevantes quanto à sua capacidade explicativa do modelo. Quanto maior o valor da medida de importância das variáveis explicativas, maior é o seu poder explicativo sobre a variável dependente. A Tabela 9 apresenta os valores resultantes no que se refere

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

à importância de cada variável, ou seja, o seu peso na explicação do modelo, à importância relativa, que representa o valor anterior numa versão normalizada (em percentagem) e à sua ordenação.

**Tabela 9: Importância relativa dos atributos explicativos da caracterização económica dos países para a explicação de cada um dos impactos do terrorismo**

Atributos explicativos	N.º total de incidentes terroristas			N.º total de mortos			N.º total de feridos		
	Importância	Importância relativa	Ordem	Importância	Importância relativa	Ordem	Importância	Importância relativa	Ordem
Combustíveis, % das exportações de mercadorias (n-1)	<b>614,1</b>	<b>47,5%</b>	<b>3</b>	<b>19035,8</b>	<b>100,0%</b>	<b>1</b>	<b>74072,0</b>	<b>100,0%</b>	<b>1</b>
Combustíveis, % das importações de mercadorias (n-1)	106,5	8,2%		1238,9	6,5%		1401,3	1,9%	
Comércio (em % do PIB) (n-1)	207,8	16,1%		1355,0	7,1%		1249,8	1,7%	
Crescimento do PIB (% anual) (n-1)	<b>430,6</b>	<b>33,3%</b>	<b>4</b>	4522,1	23,8%		21184,9	28,6%	
Despesa de consumo final (% do PIB) (n-1)	20,5	1,6%		716,0	3,8%		6184,3	8,3%	
Despesa Nacional Bruta (% do PIB) (n-1)	53,4	4,1%		756,6	4,0%		2907,0	3,9%	
Despesas gerais de consumo final do Governo (% do PIB) (n-1)	10,1	0,8%		38,5	0,2%		979,2	1,3%	
Formação bruta de capital (% do PIB) (n-1)	20,7	1,6%		298,7	1,6%		143,1	0,2%	
Formação bruta de capital (crescimento anual em%) (n-1)	7,1	0,5%		447,0	2,3%		19,6	0,0%	
Inflação, preços de consumo (% anual) (n-1)	<b>1293,5</b>	<b>100,0%</b>	<b>1</b>	<b>18852,3</b>	<b>99,0%</b>	<b>2</b>	<b>64787,5</b>	<b>87,5%</b>	<b>2</b>
Investimento estrangeiro directo, entradas líquidas (% do PIB) (n-1)	97,7	7,6%		<b>11672,3</b>	<b>61,3%</b>	<b>3</b>	<b>48069,4</b>	<b>64,9%</b>	<b>4</b>
PIB per capita (n-1)	260,8	20,2%		<b>8425,5</b>	<b>44,3%</b>	<b>5</b>	<b>47371,3</b>	<b>64,0%</b>	<b>5</b>
Quota de consumo do Governo sobre o PIB per capita (n-1)	<b>671,8</b>	<b>51,9%</b>	<b>2</b>	5065,6	26,6%		<b>38474,4</b>	<b>51,9%</b>	<b>7</b>
Quota de consumo do PIB per capita PPP (n-1)	263,8	20,4%		<b>8277,5</b>	<b>43,5%</b>	<b>6</b>	<b>48453,4</b>	<b>65,4%</b>	<b>3</b>
Quota de Investimento do PIB per capita (n-1)	274,3	21,2%		<b>8430,6</b>	<b>44,3%</b>	<b>4</b>	<b>45737,6</b>	<b>61,7%</b>	<b>6</b>
Utilização de crédito do FMI (n-1)	18,0	1,4%		1197,1	6,3%		2652,0	3,6%	

Apesar de se pretender seleccionar apenas as variáveis com poder explicativo superior a 40%, na escolha dos atributos explicativos do número total de incidentes terroristas foi decidido descer ao nível dos 30%. No primeiro nível foram unicamente apurados três atributos: a inflação, preços de consumo (% anual) (n-1); a quota de consumo do governo sobre o PIB *per capita* (n-1) e o combustíveis, % das exportações de mercadorias (n-1). Logo a seguir na quarta posição estava a variável crescimento do PIB (% anual) (n-1), com uma importância relativa de 33,3%, distanciando-se em cerca de 12 pontos relativamente à posição seguinte. Devido à relevância que este indicador representa no enquadramento da economia de um país, em conjunto com o contexto acima explanado, optou-se por manter este atributo e integrá-lo na construção do modelo final.

A variável combustíveis, % das exportações de mercadorias (n-1) apresentou valores de importância explicativa alta para todas as variáveis dependentes. Nos modelos referentes ao número total de mortos e feridos no primeiro lugar do ranking e para o número total de incidentes terroristas no

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

terceiro lugar. Outra variável de importância transversal a todos os modelos foi a inflação, preços de consumo (% anual) (n-1), com importância relativa de 100% na explicação do número total de incidentes terroristas, 99% na explicação do número total de mortos e 87,5% na explicação do número total de feridos.

Em segundo lugar na medida de importância explicativa do número total de incidentes terroristas encontra-se a quota de consumo do governo sobre o PIB *per capita* (n-1), com 51,9%, uma redução considerável comparativamente com o primeiro lugar. No modelo referente ao número total de mortos assume uma importância relativa de 26,6%, ficando desta forma excluída da construção do respectivo modelo final. Contudo, para o número total de feridos tem exactamente a mesma importância que assume no modelo referente ao número total de incidentes terroristas, apesar de no respectivo ranking ocupar o sétimo lugar.

Em contrário às expectativas, o indicador crescimento do PIB (% anual) (n-1) não demonstrou ter forte relevância como variável explicativa em qualquer um dos modelos. Apesar dos valores de importância apurados não serem muito diferentes na explicação de cada um dos modelos, apenas será considerado para o estudo da primeira variável dependente, da qual resultaram menos atributos explicativos com valores acima de 40%.

O investimento estrangeiro directo, entradas líquidas (% do PIB) (n-1) assume uma importância relativa bastante significativa para a explicação do número total de mortos e de feridos, ocupando respectivamente o terceiro (61,3%) e quarto lugar (64,9%) no ordenamento da tabela. Pelo contrário no modelo referente ao número total de incidentes terroristas detém apenas uma capacidade explicativa de 7,6%, um valor bastante inferior

O PIB *per capita* (n-1) na explicação do número total de incidentes terroristas tem uma importância relativa de 20,2%, não sendo por isso apurada para a construção do modelo final. Para as outras duas variáveis dependentes assume o quinto lugar no ranking, com 44,3% para o número total de mortos e 64% para os feridos. As quotas de consumo e investimento do PIB *per capita* (n-1) são também seleccionadas para a construção dos modelos explicativos finais do número total de mortos e número total de feridos, com 43,5% e 44,3% para o primeiro e 65,4% e 61,7% para o segundo, não tendo demonstrado importância significativa na explicação do número total de incidentes terroristas.

Numa visão global verifica-se que existe entre os três modelos um progressivo aumento de variáveis com maior poder explicativo sobre as variáveis dependentes. Para o número total de incidentes terroristas encontram-se apenas quatro atributos com importância relativa até 30% na explicação do modelo, descido o parâmetro em 10 pontos percentuais, sendo o valor posterior de 21,2%. Detém igualmente uma diferença no poder explicativo entre o primeiro e o segundo lugar de aproximadamente 48%, o que expõe a elevada capacidade explicativa da primeira comparativamente com as restantes variáveis. Para o número total de mortos, as duas primeiras variáveis na ordenação

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

têm uma capacidade explicativa bastante elevada, a atingir os 100% e 99%, assumindo as quatro variáveis seguintes valores entre os 61,3% e os 43,5%. Neste modelo num ranking até 40% seleccionaram-se seis variáveis. Na explicação do número total de feridos, dentro do mesmo nível dos 40%, apuraram-se sete variáveis com importâncias relativas entre 100% e 51,9%.

Para a determinação das correlações existentes entre as variáveis da dimensão sociodemográfica foram aplicados os coeficientes de correlação de *Pearson* e *Spearman*, o primeiro para o apuramento entre variáveis quantitativas e o segundo para a avaliação entre variáveis quantitativas e qualitativas ordinais. Numa primeira abordagem destacou-se a evidente correlação entre as variáveis rácio de mortalidade materna (por 100.000 nascidos vivos) (*Pearson* -0,864;  $p < 0,001$ ), sub-5 taxa de mortalidade (por 1.000 nascidos vivos) (*Pearson* -0,934;  $p < 0,001$ ), taxa de mortalidade neonatal (por 1.000 nascimentos) (*Pearson* -0,898;  $p < 0,001$ ), taxa de mortalidade pós-neonatal (por 1.000 nascimentos) (*Pearson* -0,923;  $p < 0,001$ ) e a variável expectativa de vida ao nascer, total (anos). Perante esta forte correlação entre estas quatro variáveis e esta última, que com os coeficientes registados indica associar em si só o significado global conferido pelas anteriores, optou-se por eliminar os indicadores relativos às taxas de mortalidade e considerar na análise unicamente o indicador expectativa de vida ao nascer, total (anos).

As variáveis média de anos de educação (mulheres) e média de anos de educação (homens) apresentam um coeficiente de correlação bastante forte (*Pearson* 0,970;  $p < 0,001$ ), contudo por representarem características de análise diferentes, entre mulheres e homens, decidiu-se por manter as duas. A mesma situação se coloca com as variáveis população de 0-14 (% do total) e população 15-64 anos (% do total), que obtiveram um coeficiente *Pearson* de -0,945,  $p < 0,001$ .

Na determinação da importância relativa dos atributos explicativos desta dimensão na explicação de cada uma das variáveis dependentes foi apurado, entre os 100% e 40%, um maior número de variáveis comparativamente com a dimensão anterior, concretamente mais onze (Tabela 10).

A densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1) constitui o atributo de maior importância na explicação das três variáveis dependentes, num valor de 100%. O crescimento populacional (% anual) (n-1) na explicação do número de incidentes terroristas assume o segundo lugar, com uma importância relativa de 75,5%, tendo igualmente relevância explicativa relativamente ao número total de mortos e de feridos, mas num peso inferior, de 52,5% e 40,9% respectivamente, ocupando a sétima e oitava posições. A população, total (n-1) assume menor valor explicativo nos três modelos relativamente às anteriores, mantendo-se em análise unicamente para a construção dos modelos finais do número total de incidentes terroristas e número total de mortos, detendo na presente fase uma importância relativa de 61,2% e 41,7% respectivamente.

O atributo população com 65 anos e acima (% do total) (n-1) também apresenta significativa capacidade explicativa transversal a todos os modelos, num valor que varia entre os 63,3% e os

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

75,2%. A população 15-64 anos (% do total) (n-1) pelo contrário não detém peso superior a 40% em qualquer um dos modelos. Por outro lado, a população de 0-14 (% do total) (n-1) está posicionada em terceiro na ordenação referente à importância explicativa do número de mortos com um valor de 65,9%, e em nono para o número total de incidentes terroristas, com 52,1%. Como atributo explicativo do número total de feridos assume apenas a importância relativa de 31,5%, não sendo por isso incluída na construção do modelo final.

A fragmentação étnica (n-1) revela uma boa capacidade explicativa de todas as variáveis dependentes, num valor de 73,7% para o número total de incidentes terroristas, de 60,9% para o número total de mortos e de 59,1% para o número total de feridos, assumindo respectivamente a quarta, quinta e terceira posições. A revisão da literatura neste campo revelou resultados opostos entre si, contudo expecta-se que na construção dos modelos finais este atributo mantenha uma capacidade explicativa significativa.

**Tabela 10: Importância relativa dos atributos explicativos da caracterização sociodemográfica dos países para a explicação de cada um dos impactos do terrorismo**

Atributos explicativos	Nº total de incidentes terroristas			Nº total de mortos			Nº total de feridos		
	Importância	Importância relativa	Ordem	Importância	Importância relativa	Ordem	Importância	Importância relativa	Ordem
Crescimento populacional (% anual) (n-1)	1134,4	75,5%	2	11527,8	52,5%	7	41278,5	40,9%	8
Densidade populacional (pessoas por km <sup>2</sup> de área de terra) (n-1)	1503,3	100,0%	1	21969,5	100,0%	1	100866,9	100,0%	1
Despesa total em saúde (% do PIB) (n-1)	620,2	41,3%	10	9603,1	43,7%	9	39802,5	39,5%	
Direitos dos Trabalhadores (n-1)	91,0	6,1%		2508,0	11,4%		8963,3	8,9%	
Direitos Económicos da Mulher (0 a 3) (n-1)	125,5	8,3%		2873,9	13,1%		4282,7	4,2%	
Escolarização, ensino primário (% bruto) (n-1)	468,7	31,2%		5253,6	23,9%		10380,8	10,3%	
Escolarização, ensino secundário (% bruto) (n-1)	430,6	28,6%		4430,8	20,2%		17006,0	16,9%	
Escolarização, ensino terciário (% bruto) (n-1)	326,7	21,7%		7026,1	32,0%		5428,2	5,4%	
Expectativa de vida ao nascer, total (anos) (n-1)	448,8	29,9%		8461,5	38,5%		44724,1	44,3%	6
Fragmentação étnica (n-1)	1107,4	73,7%	4	13386,4	60,9%	5	59627,9	59,1%	3
Fragmentação linguística (n-1)	439,7	29,3%		4908,5	22,3%		31814,1	31,5%	
Linhas telefónicas (por 100 pessoas) (n-1)	131,6	8,8%		9713,0	44,2%	8	42596,9	42,2%	7
Média de Anos de Educação (Homens) (n-1)	337,8	22,5%		7725,4	35,2%		33598,1	33,3%	
Média de Anos Educação (Mulheres) (n-1)	270,2	18,0%		8062,8	36,7%		35373,4	35,1%	
População 15-64 anos (% do total) (n-1)	232,0	15,4%		8588,9	39,1%		30050,2	29,8%	
População com 65 anos e acima (% do total) (n-1)	1130,4	75,2%	3	13896,3	63,3%	4	71042,5	70,4%	2
População de 0-14 (% do total) (n-1)	782,9	52,1%	9	14480,3	65,9%	3	31809,6	31,5%	
População urbana (% do total) (n-1)	923,1	61,4%	6	11910,2	54,2%	6	51658,0	51,2%	4
População, total (n-1)	919,9	61,2%	7	9155,8	41,7%	10	38496,7	38,2%	
Taxa de participação laboral, mulheres (% da população do sexo feminino 15+) (n-1)	976,7	65,0%	5	14873,6	67,7%	2	46384,6	46,0%	5
Utilizadores de Internet (por cada 100 pessoas) (n-1)	823,5	54,8%	8	7593,2	34,6%		27244,2	27,0%	

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

A variável população urbana (% do total) (n-1), que pretende determinar o nível de urbanização dos países, também mostrou ter peso relevante na explicação das três variáveis de terrorismo, ocupando o sexto lugar do ranking para o número total de incidentes terroristas e número total de mortos e o quarto para o número total de feridos, assumindo contudo para o primeiro uma capacidade explicativa de 61,4%, para o segundo 54,2% apesar de apresentarem o mesmo posicionamento e para o terceiro um valor de 51,2%.

A taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1), variável escassamente abordada em estudos anteriores, revela uma elevada importância na explicação dos três modelos. Detém importância relativa de 65% para o número total de incidentes terroristas, 67,7% para o número total de mortos e 46% para o número total de feridos, salientando-se que para o segundo surge imediatamente a seguir à variável com maior poder explicativo.

A despesa total em saúde apresenta um valor explicativo aproximado para todos os modelos, contudo apenas se manterá em análise para o número total de incidentes terroristas e número total de mortos, por nestes casos assumir uma importância relativa superior a 40%. A variável linhas telefónicas (por 100 pessoas) (n-1) contribui com significância de 44,2% e 42,2% para a explicação do número total de mortos e número total de feridos, detendo contudo um peso explicativo muito baixo para o número total de incidentes terroristas. Contrariamente, o atributo utilizadores de Internet (por cada 100 pessoas) (n-1) apenas revela importância explicativa significativa para o número total de incidentes terroristas, com 54,8%.

Quanto às cinco variáveis relativas à educação, nenhuma delas revelou valores significantes de capacidade explicativa das variáveis de terrorismo. Os três atributos referentes ao ensino primário, secundário e terciário revelaram maior capacidade explicativa do número total de incidentes terroristas e número total de mortos do que do número total de feridos. A média de anos de educação dos homens e das mulheres mostraram uma aproximada linearidade de valores explicativos ao longo dos três modelos, facto coerente com a revelada anteriormente correlação positiva muito forte.

Na análise das correlações às variáveis da dimensão política com o coeficiente de *Spearman*, a variável nível de democracia surgiu fortemente correlacionada com a democracia institucionalizada (*Spearman* 0,962;  $p < 0,001$ ), com os direitos políticos (*Spearman* -0,971;  $p < 0,001$ ), com as liberdades civis (*Spearman* -0,958;  $p < 0,001$ ), com as restrições executivas (*Spearman* 0,913;  $p < 0,001$ ) e com o status de liberdade (*Spearman* -0,921;  $p < 0,001$ ). A relação negativa registada com os direitos políticos, liberdades civis e status de liberdade justifica-se apenas por estarem delimitados por escalas de ordem inversa, sendo que estas variáveis têm uma correlação muito forte com o nível de democracia por esta implicar um aumento do exercício das liberdades. Com as restrições executivas verifica-se uma correlação positiva muito forte pelo facto de o regime político no poder experienciar mais ou menos restrições de acordo com os níveis de autoridade que detenham. Em contraposição



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

com os regimes autocráticos caracterizados por uma autoridade sem limites, nos regimes democráticos as restrições do governo aumentam quanto mais democráticos eles são. Neste caso, decide-se manter o nível de democracia e eliminar as outras variáveis à excepção do *status* de liberdade que se trata de um indicador global de liberdade, no sentido de que possa ser estudada a sua influência separadamente da democracia.

Nos atributos de caracterização política (Tabela 11) observa-se uma maior capacidade explicativa em mais variáveis no modelo referente ao número de incidentes terroristas do que no número total de mortos e de feridos. De facto, mesmo descendo ao nível dos 30% na importância relativa das variáveis para a explicação do número total de mortos e de feridos, só foi possível apurar três variáveis no número total de mortos e duas no número total de feridos.

O atributo fraccionamento político total (n-1) detém 100% de importância relativa na explicação do número total de incidentes e número total de feridos, assumindo apenas um valor de 20% na explicação do número total de mortos. A durabilidade do regime (n-1) é a segunda variável explicativa de maior peso no modelo referente ao número total de incidentes terroristas, com o valor de 78,3%, e também para o número total de mortos apesar de neste caso assumir unicamente uma importância relativa de 36,5%. O atributo instituições em regime (n-1) detém o terceiro lugar no ranking das variáveis explicativas do número total de incidentes, contando com uma importância relativa de 75,1%, contudo para a explicação do número total de mortos e de feridos revela valores irrelevantes. Tal como nas instituições em regime (n-1), expectava-se observar no nível de democracia (n-1) uma elevada capacidade explicativa para as três variáveis de terrorismo. Na explicação do número total de incidentes terroristas revela uma importância relativa significativa, de 66%, porém no número total de feridos detém um valor de 30,1%, ainda assim o segundo lugar no seu ranking, e no número total de mortos apenas 17,9%. O atributo democracia (n-1) não revelou índices explicativos expressivos em qualquer um dos modelos, para o que poderá ter contribuído o facto de ser apenas uma escala limitada de sim/não.

## Data Mining no Contra-terrorismo

Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

**Tabela 11: Importância relativa dos atributos explicativos da caracterização política dos países para a explicação de cada um dos impactos do terrorismo**

Atributos explicativos	Nº total de incidentes terroristas			Nº total de mortos			Nº total de feridos		
	Importância	Importância relativa	Ordem	Importância	Importância relativa	Ordem	Importância	Importância relativa	Ordem
Autocracia institucionalizada (n-1)	403,1	32,8%		2090,9	10,7%		3013,9	3,9%	
Autodeterminação Eleitoral (n-1)	456,8	37,2%		2644,5	13,5%		2064,0	2,7%	
Competitividade da Participação Política (n-1)	67,3	5,5%		663,2	3,4%		6043,6	7,9%	
Democracia (n-1)	292,7	23,8%		488,3	2,5%		15504,0	20,2%	
Direitos Políticos da Mulher (0 a 3) (n-1)	111,7	9,1%		5571,1	28,5%		12480,2	16,3%	
Durabilidade de regime (n-1)	<b>962,3</b>	<b>78,3%</b>	<b>2</b>	<b>7131,6</b>	<b>36,5%</b>	<b>2</b>	16997,1	22,2%	
Fraccionamento político total (n-1)	<b>1228,8</b>	<b>100,0%</b>	<b>1</b>	3919,7	20,0%		<b>76693,2</b>	<b>100,0%</b>	<b>1</b>
Índice de Direitos de Integridade física (n-1)	<b>562,7</b>	<b>45,8%</b>	<b>5</b>	<b>6353,4</b>	<b>32,5%</b>	<b>3</b>	14274,4	18,6%	
Instituições em Regime (n-1)	<b>923,4</b>	<b>75,1%</b>	<b>3</b>	3249,0	16,6%		19343,4	25,2%	
Liberdade de assembleia e associação (n-1)	127,8	10,4%		1678,0	8,6%		1615,1	2,1%	
Liberdade de expressão (n-1)	205,4	16,7%		1019,7	5,2%		14291,7	18,6%	
Nível de democracia (0 a10) (n-1)	<b>811,0</b>	<b>66,0%</b>	<b>4</b>	3501,9	17,9%		<b>23084,0</b>	<b>30,1%</b>	<b>2</b>
Prisão por Motivos Políticos (n-1)	292,4	23,8%		<b>19559,4</b>	<b>100,0%</b>	<b>1</b>	14802,9	19,3%	
Regulamentação de Participação Política (1 a 5) (n-1)	364,6	29,7%		1647,3	8,4%		2226,7	2,9%	
Status de Liberdade (n-1)	275,1	22,4%		534,8	2,7%		17810,6	23,2%	
Tortura (n-1)	28,4	2,3%		168,5	0,9%		1161,3	1,5%	

O índice de direitos de integridade física (n-1) detém importância explicativa significativa para o número total de incidentes terroristas e número total de mortos, no primeiro com 45,8%, na quinta posição, e no segundo com 32,5%, mas neste caso no terceiro lugar apesar de assumir valor inferior. Este índice trata-se de um indicador composto, constituído pelas variáveis: tortura, execuções extrajudiciais, prisão por motivos políticos e desaparecimento. A variável tortura, avaliada também de forma isolada neste estudo, não representa importância explicativa relevante, não passando os seus valores nos três modelos de 2,3%. Por outro lado, a prisão por motivos políticos (n-1) assume uma importância relativa de 100% na explicação do número total de mortos. Para o número total de incidentes e feridos apresenta valores inferiores e bastante menos relevantes na ordem dos 23,8% e 19,3% respectivamente.

Os indicadores relativos a liberdade, por estarem altamente correlacionados com o nível de democracia esperavam-se possuidores de importâncias explicativas mais altas. A liberdade de assembleia e associação (n-1) e a liberdade de expressão (n-1) revelaram coeficientes de correlação com o nível de democracia na ordem dos 70% e o *status* de liberdade (n-1) acima de 90%.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Na análise das correlações entre as variáveis de caracterização religiosa, através da aplicação do coeficiente *Pearson*, nenhuma apresentou coeficiente com registo acima do moderado (máximo *Pearson* -0,520;  $p < 0,001$ ), verificado entre as variáveis religião muçulmana (percentagem da população total) (n-1) e outras religiões (percentagem da população total) (n-1).

As variáveis de caracterização religiosa (Tabela 12) apresentam na sua maioria elevada capacidade explicativa sobre as variáveis de terrorismo. A religião muçulmana (percentagem da população total) (n-1) representa uma importância relativa de 100% na explicação das três variáveis dependentes. A fragmentação religiosa (n-1) ocupa o segundo lugar na explicação do número total de incidentes terroristas, com 95,5%, e também na explicação do número total de feridos, com 87,7%. No número total de mortos a religião católica (percentagem da população total) (n-1) é a segunda variável com maior importância explicativa, com 80,8%.

O terceiro lugar é divergente nos três modelos, sendo ocupado pela variável outras religiões (percentagem da população total) (n-1) com 93,6% na explicação do número total de incidentes terroristas, pela fragmentação religiosa (n-1) no número total de mortos com 70,4% e pela religião católica (percentagem da população total) (n-1) na explicação do número total de feridos, com 80,4%. No número total de feridos a quarta variável já não demonstrou capacidade explicativa acima dos 40%, não sendo por isso incluída na construção do modelo final. Para o número total de incidentes terroristas a quarta posição na ordenação das variáveis foi ocupada pela religião católica (percentagem da população total) (n-1), com 67,4% e para o número total de mortos pela variável outras religiões (percentagem da população total) (n-1), com 48,2%.

Para o número total de incidentes terroristas as três posições com maior peso explicativo assumem uma importância relativa acima dos 93%, perfazendo a variável de menor capacidade explicativa o valor de 67,4%. No número total de mortos há uma diferença superior entre a variável de maior importância relativa, de 100%, e a de menor valor, de 48,2%. No número total de feridos três variáveis detêm valores de importância relativa superiores a 80%, tendo a variável seguinte um valor bastante inferior, de 24,5%.

**Tabela 12: Importância relativa dos atributos explicativos da caracterização religiosa dos países para a explicação de cada um dos impactos do terrorismo**

Atributos explicativos	Nº total de incidentes terroristas			Nº total de mortos			Nº total de feridos		
	Importância	Importância relativa	Ordem	Importância	Importância relativa	Ordem	Importância	Importância relativa	Ordem
Fragmentação religiosa (n-1)	961,5	95,5%	2	8390,8	70,4%	3	44170,5	87,7%	2
Outras religiões (percentagem da população total) (n-1)	941,6	93,6%	3	5750,9	48,2%	4	12352,9	24,5%	
Religião Católica (percentagem da população total) (n-1)	678,0	67,4%	4	9639,8	80,8%	2	40505,8	80,4%	3
Religião Muçulmana (percentagem da população total) (n-1)	1006,4	100,0%	1	11925,8	100,0%	1	50375,8	100,0%	1

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Quanto à dimensão referente às outras variáveis de caracterização dos países, verifica-se que as variáveis relativas à área estão bastante relacionadas, estabelecendo entre si uma correlação positiva perfeita ( $Pearson= 1,000$ ;  $p<0,001$ ). Apenas diferem pela inclusão ou não dos terrenos cobertos por água e assim inabitáveis, optando-se desta forma pela manutenção da variável referente à área geográfica total dos países, eliminando-se a outra.

Os restantes coeficientes de correlação calculados revelaram valores de relação muito fracos, mantendo-se as restantes variáveis em estudo.

Relativamente aos resultados obtidos na geração das árvores parciais para esta dimensão (Tabela 13), agora com três variáveis, constata-se um elevado peso explicativo da área de superfície (km<sup>2</sup>) - área total geográfica do país (n-1), com importância relativa de 100% em todos os modelos. O segundo lugar para as três variáveis de terrorismo é ocupado pelo KOF - Índice de Globalização (0 a 100) (n-1), para o número total de incidentes terroristas com uma importância relativa de 56,3%, para o número total de mortos de 60,2% e número total de feridos com 81,9%, revelando nesta última uma capacidade explicativa bastante superior à assinalada no primeiro. A variável despesas militares (em % do PIB) (n-1) tem uma importância relativa nos três modelos entre 38% e 44,1%, não difere muito entre eles, mas assume relativamente às outras variáveis de caracterização um poder explicativo de peso inferior, não chegando mesmo a ficar apurada para o modelo final relativo ao número total de mortos.

**Tabela 13: Importância relativa de outros atributos explicativos de caracterização dos países para a explicação de cada um dos impactos do terrorismo**

Atributos explicativos	Nº total de incidentes terroristas			Nº total de mortos			Nº total de feridos		
	Importância	Importância relativa	Ordem	Importância	Importância relativa	Ordem	Importância	Importância relativa	Ordem
Área de superfície (km <sup>2</sup> ) - área total geográfica do país (n-1)	1079,5	100,0%	1	13032,2	100,0%	1	54928,2	100,0%	1
Despesas militares (em % do PIB) (n-1)	456,1	42,2%	3	4947,8	38,0%		24203,9	44,1%	3
KOF - Índice de Globalização (0 a 100) (n-1)	607,7	56,3%	2	7851,9	60,2%	2	44969,4	81,9%	2

### 6.3.2 Modelo Explicativo do Número Total de Incidentes Terroristas

Na construção dos modelos explicativos foram seleccionadas as variáveis apuradas no ponto anterior e testadas várias árvores de acordo com regras de parametrização diferentes, no sentido de encontrar o modelo que consista na melhor conjugação entre capacidade explicativa em relação à amostra e capacidade preditiva na aplicação a novos dados.

Na fase de geração dos modelos ainda se incluiu como possível atributo explicativo a região, que perfaz um total de treze, pelas quais se distribuem os 173 países em análise, analisada em detalhe nos pontos 6.1 e 6.2. do presente capítulo.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Na avaliação aos resultados do modelo para o número total de incidentes terroristas, apresentados na Tabela 14, foi escolhido o modelo 3 por espelhar a melhor e mais aproximada disposição de valores. No sentido de conhecer os factores determinantes no fomento do terrorismo pretende-se um modelo com uma boa capacidade explicativa da amostra, contudo um sobreajustamento aos dados reduz o seu poder preditivo. A árvore apurada detém como regras de paragem um nível máximo de profundidade de cinco níveis abaixo do nó raiz, um número mínimo de casos nos nós *pais* de quatro e nos nós *filhos* de dois e um ganho mínimo na qualidade do modelo (decréscimo na variância da variável dependente resultando de uma nova ramificação) de 0,0001. A proporção de variância explicada pelo modelo para a totalidade da amostra é de 72,6% e medida através de validação cruzada (10-*fold*) é de 55,1%.

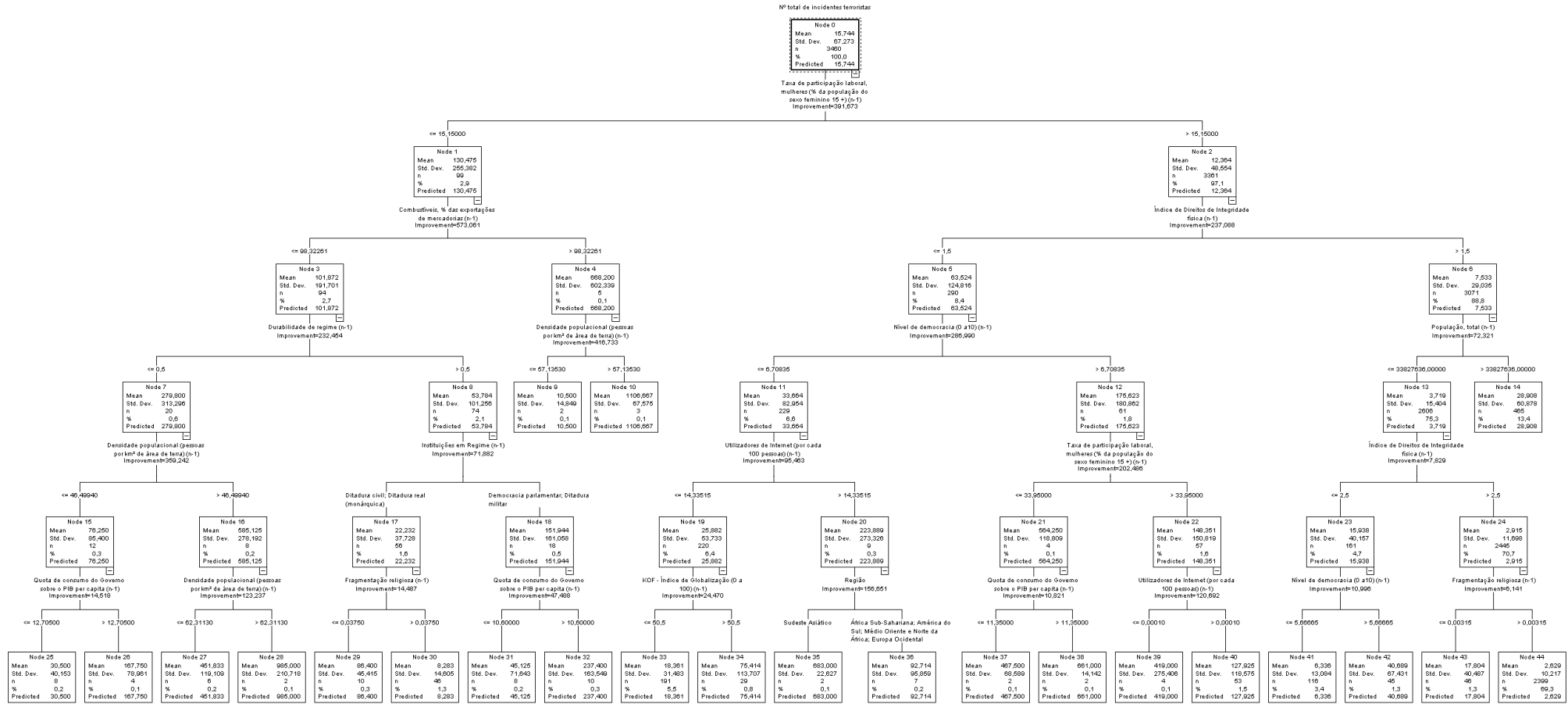
**Tabela 14: Resultados de vários modelos para o número total de incidentes terroristas**

Parametrização	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6	Modelo 7	Modelo 8
N.º máximo de níveis da árvore	6	6	5	5	4	4	3	3
N.º mínimo de observações em nó <i>pai</i>	4	2	4	2	4	2	4	2
N.º mínimo de observações em nó <i>filho</i>	2	1	2	1	2	1	2	1
% de variância explicada (para toda a amostra)	76,5	84,9	72,6	75,1	61,6	63,2	45,8	45,8
% de variância explicada (Validação Cruzada)	47,0	15,3	55,1	16,5	31,6	29,9	25,4	27,9

# Data Mining no Contra-terrorismo

## Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

**Figura 5: Árvore de regressão do número total de incidentes terroristas**



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

A árvore obtida (Figura 5) apresenta 45 nós, dos quais 23 são nós folha. A importância relativa das diferentes variáveis explicativas é apresentada na Tabela 15. A variável densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1) detém uma importância explicativa no modelo de 100%, isto é, é o factor com maior capacidade explicativa do número de incidentes terroristas. A taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) é a segunda variável com maior poder explicativo (75,8%) e constitui a primeira ramificação da árvore, correspondendo um ramo a uma taxa menor ou igual a 15,15% da população do sexo feminino com mais de quinze anos (nó 1) e o outro a uma taxa superior a 15,15% (nó 2). No nó 1 os países/anos tendem a apresentar uma média superior de incidentes terroristas (130,475) e o nó 2 uma média bastante menor, de 12,364 incidentes. O terceiro e o quarto nós em que se ramifica o nó 1 têm por base a variável combustíveis, % das exportações de mercadorias (n-1), que ocupa o terceiro lugar no *ranking* com 73%. Países que em determinado ano tenham exportações de combustíveis superiores a 98,32% e que tenham uma taxa de participação laboral das mulheres menor ou igual a 15,15% da população do sexo feminino com mais de quinze anos tendem a experienciar em média 668,2 incidentes terroristas, enquanto que a valores iguais ou inferiores a essa percentagem e com uma taxa de participação laboral das mulheres menor ou igual a 15,15% da população do sexo feminino com mais de quinze anos tende a corresponder uma média de 101,872 incidentes.

O nó 2 ramifica-se nos nós 5 e 6 que se referem ao índice de direitos de integridade física (n-1), o primeiro para valores inferiores ou iguais a 1,5 (de 0 - nenhum, a 8 - total) e o segundo para valores superiores a 1,5. Países que em determinado ano tenham índice de direitos de integridade física inferior ou igual a 1,5 e taxa de participação laboral das mulheres superior a 15,15% da população do sexo feminino com mais de quinze anos podem vir a ser alvo de uma média prevista de 63,52 incidentes terroristas, enquanto que com índice de direitos de integridade física de valor superior a 1,5 e taxa de participação laboral das mulheres superior a 15,15% da população do sexo feminino com mais de quinze anos têm tendência a uma média de 7,53 incidentes terroristas.

Destacam-se as variáveis densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1) e quota de consumo do governo sobre o PIB *per capita* (n-1) por surgirem mais vezes com ramificadoras, respectivamente nos nós 4,7 e 16 e nos nós 15, 18 e 21.

De uma forma sintetizada, a avaliação ao modelo gerado permite dizer por exemplo que se taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) for menor ou igual a 15,15% e a percentagem de exportações de combustíveis (% das exportações de mercadorias) (n-1) for menor ou igual a 98,32 e a durabilidade de regime for superior a meio ano e se vigorar ditadura civil ou real (monárquica) e verificar valores de fragmentação religiosa superiores a 0,0375 (de 0 a 1), então a média de incidentes previstos é de 8,283 (nó 30). Por outro lado, se um país pertencer à região do Sudeste Asiático e tiver mais de 14,34 utilizadores de internet (por cada 100 pessoas (n-1) e nível de democracia inferior ou igual a 6,71 (de 0 - baixo a 10 - elevado) e verificar índice de direitos

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

de integridade física (n-1) inferior ou igual a 1,5 (de 0 - nenhum, a 8 - total) e taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) superior a 15,15%, então a média de incidentes previstos é de 683 (nó 35). A média de incidentes terroristas prevista no modelo com o segundo valor mais alto é de 985 (nó 28), que poderá resultar de um contexto de densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1) superior a 62,311 e com durabilidade de regime político inferior a meio ano e com percentagem de exportações de combustíveis (% das exportações de mercadorias) (n-1) inferior ou igual a 98,32% e taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) inferior ou igual a 15,15%.

A média de incidentes terroristas prevista mais baixa é de 2,629 (nó 44) e verifica-se no caso de a fragmentação religiosa (n-1) ser maior que 0,00315 (0 a 1) e o índice de direitos de integridade física (n-1) ser superior a 2,5 (de 0 - nenhum, a 8 - total) e população total ser menor ou igual a 33.827.636 pessoas e taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) superior a 15,15%.

**Tabela 15: Importância das variáveis explicativas para o modelo do número total de incidentes terroristas**

Atributo Explicativo	Dimensão	Importância	Importância Relativa
Densidade populacional (pessoas por km <sup>2</sup> de área de terra) (n-1)	Sociodem.	1480,648	<b>100,0%</b>
Taxa de participação laboral, mulheres (% população do sexo feminino 15 +) (n-1)	Sociodem.	1122,643	<b>75,8%</b>
Combustíveis, % das exportações de mercadorias (n-1)	Económ.	1081,073	<b>73,0%</b>
População, total (n-1)	Sociodem.	1074,439	<b>72,6%</b>
Instituições em Regime (n-1)	Política	972,757	<b>65,7%</b>
Durabilidade de regime (n-1)	Política	866,653	<b>58,5%</b>
Fraccionamento político total (n-1)	Política	837,125	<b>56,5%</b>
KOF - Índice de Globalização (0 a 100) (n-1)	Outras	766,773	<b>51,8%</b>
Área de superfície (km <sup>2</sup> ) - área total geográfica do país (n-1)	Outras	765,744	<b>51,7%</b>
Quota de consumo do Governo sobre o PIB per capita (n-1)	Económ.	737,137	49,8%
Índice de Direitos de Integridade física (n-1)	Política	694,027	46,9%
Religião Muçulmana (percentagem da população total) (n-1)	Religiosa	679,880	45,9%
Inflação, preços de consumo (% anual) (n-1)	Económ.	678,946	45,9%
Outras religiões (percentagem da população total) (n-1)	Religiosa	678,785	45,8%
Fragmentação religiosa (n-1)	Religiosa	619,524	41,8%
Crescimento do PIB (% anual) (n-1)	Económ.	618,748	41,8%
Religião Católica (percentagem da população total) (n-1)	Religiosa	614,544	41,5%
População urbana (% do total) (n-1)	Sociodem.	581,742	39,3%
Região	Outras	551,067	37,2%
População de 0-14 (% do total) (n-1)	Sociodem.	542,464	36,6%
Nível de democracia (0 a 10) (n-1)	Política	539,476	36,4%
Utilizadores de Internet (por cada 100 pessoas) (n-1)	Sociodem.	525,129	35,5%
Despesa total em saúde (% do PIB) (n-1)	Sociodem.	523,122	35,3%
Fragmentação étnica (n-1)	Sociodem.	455,391	30,8%
Crescimento populacional (% anual) (n-1)	Sociodem.	415,093	28,0%
População com 65 anos e acima (% do total) (n-1)	Sociodem.	333,703	22,5%
Despesas militares (em % do PIB) (n-1)	Outras	272,215	18,4%



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

O nó 10 apresenta o valor médio de incidentes terroristas mais elevado, de 1106,667, que surge na sequência de uma taxa de participação laboral, mulheres (% da população do sexo feminino 15+) (n-1) inferior ou igual a 15,15% e de uma percentagem de exportações de combustíveis (% das exportações de mercadorias) (n-1) superior a 98,32% e de uma densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1) superior a 57,135.

Ainda se destacam com importâncias relativas acima de 50% as variáveis população, total (n-1), com uma importância de 72,6%; instituições em regime (n-1), com importância de 65,7%; durabilidade de regime (n-1), com importância de 58,5%; fraccionamento político total (n-1), com importância de 56,5%; KOF - índice de globalização (0 a 100) (n-1), com importância de 51,8% e área de superfície (km<sup>2</sup>) - área total geográfica do país (n-1), com importância de 51,7%.

### 6.3.3 Modelo Explicativo do Número Total de Mortos

Para a concepção do modelo explicativo do número total de mortos usou-se a mesma metodologia do ponto anterior. Foram gerados oito modelos com parametrizações diferentes, no sentido de encontrar o modelo que detenha as melhores percentagens de variância explicativa. A Tabela 16 apresenta as parametrizações de cada modelo e as respectivas percentagens de variância explicadas, tanto para a totalidade da amostra, como no resultado da validação cruzada (10-fold). O modelo escolhido foi o número seis por apresentar uma boa capacidade explicativa para a amostra (76,4%), em conjunto com a melhor percentagem de capacidade explicativa resultante da validação cruzada (39,5%). A árvore apurada detém como regras de paragem um nível máximo de profundidade de quatro níveis abaixo do nó raiz, um número mínimo de casos nos nós *pai* de dois e nos nós *filhos* de um e um ganho mínimo na qualidade do modelo (decréscimo na variância da variável dependente resultando de uma nova ramificação) de 0,0001.

**Tabela 16: Resultados de vários modelos para o número total de mortos**

Parametrização	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6	Modelo 7	Modelo 8
N.º máximo de níveis da árvore	6	6	5	5	4	4	3	3
N.º mínimo de observações em nó <i>pai</i>	4	2	4	2	4	2	4	2
N.º mínimo de observações em nó <i>filho</i>	2	1	2	1	2	1	2	1
% de variância explicada (para toda a amostra)	73,1	88,7	70,9	85,1	68,7	76,4	57,1	57,1
% de variância explicada (Validação Cruzada)	21,8	-16,6	12,4	12,4	14,5	39,5	27,0	7,1

A árvore obtida (Figura 6) apresenta 31 nós, dos quais 16 são nós folha. A importância relativa das diversas variáveis explicativas é apresentada na Tabela 17. Neste modelo a variável combustíveis, % das exportações de mercadorias (n-1) é o factor com maior capacidade explicativa, com uma importância relativa de 100%, sendo imediatamente seguida da densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1) com um valor de 95,8%. Na faixa dos 80% encontram-se a população urbana (% do total) (n-1) com uma importância de 87,7% e a fragmentação religiosa (n-1) com uma importância de 81,3%. As restantes variáveis de caracterização religiosa também se encontram entre

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

as variáveis de maior importância explicativa relativa, com valores entre 61,6% e 62,8%. A taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) neste modelo aparece no quinto lugar do *ranking* com uma importância de 67,9%, mas constitui similarmente a primeira ramificação da árvore, subdividindo-se no nó 1 para valores iguais ou inferiores a 14,45% e no nó 2 para valores superiores a 14,45%, correspondendo a este último um valor médio previsto de mortos de valor inferior ao anterior. Com uma importância relativa de 64,8% ainda se destaca a área de superfície (km<sup>2</sup>) - área total geográfica do país (n-1). A inflação, preços de consumo (% anual) (n-1) é a segunda variável de caracterização económica com maior capacidade explicativa, com uma importância de 61,6%. Com valores de importância relativa acima de 50% ainda se encontram as variáveis durabilidade de regime (n-1), população de 0-14 (% do total) (n-1), fragmentação étnica (n-1) e a região. Das 26 variáveis consideradas na construção deste modelo, 19 têm importância explicativa relativa superior a 40%.

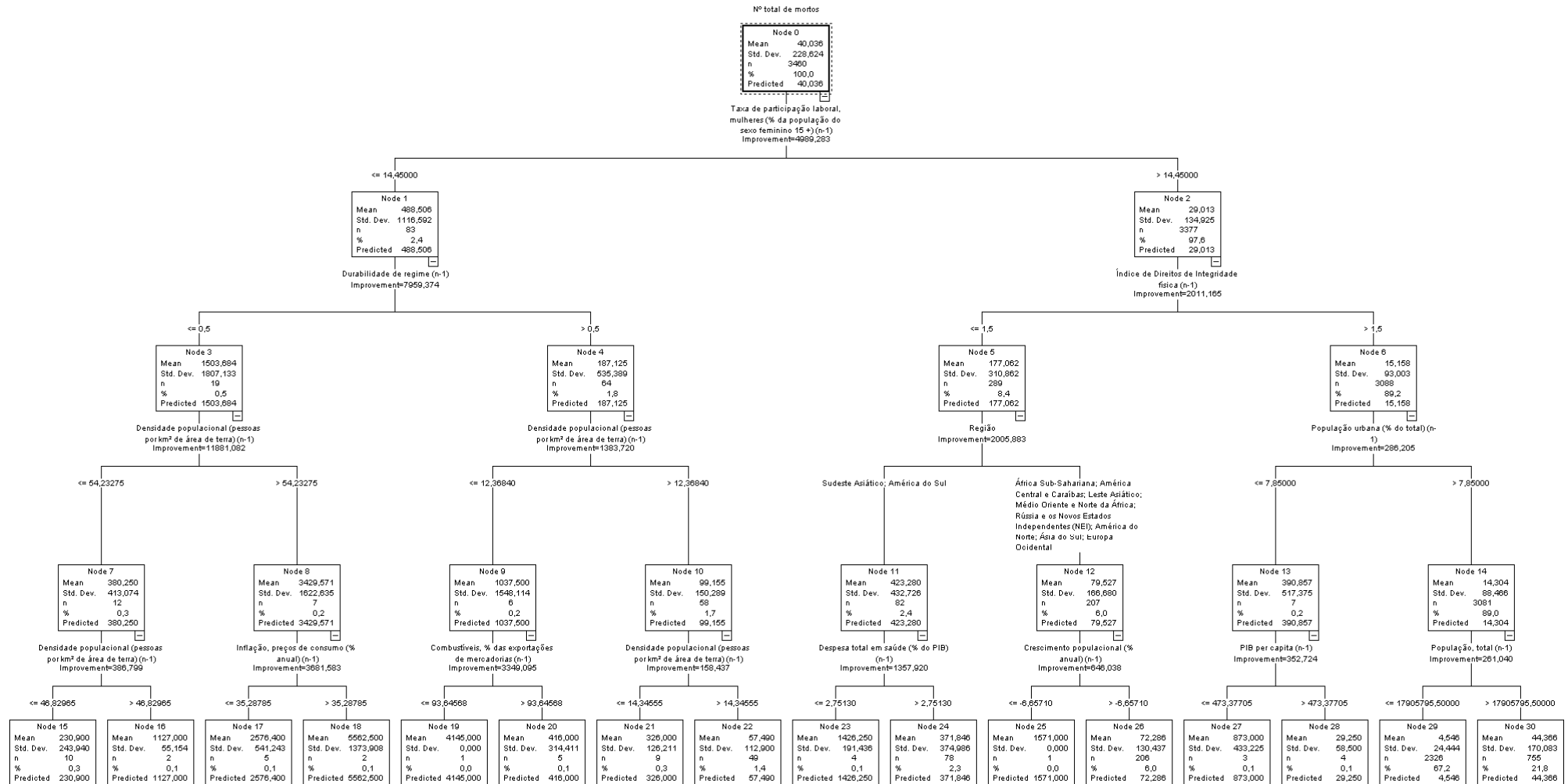
**Tabela 17: Importância das variáveis explicativas para o modelo do número total de mortos**

Atributo Explicativo	Dimensão	Importância	Importância Relativa
Combustíveis, % das exportações de mercadorias (n-1)	Económ.	15044,199	<b>100,0%</b>
Densidade populacional (pessoas por km <sup>2</sup> de área de terra) (n-1)	Sociodem.	14418,931	<b>95,8%</b>
População urbana (% do total) (n-1)	Sociodem.	13190,428	<b>87,7%</b>
Fragmentação religiosa (n-1)	Religiosa	12234,117	<b>81,3%</b>
Taxa participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1)	Sociodem.	10207,710	<b>67,9%</b>
Área de superfície (km <sup>2</sup> ) - área total geográfica do país (n-1)	Outras	9744,854	<b>64,8%</b>
Outras religiões (percentagem da população total) (n-1)	Religiosa	9444,683	<b>62,8%</b>
Religião Católica (percentagem da população total) (n-1)	Religiosa	9410,705	<b>62,6%</b>
Religião Muçulmana (percentagem da população total) (n-1)	Religiosa	9268,403	<b>61,6%</b>
Inflação, preços de consumo (% anual) (n-1)	Económ.	9259,859	<b>61,6%</b>
Durabilidade de regime (n-1)	Política	8943,934	<b>59,5%</b>
População de 0-14 (% do total) (n-1)	Sociodem.	8231,527	<b>54,7%</b>
Fragmentação étnica (n-1)	Sociodem.	8080,963	<b>53,7%</b>
Região	Outras	7798,580	<b>51,8%</b>
Despesa total em saúde (% do PIB) (n-1)	Sociodem.	7415,426	49,3%
Investimento estrangeiro directo, entradas líquidas (% do PIB) (n-1)	Económ.	7263,092	48,3%
Crescimento populacional (% anual) (n-1)	Sociodem.	6955,963	46,2%
KOF - Índice de Globalização (0 a 100) (n-1)	Outras	6209,127	41,3%
População com 65 anos e acima (% do total) (n-1)	Sociodem.	6033,881	40,1%
Linhas telefónicas (por 100 pessoas) (n-1)	Sociodem.	4691,009	31,2%
Quota de consumo do PIB per capita PPP (n-1)	Económ.	4621,178	30,7%
Índice de Direitos de Integridade física (n-1)	Política	4087,538	27,2%
PIB per capita (n-1)	Económ.	3341,738	22,2%
População, total (n-1)	Sociodem.	1842,714	12,2%
Quota de Investimento do PIB per capita (n-1)	Económ.	1742,151	11,6%
Prisão por Motivos Políticos (n-1)	Política	980,810	6,5%

# Data Mining no Contra-terrorismo

## Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

**Figura 6: Árvore de regressão do número total de mortos**



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Do nó 1, a que corresponde a taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) de valores iguais ou inferiores a 14,45%, surge ramificação referente à durabilidade do regime, no nó 3 menor ou igual a meio ano e no nó 4 para valores superiores a meio ano. Estes valores sugestionam que regimes de duração de meio ano ou menos e com taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) de valores iguais ou inferiores a 14,45% têm uma média de mortos prevista de 1503,864 (nó 3), superior à média prevista no nó 4 (187,125) na situação de regimes de duração superior a meio ano e com taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) de valores superiores a 14,45%.

O nó 2 ramifica-se nos nós 5 e 6 que se referem ao índice de direitos de integridade física (n-1), o primeiro para valores inferiores ou iguais a 1,5 (de 0 - nenhum, a 8 - total) e o segundo para valores superiores a 1,5. Países que em determinado ano tenham índice de direitos de integridade física inferior ou igual a 1,5 e taxa de participação laboral das mulheres superior a 14,45% da população do sexo feminino com mais de quinze anos podem vir a ser alvo de uma média prevista de 177,062 mortos (nó 5), enquanto que com índice de direitos de integridade física de valor superior a 1,5 e taxa de participação laboral das mulheres superior a 14,45% da população do sexo feminino com mais de quinze anos têm tendência a uma média de 15,158 mortos (nó 6).

A densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1), a segunda no *ranking* das importâncias relativas das variáveis explicativas, é também a variável que surge em mais ramificações, nos nós 3, 4, 7 e 10.

A variável combustíveis, % das exportações de mercadorias (n-1) que se ramifica para os nós 19 e 20, o primeiro para valores inferiores ou iguais a 93,646% e o segundo para valores superiores a 93,646%, apresenta entre os dois nós a mais destacada diferença na média prevista de mortos. Se um país em determinado ano tiver uma taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) de valores iguais ou inferiores a 14,45% e durabilidade do regime político superior a meio ano e densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1) inferior ou igual a 12,368 e percentagem de vendas de combustíveis (% das exportações de mercadorias) (n-1) inferior ou igual a 93,656, então a média de mortos prevista é de 4145 (nó 19). Por outro lado, se esse mesmo país apresentar ao invés um valor de vendas de combustíveis (% das exportações de mercadorias) (n-1) superior a 93,646%, a média prevista passa para 416 mortos (nó 20).

O valor médio previsto de mortos mais alto surge no nó 18 (5562,5), o qual poderá resultar no caso de um país/ano com uma taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) de valor igual ou inferior a 14,45% e durabilidade do regime político inferior ou igual a meio ano e densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1) superior a 54,233 e inflação, preços de consumo (% anual) (n-1) superior a 35,288%. No nó 29 está o valor médio previsto de mortos mais reduzido, de 4,546, que surge na sequência de uma conjuntura de uma população total

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

inferior ou igual a 17.905.795,5 pessoas e valor de população urbana (% do total) (n-1) superior a 7,85% e índice de direitos de integridade física (n-1) de valor superior a 1,5 (de 0 - nenhum, a 8 - total) e taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) de valor superior a 14,45%. Nesta mesma sequência mas para uma população total superior a 17.905.795,5 pessoas, figura no nó 30 o número médio de mortos previsto, com o valor de 44,366, mais aproximado da média global (40,036).

#### 6.3.4 Modelo Explicativo do Número Total de Feridos

Em conformidade com os pontos anteriores foram gerados oito modelos com parametrizações diferentes, no sentido de encontrar o modelo que detenha as melhores percentagens de variância explicativa. A Tabela 18 apresenta as parametrizações de cada modelo e as respectivas percentagens de variância explicadas, tanto para a totalidade da amostra, como no resultado da validação cruzada (10-*fold*). Neste caso, o modelo escolhido foi o número três por apresentar a melhor combinação entre a capacidade explicativa para a amostra (59,7%) e a capacidade explicativa resultante da validação cruzada (36,1%). A árvore apurada detém como regras de paragem um nível máximo de profundidade de cinco níveis abaixo do nó raiz, um número mínimo de casos nos nós *pais* de quatro e nos nós *filhos* de dois e um ganho mínimo na qualidade do modelo (decréscimo na variância da variável dependente resultando de uma nova ramificação) de 0,0001.

**Tabela 18: Resultados de vários modelos para o número total de feridos**

Parametrização	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6	Modelo 7	Modelo 8
N.º máximo de níveis da árvore	6	6	5	5	4	4	3	3
N.º mínimo de observações em nó <i>pai</i>	4	2	4	2	4	2	4	2
N.º mínimo de observações em nó <i>filho</i>	2	1	2	1	2	1	2	1
% de variância explicada (para toda a amostra)	61,3	84,0	59,7	79,8	57,0	65,3	50,4	52,7
% de variância explicada (Validação Cruzada)	13,5	4,1	36,1	0,4	20,8	7,7	8,9	3,4

A árvore obtida (Figura 7) apresenta 27 nós, dos quais 14 são nós folha. A importância relativa das diversas variáveis explicativas é apresentada na Tabela 19. A variável com maior capacidade explicativa do número de feridos, com importância relativa de 100%, é a quota de consumo do governo sobre o PIB *per capita* (n-1), constituindo as ramificações dos nós 3 e 7. A taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) é a segunda variável explicativa com importância relativa mais alta, de 85,2%, com origem no nó 1 e dividindo-se no nó 3 para valores menores ou iguais a 14,45% e no nó 4 para valores superiores a 14,45%. A região com 76,7% ocupa o terceiro lugar no ranking e aparece na árvore nas ramificações dos nós 2 e 4. A fragmentação étnica (n-1), a religião católica (percentagem da população total) (n-1) e os combustíveis, % das exportações de mercadorias (n-1) estão ainda na faixa dos 70%, esta última constituindo a primeira ramificação da árvore, na origem dos nós 1 e 2, o primeiro para percentagens iguais ou abaixo de 99,519 e o segundo para percentagens superiores a 99,519. A área de superfície

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

(km<sup>2</sup>) - área total geográfica do país (n-1), que apenas apresenta uma importância relativa de 12,7%, é a variável ramificadora de maior relevo surgindo nos nós 9,10 e 15.

**Tabela 19: Importância das variáveis explicativas para o modelo do número total de feridos**

Atributo Explicativo	Dimensão	Importância	Importância Relativa
Quota de consumo do Governo sobre o PIB per capita (n-1)	Económ.	52146,214	<b>100,0%</b>
Taxa participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1)	Sociodem.	44408,087	<b>85,2%</b>
Região	Outras	39999,814	<b>76,7%</b>
Fragmentação étnica (n-1)	Sociodem.	39206,885	<b>75,2%</b>
Religião Católica (percentagem da população total) (n-1)	Religiosa	38305,998	<b>73,5%</b>
Combustíveis, % das exportações de mercadorias (n-1)	Económ.	37849,433	<b>72,6%</b>
Religião Muçulmana (percentagem da população total) (n-1)	Religiosa	35790,286	<b>68,6%</b>
População com 65 anos e acima (% do total) (n-1)	Sociodem.	35759,533	<b>68,6%</b>
Inflação, preços de consumo (% anual) (n-1)	Económ.	33595,536	<b>64,4%</b>
Fraccionamento político total (n-1)	Política	26476,000	<b>50,8%</b>
População urbana (% do total) (n-1)	Sociodem.	22823,752	43,8%
KOF - Índice de Globalização (0 a 100) (n-1)	Outras	20746,508	39,8%
Quota de Investimento do PIB per capita (n-1)	Económ.	19611,000	37,6%
Crescimento populacional (% anual) (n-1)	Sociodem.	19145,682	36,7%
Investimento estrangeiro directo, entradas líquidas (% do PIB) (n-1)	Económ.	15312,479	29,4%
Área de superfície (km <sup>2</sup> ) - área total geográfica do país (n-1)	Outras	6628,725	12,7%
Despesas militares (em % do PIB) (n-1)	Outras	4756,483	9,1%
Fragmentação religiosa (n-1)	Religiosa	2680,605	5,1%
Expectativa de vida ao nascer, total (anos) (n-1)	Sociodem.	2410,601	4,6%
Nível de democracia (0 a10) (n-1)	Política	536,564	1,0%
Densidade populacional (pessoas por km <sup>2</sup> de área de terra) (n-1)	Sociodem.	415,323	0,8%
PIB per capita (n-1)	Económ.	360,363	0,7%
Quota de consumo do PIB per capita PPP (n-1)	Económ.	193,202	0,4%
Linhas telefónicas (por 100 pessoas) (n-1)	Sociodem.	116,266	0,2%

Do nó 2 descende a variável região, que no nó 5 representa o Leste Asiático e no nó 6 a África Sub-Sahariana e a Rússia e os Novos Estados Independentes. O nó 5 apresenta a média prevista de feridos mais alta, que poderá ocorrer quando um país/ano pertencer ao Leste Asiático e detiver percentagem na exportação de combustíveis (% das exportações de mercadorias) (n-1) superior a 99,519. A média prevista de feridos mais baixa encontra-se por oposição no nó 6, para países/ano que pertençam às regiões da África Sub-Sahariana e da Rússia e Novos Estados Independentes e tenham percentagem de exportação de combustíveis (% das exportações de mercadorias) (n-1) superior a 99,519.

No nó 3 encontra-se uma média de feridos prevista de 656,925, para países/ano que tenham percentagem de exportação de combustíveis (% das exportações de mercadorias) (n-1) inferior ou igual a 99,519 e taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) com valores inferiores ou iguais a 14,45%. Da mesma forma, mas para uma taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1) com valores superiores a 14,45% já se prevê uma média de 41,659 feridos (nó 4).

O nó 15 estabelece uma média prevista de feridos de 24,336 para a grande maioria das observações da amostra, ramificando-se através da variável área de superfície (km<sup>2</sup>) - área total geográfica do país



### 6.3.5 Comparação dos Modelos Explicativos

As percentagens de variância explicadas para a amostra em todos os modelos são significativamente mais altas que as percentagens de variância explicadas pela aplicação da validação cruzada (*10-fold*). A escolha dos modelos procurou encontrar o equilíbrio entre um modelo que ajude a conhecer os determinantes do terrorismo com base na amostra recolhida, e um modelo que possa servir de base para a previsão de ocorrências futuras através da aplicação a novos dados. Contudo, a capacidade de generalização em todos os modelos, principalmente no que se refere ao número total de mortos e ao número total de feridos, não ultrapassou níveis medianos, o que se poderá dever à existência de contextos internos muito diferentes de país para país, ou de alguns outros factores mais específicos que não se tornam possíveis de ser abordados em estudos de carácter global.

A Tabela 20 apresenta as regras de parametrização e resultados dos três modelos, bem como o número de variáveis explicativas de cada dimensão de caracterização dos países com importâncias relativas acima de 40% identificadas em cada um. O modelo relativo ao número total de mortos é o que apresenta melhor percentagem de variância explicada para a amostra, com 76,4%. Para o número total de incidentes terroristas o modelo tem uma capacidade explicativa para a amostra de 72,6% e para o número total de feridos o modelo apresenta apenas uma capacidade explicativa para a amostra de 59,7%, de todos o menor. No que refere à proporção de variância explicada medida através da validação cruzada, o modelo do número total de incidentes terroristas detém o valor mais relevante, de 55,1%. Para o número total de mortos e número total de feridos os valores não ultrapassaram os 40%, perfazendo respectivamente 39,5% e 36,1%.

**Tabela 20: Parametrização e resultados dos três modelos e número de variáveis explicativas acima de 40% de importância relativa em cada dimensão de caracterização dos países**

Parametrização	N.º total incidentes	N.º total mortos	N.º total feridos
N.º máximo de níveis da árvore	5	4	5
N.º mínimo de observações em nó <i>pai</i>	4	2	4
N.º mínimo de observações em nó <i>filho</i>	2	1	2
% de variância explicada (para toda a amostra)	72,6	76,4	59,7
% de variância explicada (Validação Cruzada)	55,1	39,5	36,1
N.º variáveis económicas	4	3	3
N.º variáveis sociodemográficas	3	8	4
N.º variáveis políticas	4	1	1
N.º variáveis de religião	4	4	2
N.º outras variáveis	2	3	1

A variável com maior capacidade explicativa no número de incidentes terroristas é a densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1) com uma importância de 100%. No modelo do número total de mortos assume igualmente uma elevada capacidade explicativa, com uma importância de 95,8%. Pelo contrário, no modelo explicativo do número total de feridos detém apenas uma importância de 0,8%. Neste modelo a variável com maior poder explicativo, 100%, é a quota de



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

consumo do governo sobre o PIB *per capita* (n-1), a qual demonstrou no modelo do número total de incidentes uma capacidade explicativa de 49,8%, enquanto que para o modelo do número total de mortos não chegou a ser considerada por apresentar na selecção de atributos por dimensão uma importância relativa inferior a 40%.

No modelo do número total de mortos a variável combustíveis, % das exportações de mercadorias (n-1) assume a maior capacidade explicativa, com uma importância de 100%, detendo também importâncias relativas relevantes nos modelos explicativos do número total de incidentes terroristas, com 73%, e do número total de feridos, com 72,6%, apesar de no segundo estar na terceira posição do *ranking* e no terceiro na sexta posição.

A taxa de participação laboral, mulheres (% população do sexo feminino 15 +) (n-1) apresenta em todos os modelos uma capacidade explicativa de valor elevado, aparecendo transversalmente como uma variável de relevante significado. No modelo do número total de incidentes terroristas ocupa a segunda posição no *ranking* com uma importância de 75,8%, da mesma forma que no modelo do número total de feridos, com uma capacidade explicativa de 85,2%. Para o modelo do número total de mortos tem uma importância de 67,9%, ficando na quinta posição.

Ainda dentro das primeiras três posições surgem no modelo do número total de mortos a variável população urbana (% do total) (n-1) com uma importância de 87,7% e no modelo do número total de feridos a região, com uma importância de 76,7%. A população urbana (% do total) (n-1), que mede o nível de urbanização dos países, apresenta no modelo explicativo do número total de incidentes uma importância de apenas 39,3% e no modelo do número total de feridos uma importância de 43,8% estando neste último sete posições acima no *ranking*. A região no número total de incidentes terroristas tem uma importância explicativa relativa de 37,2% e no número total de mortos uma importância de 51,8%, ocupando em ambas posições abaixo de meio na tabela de *ranking*.

Dentro do nível de capacidade explicativa até 50%, para além nas três posições cimeiras ainda se podem encontrar seis variáveis no modelo do número total de incidentes terroristas, onze variáveis no modelo do número total de mortos e sete no modelo do número total de feridos. Para o modelo do número total de incidentes terroristas a variável população, total (n-1) assume uma importância de 72,6%, surgindo para o modelo do número total de mortos com uma importância de apenas 12,2% e não tendo sequer sido incluída no modelo do número total de feridos por não apresentar importância relevante aquando da geração das árvores de regressão para a selecção de atributos por dimensão de caracterização.

As instituições em regime (n-1) na explicação do número total de incidentes terroristas apresenta uma importância de 65,7%, enquanto que para os restantes modelos não demonstrou ter importância explicativa relevante. A durabilidade de regime (n-1) apenas demonstrou capacidade explicativa com significância para o modelo do número total de incidentes terroristas com 58,5% e para o modelo do

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

número total de mortos com 59,5%. O fraccionamento político total (n-1) não surge com significado no modelo do número total de mortos, mas para o modelo do número total de incidentes detém uma capacidade explicativa de 56,5% e no modelo do número total de feridos 50,8%.

O KOF - Índice de Globalização (n-1) foi abordado nos três modelos, tendo demonstrado maior capacidade explicativa no modelo do número total de incidentes terroristas, com uma importância de 51,8%. No modelo explicativo do número total de mortos assume uma importância de 41,3% e no modelo do número total de feridos uma capacidade explicativa de 39,8%.

A área de superfície (km<sup>2</sup>) - área total geográfica do país (n-1) foi igualmente tida em linha de conta na construção dos três modelos, possuindo maior nível de capacidade explicativa no modelo do número total de mortos, com uma importância de 64,8%, seguido do modelo relativo ao número total de incidentes terroristas com uma importância de 51,7% e do modelo explicativo do número total de feridos em que apenas deteve uma importância relativa de 12,7%.

A religião católica (percentagem da população total) (n-1) apresenta maior capacidade explicativa ao nível dos impactos, nos modelos relativos ao número total de mortos e ao número total de feridos, com importâncias de 62,6% e 73,5%, do que no modelo do número total de incidentes terroristas, em que detém uma importância relativa de 41,5%. O mesmo se retrata na religião muçulmana (percentagem da população total) (n-1) que apresenta uma importância de 61,6% e 68,6% para os modelos explicativos do número total de mortos e número total de feridos, e para o modelo do número total de incidentes uma capacidade explicativa de 45,9%.

A variável outras religiões (percentagem da população total) (n-1) não demonstrou importância relativa relevante na explicação do número total de feridos dentro da sua dimensão de caracterização, pelo que não foi incluída no modelo final. Nos modelos relativos ao número total de incidentes terroristas e número total de mortos apresentou importâncias relativas acima dos 40%, e dentro dos níveis das outras variáveis de religião anteriormente referidas, ou seja, no modelo relativo ao número total de incidentes terroristas deteve uma importância de 45,8% e no modelo relativo ao número total de mortos um valor superior de 62,8%. A fragmentação religiosa (n-1) aparece como uma variável de capacidade explicativa bastante elevada no modelo do número total de mortos, com uma importância 81,3%, enquanto que no modelo do número total de incidentes terroristas já detém uma importância de 41,8% e para o modelo do número total de feridos um valor bastante inferior de 5,1%.

A fragmentação étnica (n-1) apresenta uma capacidade explicativa para o modelo do número total de incidentes de 30,8%, para o modelo do número total de mortos um valor um pouco maior, com uma importância de 53,7%, e para o modelo do número total de feridos um valor bastante superior, de 75,2%. Apesar de ter uma importância relativa significativa em todos os modelos, explica sem dúvida melhor o modelo do número total de feridos.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

A inflação, preços de consumo (% anual) (n-1) também demonstrou uma capacidade explicativa em todos os modelos de algum relevo, para o número total de incidentes terroristas com uma importância de 45,9%, e para o número total de mortos e número total de feridos com valores na ordem dos 60%, respectivamente com importâncias de 61,6% e 64,4%.

A variável crescimento populacional (% anual) (n-1) apenas detém capacidade explicativa acima de 40% no modelo do número total de mortos, com uma importância de 46,2%, tendo no modelo explicativo do número total de incidentes uma importância de 28% e no modelo relativo ao número total de feridos uma importância explicativa de 36,7%. A despesa total em saúde (% do PIB) (n-1), que na selecção de variáveis inicial não demonstrou relevância explicativa para o número total de feridos e por isso não incluída no modelo final, apresenta uma importância relativa de 49,3% no modelo do número total de mortos e um valor um pouco inferior, de 35,3%, na explicação do modelo relativo ao número total de incidentes terroristas.

O índice de direitos de integridade física (n-1) apresenta uma importância relativa de 46,9% na explicação do modelo do número total de incidentes terroristas e uma importância de 27,2% para o modelo do número total de mortos, sendo que para o modelo do número total de feridos não chegou a reflectir capacidade explicativa relevante no sentido da sua inclusão no modelo final.

O crescimento do PIB (% anual) (n-1) apenas demonstrou capacidade explicativa relevante para o modelo do número total de incidentes terroristas, com uma importância relativa de 41,8%. Pelo contrário, o investimento estrangeiro directo, entradas líquidas (% do PIB) (n-1) apenas surge nos modelos relativos ao número total de mortos e número total de feridos, com importância relativa de 48,3% e 29,4%, respectivamente.

As dimensões que explicam melhor o modelo do número total de incidentes terroristas são a sociodemográfica, económica e política, com maior relevância na capacidade explicativa para as variáveis de caracterização sociodemográfica, seguidas das políticas. A dimensão económica é representada pelas variáveis combustíveis, % das exportações de mercadorias (n-1), quota de consumo do governo sobre o PIB *per capita* (n-1), inflação, preços de consumo (% anual) (n-1) e crescimento do PIB (% anual) (n-1), sendo que apenas a primeira assume importância acima de 50%.

A dimensão sociodemográfica está representada pelas variáveis densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1), taxa de participação laboral, mulheres (% população do sexo feminino 15+) (n-1) e população, total (n-1), verificando-se em todas importâncias explicativas acima dos 70%. A dimensão política detém as variáveis instituições em regime (n-1), com uma importância de 65,7%, durabilidade de regime (n-1), com uma importância de 58,5%, o fraccionamento político total (n-1), com uma importância de 56,5% e o índice de direitos de integridade física (n-1), com uma importância de 46,9%.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

As quatro variáveis integrantes da dimensão religiosa constam todas no modelo com importâncias relativas na ordem dos 40%. Desta forma, pode-se concluir que as dimensões sociodemográfica e política são as que melhor explicam o modelo, com destaque para a dimensão sociodemográfica que detém importâncias explicativas mais elevadas. A dimensão económica apenas sobressai relativamente à dimensão religiosa e ao KOF - Índice de Globalização (0 a 100) (n-1) e à área de superfície (km<sup>2</sup>) - área total geográfica do país (n-1) devido ao peso da variável correspondente à percentagem de exportação de combustíveis, que demonstra uma forte relação na explicação do número total de incidentes terroristas.

Quanto ao modelo relativo ao número total de mortos as dimensões que melhor explicam o modelo são a sociodemográfica, religiosa e económica, com maior relevância para as duas primeiras. As variáveis sociodemográficas de maior relevo são a densidade populacional (pessoas por km<sup>2</sup> de área de terra) (n-1), a população urbana (% do total) (n-1), a taxa de participação laboral, mulheres (% da população do sexo feminino 15+) (n-1), a população de 0-14 (% do total) (n-1), a fragmentação étnica (n-1), a despesa total em saúde (% do PIB) (n-1) e o Crescimento populacional (% anual) (n-1), salientando-se as três primeiras com valores de importância relativa mais altos entre elas, respectivamente de 95,8%, 87,7% e 67,9%.

As variáveis religiosas estão todas abrangidas, com uma importância relativa acima de 60%. A dimensão económica está representada pelas variáveis combustíveis, % das exportações de mercadorias (n-1), inflação, preços de consumo (% anual) (n-1) e investimento estrangeiro directo, entradas líquidas (% do PIB) (n-1). Apesar da variável explicativa de maior relevo, com importância de 100%, pertencer à dimensão económica, ela é apenas um indicador económico parcial que não veicula uma caracterização da economia do país como um todo, referindo-se exclusivamente à percentagem de exportação de combustíveis no âmbito da percentagem total de exportação de mercadorias. Esta variável de facto diz-nos que a economia de um país é rica, mas é um indicador que poderá estar relacionado com muitos outros factores ligados a contextos geopolíticos que poderão influenciar a sua relação com o terrorismo. A inflação, preços de consumo (% anual) (n-1) tem uma importância de 61,6%, inferior a todas as variáveis de caracterização religiosa, e o investimento estrangeiro directo, entradas líquidas (% do PIB) (n-1) uma importância relativa de 48,3%.

A dimensão política está apenas representada pela variável durabilidade de regime (n-1), com uma importância de 59,5%, indicador que poderá relacionar-se com contextos de instabilidade política. Neste modelo as restantes variáveis de caracterização política não demonstraram boa capacidade explicativa do modelo.

No modelo relativo ao número total de feridos as variáveis de maior poder explicativo pertencem às várias dimensões. À dimensão política pertence apenas a variável fraccionamento político total (n-1),

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

com uma importância relativa de 50,8%. A região surge com uma capacidade explicativa muito alta neste modelo, de 76,7%, e a mais elevada de entre os três modelos. As variáveis económicas mais relevantes neste modelo são a quota de consumo do governo sobre o PIB *per capita* (n-1), com 100% de importância explicativa, a combustíveis, % das exportações de mercadorias (n-1), com uma importância de 72,6% e a inflação, preços de consumo (% anual) (n-1), com uma importância de 64,4%.

Pertencentes à dimensão sociodemográfica encontramos a taxa de participação laboral, mulheres (% da população do sexo feminino 15 +) (n-1), com uma importância explicativa de 85,2%, a fragmentação étnica (n-1), com uma importância de 75,2%, a população com 65 anos e acima (% do total) (n-1), com uma importância de 68,6% e a população urbana (% do total) (n-1), com uma importância de 43,8%. Da dimensão religiosa as variáveis mais significativas são a religião católica (percentagem da população total) (n-1), com uma importância de 73,5% e a religião muçulmana (percentagem da população total) (n-1), com uma importância de 68,6%. Pode-se concluir que neste modelo as dimensões de maior poder explicativo são a económica e a sociodemográfica.

De uma forma geral na avaliação dos três modelos destaca-se a dimensão sociodemográfica como a que melhor explica as variáveis de terrorismo, apresentando além de importâncias explicativas mais altas, um número maior de variáveis com capacidade explicativa acima de 40%, perfazendo um total de quinze. Em segundo vem a dimensão económica com um conjunto de dez variáveis com capacidade explicativa superior a 40% e também com importâncias relativas na explicação dos modelos mais altas, muito embora o maior peso se tenha verificado na variável combustíveis, % das exportações de mercadorias (n-1), indicador conotado com uma certa subjectividade no contexto geopolítico. A dimensão religiosa demonstrou ter uma capacidade explicativa mais relevante nos modelos do número total de mortos e número total de feridos do que no modelo do número de incidentes terroristas, indiciando ter maior relação com os impactos dos incidentes do que com a sua frequência.

## **6.4 Discussão de Resultados**

A revisão da literatura sobre o tema abordado na presente investigação permitiu constatar a não existência de estudos anteriores em que tivesse sido utilizada a metodologia CART, o que invalida a comparação de resultados com base nas métricas usadas. A generalização dos factores determinantes do terrorismo é uma tarefa de elevada complexidade tendo em conta que as diversas regiões e países têm contextos geopolíticos diferentes e que as motivações e ideologias dos vários grupos terroristas por todo o mundo são também bastante desiguais. Na avaliação sobre a proporção de variância explicada pelos modelos para toda a amostra, mas principalmente quando medida através da validação cruzada, ou seja, quanto à sua capacidade preditiva na aplicação a novos dados, não se expectavam valores altos, sendo considerados satisfatórios na previsão valores ao nível dos 50%.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Na explicação das características do terrorismo em análise pode-se concluir que factores económicos têm um papel relevante. Um baixo consumo do governo sobre o PIB *per capita* e baixos níveis de inflação revelam menos terrorismo. Pelo contrário, menores valores de PIB *per capita* e quota de investimento do PIB *per capita* reduzida resultam em mais terrorismo. Economias de menor estabilidade, de fraco crescimento e conducentes a fracas condições económicas demonstram assim ter impactos no aumento do terrorismo, o que vai ao encontro das conclusões de Azam e Delacroix (2006), Azam e Thelen (2008), Blomberg *et al.* (2004a, 2004b), Blomberg e Hess (2008a, 2008b), Feldmann e Perälä (2004), Frey e Luechinger (2003), Freytag *et al.* (2011), Goldstein (2005), Krieger e Meierrieks (2011), Krueger e Laitin (2008), Lai (2007) e Piazza (2011), contrariando por outro lado os resultados de Basuchoudhary e Shughart (2010), Krueger e Maleckova (2003), Kurrild-Klitgaard *et al.* (2006), Meierrieks e Gries (2012) e Plümper e Neumayer (2010), que não encontraram indícios de que o crescimento económico possa influir de forma relevante nos padrões de terrorismo.

No que se refere à dimensão política foi possível concluir que a instabilidade política cria condições favoráveis ao crescimento do terrorismo, prevendo-se para regimes com duração inferior a meio ano números de terrorismo mais elevados, o que reitera os resultados de Gassebner *et al.* (2008), Gassebner e Luechinger (2011), Krueger e Maleckova (2003), Li (2005), Meierrieks e Gries (2012) e Piazza (2008b, 2009). A valores mais altos de despesas militares corresponde mais terrorismo, sendo que, tal como menciona Campos e Gassebner (2009), clima de guerra civil constitui igualmente ambiente ideal à implementação de campos de treino para potenciais terroristas.

Os resultados da presente investigação contrariam as conclusões de Krueger (2007) e Pape (2006), que afirmam que o terrorismo tem origem em sociedades que estão sob regimes repressores, enfatizando as causas políticas ao invés das económicas ou culturais. Pelo contrário, no presente estudo as dimensões sociodemográfica e económica detêm bastante mais relevância na explicação do terrorismo do que a dimensão política. Relativamente ao tipo de regime político, pôde-se apurar que a níveis mais altos de democracia corresponde mais terrorismo. Resultado idêntico foi obtido por Frey e Luechinger (2003) e Li (2005), o que se poderá justificar com o facto de que a repressão política e acções militares contra grupos terroristas são menos prováveis quando governos têm que respeitar as liberdades civis e ter em conta uma ampla gama de interesses políticos. Contudo, Gassebner e Luechinger (2011) no seu estudo empírico não encontraram qualquer relação entre a democracia e o terrorismo. Concluíram que a ausência de abusos no que se refere ao respeito pelos direitos humanos e direitos de integridade física está associado a menos terrorismo, o que vai ao encontro do estudo ora efectuado que revela que a índices de direitos de integridade física mais baixos correspondem valores de terrorismo previstos mais altos. Também se tem que ter em consideração, conforme referem Li (2005) e Drakos e Gofas (2006), que em países de baixo nível de desenvolvimento político, com regimes repressivos, existe maior susceptibilidade de controlo das notícias veiculadas na imprensa. Nesse sentido, a cobertura da informação sobre o terrorismo tende

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

a ser restrita, e conseqüentemente diferente do número de incidentes relatados e posteriormente registados nas bases de dados de terrorismo, e daí parecer haver menos terrorismo nesses casos.

Relativamente à existência de mais terrorismo nos países com valores muito elevados de exportação de combustíveis, este resultado acompanha as conclusões de Le Billon (2001), Tavares (2004), Bravo e Dias (2005) e Sambanis (2008), que encontraram uma relação positiva no que se refere à importância da existência de recursos naturais (minerais) estratégicos no surgimento de conflitos armados, onde se inclui o terrorismo.

No modelo explicativo do número total de incidentes terroristas a globalização, medida através das variáveis KOF - Índice de Globalização e utilizadores de internet, demonstrou revelar menos terrorismo em níveis mais baixos de globalização. A globalização é um factor interdisciplinar, sendo o KOF - Índice de Globalização um indicador composto que detém itens de diversas áreas, incluindo a económica, espelhando também o nível de integração económica global de um país. A globalização e modernização acarretam mudanças económicas, novas formas de comunicação e estilos de vida pelo avanço das novas tecnologias, urbanização, e novas ideias ou ideologias. Estes factores podem criar sentimentos de revolta e injustiça quando associados a tensões sociodemográficas e económicas. O uso das novas tecnologias pode ser um obstáculo para muitos, e para as organizações terroristas um meio favorecedor e disseminador de ideologias e incitamento ao recrutamento, como concluiu Robison *et al.* (2006). Blomberg e Hess (2008a, 2008b), Drakos e Gofas (2006), Koch e Cranmer (2007) e Kurtil-Klitgaard *et al.* (2006) também encontraram fortes evidências da influência da globalização no aumento do número de terrorismo, ao contrário de Braithwaite e Li (2007), Bravo e Dias (2005), Burgoon (2006), Campos e Gassebner (2009) e Li e Schaub (2004) que nos seus estudos não apuraram significâncias relevantes no que se refere à globalização como factor explicativo do terrorismo.

A dimensão populacional mostrou ser um relevante factor explicativo do terrorismo, com maiores números de terrorismo a corresponder a maiores níveis populacionais, o que confirma resultados de estudos anteriores que demonstraram uma correlação positiva com o terrorismo (Burgoon, 2006; Freytag *et al.*, 2008; Gassebner e Luechinger, 2011; Krieger e Meierrieks, 2011; Krueger e Laitin, 2008; Krueger e Maleckova, 2003; Lai, 2007; Meierrieks e Gries, 2012; Piazza, 2008b e Plümper e Neumayer, 2010). Tal relação pode justificar-se pelo facto de em alguns casos serem países maiores e haver maior probabilidade de ataques por disporem de mais potenciais alvos, vítimas, ou perpetradores, ou até por uma grande população poder reflectir maior captação de recursos para o policiamento em detrimento de maior investimento em contra-terrorismo, ou poder ser alvo de maiores níveis de tensão demográfica. Apenas nos estudos de Azam e Delacroix (2006) e Azam e Thelen (2008) a dimensão populacional não demonstrou ter relação com o terrorismo.

Tal como em Dreher e Fischer (2010) a presente análise demonstrou que o crescimento populacional tem significativa capacidade explicativa do terrorismo, concluindo que países com população em

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

crescimento são alvo de menos terrorismo e de menor intensidade. Nestes casos, o crescimento populacional pode estar relacionado com equilíbrio e bem-estar socioeconómico. Variáveis ligadas à dinâmica populacional, como o crescimento populacional, a dimensão e densidade populacional, a estrutura etária, ou a urbanização (população urbana), que apresentaram capacidades explicativas significativas nos modelos de terrorismo em estudo, também revelaram fortes evidências como determinantes do terrorismo nos estudos de Braithwaite e Li (2007), Burgoon (2006), Campos e Gassebner (2009), Drakos e Gofas (2006), Dreher e Gassebner (2008), Koch e Cranmer (2007), Krueger e Laitin (2008), Krueger e Maleckova (2003), Lai (2007), Li (2005), Li e Schaub (2004), Piazza (2006, 2008a, 2008b), Plümper e Neumayer (2010), Sambanis (2008) e Tavares (2004).

A educação é um factor que na literatura demonstra diversas posições antagónicas quanto à sua relação com o terrorismo. Berrebi (2007), Krueger e Maleckova (2003) e Krueger (2008) encontram uma relação positiva entre terrorismo e educação, apesar da significância ser baixa os autores concluem que o capital humano é um preditor positivo da participação no terrorismo. Freytag *et al.* (2011) detectaram um efeito significativo entre o número de anos de estudo e o terrorismo. Alguns outros autores ainda encontraram na educação um factor determinante para o terrorismo como Azam e Thelen (2008), Bravo e Dias (2005) e Tavares (2004), contudo uma grande parte das análises empíricas teve neste campo resultados idênticos à investigação em análise, ou seja, a educação, analisada através das variáveis anos de educação e frequência escolar, não demonstrou ter significativa capacidade explicativa do terrorismo, não constituindo desta forma factor preditivo do desenvolvimento do terrorismo.

Na maioria dos estudos encontrados na revisão da literatura foi possível constatar fracas evidências de correlações significativas entre religião e terrorismo. Piazza (2006) concluiu que existe significativa relação entre factores ligados à religião e os níveis de terrorismo, principalmente do que se refere aos impactos do terrorismo (número de mortos e feridos) do que na sua frequência, o que vai ao encontro dos resultados obtidos neste estudo. Também Gassebner e Luechinger (2011) encontraram uma relação positiva entre as tensões religiosas e o terrorismo. Contudo, diversos foram os autores que não detectaram relação significativa entre a religião e o terrorismo, como Pape (2006), ou relação fraca como Abadie (2006), Blomberg e Hess (2008a, 2008b), Bravo e Dias (2005), Krueger e Laitin (2008), Krueger e Maleckova (2003), Kurriil-Klitgaard *et al.* (2006) e Tavares (2004).

Nos modelos apurados no presente estudo a fragmentação étnica revelou-se uma variável com alguma relevância na explicação do número total de incidentes terroristas, mas principalmente na explicação do número total de mortos e feridos, nos quais obteve valores de importância mais elevados. Estes resultados dão robustez às análises de Abadie (2006), Basuchoudhary e Shughart (2010), Bravo e Dias (2005), Gassebner e Luechinger (2011), Kurriil-Klitgaard *et al.* (2006), Piazza (2006, 2008b), Sambanis (2008) e Tavares (2004), que encontraram fortes evidências da relação positiva entre a fragmentação étnica e o terrorismo.



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

A variável expectativa de vida ao nascer, em sintonia com os estudos de Drakos e Gofas (2006) e Gassebner e Luechinger (2011), não demonstrou importância explicativa em qualquer um dos modelos propostos. A taxa de participação laboral das mulheres, factor anteriormente abordado por Gassebner e Luechinger (2011) e Robison *et al.* (2006), deteve neste estudo importâncias explicativas bastante elevadas para os três indicadores do terrorismo. Nos estudos anteriores revelou da mesma forma uma relação negativa com o terrorismo, mas com valores de significância bastante menos relevantes.

Uma visão global sobre a totalidade dos estudos revistos permitiu apurar que os factores sóciodemográficos foram os que obtiveram maiores evidências como potenciais determinantes do terrorismo. Em segundo lugar destacam-se os factores económicos, que surgem intrinsecamente ligados à dimensão social, pela influência que a melhoria ou decréscimo das condições económicas tem na vivência da população. Seguidamente, mas também com relevância, figuram os factores de cariz político. Os factores religiosos demonstraram transversalmente uma fraca significância como determinantes do terrorismo. A dimensão que demonstrou mais controvérsia nos resultados foi a económica, destacando-se contudo a vertente que destaca a sua importância. Os factores sociodemográficos foram distintamente os que reuniram mais consenso no âmbito de todos os estudos analisados.

O presente estudo acompanha as grandes linhas apresentadas. Contudo, na explicação dos impactos do terrorismo, revelados pelo número de mortos e feridos, a dimensão política não demonstrou tão significativa capacidade explicativa, à qual se sobrepôs a dimensão religiosa, dimensão que revelou nesta análise maior peso como factor preditor do terrorismo do que na generalidade dos estudos anteriores.

### **6.5 Sumário**

Caracterizou-se o terrorismo em três indicadores: o número total de incidentes terroristas, o número total de mortos e o número total de feridos. Apresentou-se uma breve caracterização das dimensões económica, sociodemográfica, política, religiosa e outras variáveis dos países/anos que constituem a amostra.

Construíram-se árvores de regressão com o algoritmo CART para explicar os três indicadores do terrorismo em função dos atributos de cada uma das dimensões com o objectivo de identificar, em cada dimensão e em cada característica do terrorismo, os atributos mais importantes (com importância relativa acima de 40% e, em alguns casos, acima de 30%).

No final foram construídos três modelos de árvores de regressão (CART), um para cada característica do terrorismo e tendo como atributos explicativos o conjunto identificado nos modelos parciais. Os modelos obtidos apresentam capacidade explicativa entre 59,7% e 76,4% e capacidade

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

preditiva entre 36,1% e 55,1%, mas como apontam estudos anteriores, conseguir melhor é difícil dado que para o terrorismo a generalização torna-se uma tarefa complexa tendo em conta que as diversas regiões e países têm contextos geopolíticos diferentes e que as motivações e ideologias dos vários grupos terroristas por todo o mundo são também bastante desiguais. Em todos os modelos surgem atributos das diferentes dimensões, mas destacam-se as sociodemográficas e as económicas e com menos influência as políticas e religiosas.

Os resultados, embora obtidos recorrendo a uma técnica ainda não utilizada nos estudos analisados na revisão da literatura, vêm confirmar, por vezes, resultados de outros estudos. Noutros, vêm contrariar, o que evidencia que mais investigação deve ser desenvolvida e que as amostras (países e anos) utilizadas em cada estudo podem conduzir a resultados diferentes.

## **7. Conclusões**

---

A investigação desenvolvida procurou constituir uma abordagem ao estudo do terrorismo no sentido de estabelecer mais um passo para a compreensão dos processos ou indicadores subjacentes à radicalização nas suas formas mais extremas. Tomaram-se como objectivos a identificação dos factores determinantes do terrorismo e a concepção de um modelo explicativo e preditivo para a ocorrência e impacto de actos terroristas.

O processo de recolha de dados exigiu uma pesquisa aturada por diversas fontes abertas devido à grande diversidade de variáveis necessárias para a análise proposta. A possibilidade de junção de uma vasta série de variáveis revelou-se uma mais-valia, obrigando contudo a um árduo trabalho de uniformização e validação dos dados, que tendo origens diferentes traziam também formatação e organização heterogéneas. Com base nestas fontes conseguiu-se uma grande amostra, com 3460 observações, correspondentes a 173 países, ao longo de 20 anos decorridos entre 1990 e 2010.

Perante uma grande amostra e uma elevada quantidade de atributos potencialmente explicativos do terrorismo (85), agrupados em cinco dimensões, económica, sociodemográfica, política, religiosa e de carácter multidimensional, as técnicas de *data mining*, nomeadamente as árvores de regressão com o algoritmo CART, foram consideradas as adequadas neste estudo.

Os objectivos foram plenamente atingidos e a técnica de análise utilizada demonstrou ser eficaz para a concretização deste e de futuros estudos na área do terrorismo. Foram identificados diversos factores considerados como determinantes do terrorismo, pertencentes às cinco dimensões de variáveis, tendo-se confirmado a hipótese 2, já que os factores sociodemográficos são os que melhor explicam a actividade terrorista. Os modelos explicativos para os três indicadores de terrorismo não traduziram capacidades explicativas e preditivas muito elevadas, mas consideradas importantes dada a complexidade e subjectividade que envolve a área do terrorismo.

### **7.1 Contributos**

Diversas organizações de variadas áreas publicam periodicamente bases de dados no sentido de estimular o desenvolvimento de estudos científicos de cariz multidisciplinar. O estudo desenvolvido demonstrou que a utilização de dados de fontes abertas permite obter resultados válidos e interessantes para uma realidade muito complexa. Perante o incentivo à utilização destas fontes em investigação científica por parte de organismos internacionais de reconhecido prestígio, este trabalho, que se pretende divulgar em conferências científicas ou não, de âmbito mais geral ou mais relacionadas com o tema do terrorismo e contra-terrorismo, é mais um contributo válido para o conhecimento sobre o tema.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

O terrorismo é sempre um tema actual e em permanente evolução, pelo que a inclusão neste estudo dos dados mais recentes (até 2010) disponíveis é mais um factor importante para o conhecimento.

A investigação desenvolvida aplicou uma nova técnica no estudo deste tema, especialmente adaptada à análise de um grande volume de dados, tendo mostrado ser eficaz. O estudo focou e pretendeu explicar para além do número de incidentes terroristas, o número de mortos e feridos, variáveis de terrorismo menos abordadas nos estudos anteriores. Contribuiu igualmente para a identificação dos atributos económicos, sociodemográficos, políticos, religiosos, e outros, mais importantes, isto é, mais relacionados com a incidência do terrorismo. Tendo em vista um estudo holístico sobre o terrorismo, procurou também reunir num estudo global as variáveis utilizadas nos diversos estudos anteriores, algumas delas introduzidas em tempo muito recente e ainda parcamente analisadas, aplicadas a uma ampla amostra de países de todas as regiões do mundo.

O maior conhecimento sobre os factores determinantes do terrorismo fornece argumentação teórica e evidência empírica que proporcionam bases conducentes à elaboração de directrizes para projectar e melhorar as intervenções para a prevenção, detenção e reversão dos processos de radicalização onde ela possa vir a gerar actos de violência.

### **7.2 Limitações**

A complexidade e interdisciplinaridade que envolve o tema do terrorismo exigem o dispêndio e investimento de mais meios e tempo no sentido da concretização de estudos de nível mais aprofundado.

Algumas variáveis de alguma importância para a caracterização dos países tiveram que ser excluídas por terem demasiados valores omissos, principalmente em certos países, pelo que teria sido importante que todas as variáveis estivessem actualizadas para um maior número de países.

Ainda relativamente aos dados importa referir que apesar do esforço e diligência na escolha das várias fontes, que se pretenderam ser o mais fiáveis e íntegras possível, existe sempre uma certa subjectividade a eles associada sobre a qual não se pôde ter controlo.

### **7.3 Pistas Futuras**

Em futuras investigações seria interessante a utilização de outras técnicas de *data mining* e testar a sua eficácia na área do estudo do terrorismo. Na continuidade do presente trabalho poderiam ser comparados os resultados obtidos através das diferentes técnicas de *data mining*.

Sugere-se também a implementação dos modelos gerados e posterior validação com os dados reais dos anos seguintes.

## Referências Bibliográficas

---

- Abadie, A. (2006). Poverty, political freedom, and the roots of terrorism. *American Economic Review*, 96(2), pp. 50–56.
- Abadie, A. e Gardeazabal, J. (2003). The economic costs of conflict: a case study of the Basque Country. *American Economic Review* 93(1), pp. 113–132.
- Abadie, A. e Gardeazabal, J. (2008). Terrorism and the world economy. *European Economic Review*, 52, pp. 1–27.
- Alesina, A.; Devleeschauwer, A.; Easterly, W.; Kurlat, S. e Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8: pp. 155-194.
- Araz-Takay, B., Arin, K. e Omay, T. (2009). The endogenous and non-linear relationship between terrorism and economic performance: Turkish evidence. *Defence and Peace Economics* 20(1), pp. 1–10.
- Azam, J., e Delacroix, A. (2006). Aid and the delegated fight against terrorism. *Review of Development Economics*, 10(2), pp.330–344.
- Azam, J. e Thelen, V. (2008). The roles of foreign aid and education in the war on terror. *Public Choice*, 135(3), pp. 375–397.
- Azam, J. e Thelen, V. (2010). Foreign aid versus military intervention in the war on terror. *Journal of Conflict Resolution* 54(2), pp. 237–261.
- Basuchoudhary, A. e Shughart, W. (2010). On ethnic conflict and the origins of transnational terrorism. *Defence and Peace Economics* 21(1), pp. 65–87.
- Beck, T.; Clarke, G.; Groff, A; Keefer, P. e Walsh, P. (2001). New tools in comparative political economy: The Database of Political Institutions. 15:1, pp. 165-176 (September), *World Bank Economic Review*.
- Berman, E., e Laitin, D. (2008). Religion, terrorism and public goods: testing the club model. *Journal of Public Economics*, 92(10–11), pp. 1942–1967.
- Berrebi, C. (2007). Evidence about the link between education, poverty and terrorism among Palestinians. *Peace Economics, Peace Science and Public Policy*, 13(1), Article 2.
- Blomberg, S., Hess, G. e Weerapana, A. (2004a). An economic model of terrorism. *Conflict Management and Peace Science* 21(1), pp. 17–28.
- Blomberg, S., Hess, G. e Weerapana, A. (2004b). Economic conditions and terrorism. *European Journal of Political Economy* 20, pp. 463–478.
- Blomberg, S. e Hess, G. (2008a). The Lexus and the olive branch: globalization, democratization and terrorism. In P. Keefer e N. Loayza (Eds.), *Terrorism, economic development, and political openness*. Cambridge: Cambridge University Press, pp. 116–147.
- Blomberg, S. e Hess, G. (2008b). From (no) butter to guns? Understanding the economic role in transnational terrorism. In *Economic Development, and Political Openness*, edited by P. Keefer and N. Loayza. Cambridge: Cambridge University Press, pp. 83–115.
- Blomberg, S. e Rosendorff, B. (2009). A gravity model of globalization, democracy and transnational terrorism. In G. D. Hess (Ed.), *Guns and butter*. Cambridge: MIT Press, pp. 25–156.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

- Braithwaite, A., e Li, Q. (2007). Transnational terrorism hot spots: identification and impact evaluation. *Conflict Management and Peace Science*, 24(4), pp. 281–296.
- Bravo, A. e Dias, C. (2005) Os factores determinantes do terrorismo num contexto geopolítico. PROELIUM, *Revista da Academia Militar* (ISSN 1645/ 8826), 6ª Série, nº 2, Janeiro.
- Braz, J. (2009). *Investigação Criminal, a organização, o método e a prova, os desafios da nova criminalidade*. Edições Almedina S.A..
- Breiman, L., Friedman, J., Olshen, R. e Stone, C. (1984). *Classification and Regression Trees*. Chapman e Hall/CRC Press.
- Brown, M. e Kros, J. (2003). Data mining and the impact of missing data. *Industrial Management e Data Systems*, vol. 103 (8), pp.611 - 621.
- Bueno de Mesquita, E. (2005). The quality of terror. *American Journal of Political Science*, 49(3), pp. 515–530.
- Burgoon, B. (2006). On welfare and terror: social welfare policies and political-economic roots of terrorism. *Journal of Conflict Resolution*, 50(2), pp. 176–203.
- Cabete, N.; Cardoso, M. (2006). Algoritmo CART: Previsão do Desempenho na Matemática do Secundário. *Revista de Ciências da Computação (Universidade Aberta)*, Vol. 1(1), pp. 11-23.
- Campos, N. e Gassebner, M. (2009). International terrorism, political instability and the escalation effect. *CEPR Discussion Paper 7226*.
- Capul, J. e Garnier, O. (1998). *Dicionário de Economia e Ciências Sociais*. Tradução: Germano Rio Tinto, 1ª Edição, Lisboa, Plátano Edições Técnicas.
- Cardoso, T. (2011). Os impactos da globalização no terrorismo. *CEDIN - revista eletrónica de direito internacional*, volume 8, 1º semestre, pp. 289 - 329.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. e Wirth, R. (2000). *CRISP-DM 1.0 - Step-by-step data mining guide*, CRISP-DM Consortium. URL: <http://www.crisp-dm.org>, acedido em Janeiro de 2012.
- Cheibub, J.; Gandhi, J. e Vreeland, J. (2009). Democracy and dictatorship revisited. *Public Choice*, 143(1-2), pp. 67-101.
- Cingranelli, D. e Richards, D. (2010). *The Cingranelli-Richards (CIRI) Human Rights Dataset*. Version 2010.05.17, <http://www.humanrightsdata.org> (Fevereiro de 2012).
- Crain, N. e Crain, W. (2006). Terrorized economies. *Public Choice* 128, pp. 317–349.
- Crenshaw, E., Robison, K., e Jenkins, J. (2007). The “roots” of transnational terrorism: a replication and extension of Burgoon. *Paper presented to the annual meetings of the American Sociological Association*, New York City (August 2007).
- Dantas, G. e Júnior, C. (2009). A descoberta e a análise de vínculos na complexidade da investigação criminal moderna. *White Paper*, Tempo Real Tecnologias de Informação.
- DeRosa, M. (2004). Data Mining and Data Analysis for Counterterrorism, *CSIS Report (Center for Strategic & International Studies)*.
- Dias, M. (2002). Parâmetros na escolha de técnicas e ferramentas de mineração de dados. *Acta Scientiarum*. Maringá, v.24, n.º 6, p. 1715-1725.
- Drakos, K. e Gofas, A. (2006). In search of the average transnational terrorist attack venue. *Defence and Peace Economics* 17(2), pp. 73–93.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Dreher, Axel (2006). Does Globalization Affect Growth? Evidence from a new Index of Globalization. *Applied Economics* 38(10), pp. 1091-1110.

Dreher, A. e Fischer, J. (2010). Government decentralization as a disincentive for transnational terror? An empirical analysis. *International Economic Review*, 51(4), pp. 981–1002.

Dreher, A. e Fischer, J. (2011). Does government decentralization reduce domestic terror? An empirical test. *Economics Letters*, 113(3), pp. 223–225.

Dreher, A. e Gassebner, M. (2008). Does political proximity to the US cause terror? *Economics Letters*, 99(1), pp. 27–29.

Dreher, A.; Gaston, N. e Martens, P. (2008). *Measuring Globalisation - Gauging its Consequences* (New York: Springer).

Dreher, A., Krieger, T. e Meierrieks, D. (2011). Hit and (they will) run: the impact of terrorism on migration. *Economics Letters* 113, pp. 42–46.

Eckstein, Z. e Tsiddon, D. (2004). Macroeconomic consequences of terror: theory and the case of Israel. *Journal of Monetary Economics* 51, pp. 971–1002.

Ehrlich, P.; Liu, J. (2002). Some roots of terrorism. *Population and Environment: A Journal of Interdisciplinary Studies*, vol 24(2), pp. 183-192.

Enders, W. e Sandler, T. (1996). Terrorism and foreign direct investment in Spain and Greece. *Kyklos* 49(3), pp. 331–352.

Enders, W., Sachida, A. e Sandler, T. (2006). The impact of transnational terrorism on US foreign direct investment. *Political Research Quarterly* 59(4), pp. 517–531.

Evans, A. (2006). Understanding Madrasahs. Foreign Affairs (site)  
<http://www.foreignaffairs.com/articles/61369/alexander-evans/understanding-madrasahs> (Acedido em 15/12/2012)

Eyerman, J. (1998). Terrorism and democratic states: soft targets or accessible systems? *International Interactions*, 24(2), 151–170.

Fayyad, U.; Piatetsky-Shapiro, G. e Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *American Association for Artificial Intelligence. AI Magazine*, pp. 37-54.

Feldmann, A. e Perälä, M. (2004). Reassessing the causes of nongovernmental terrorism in Latin America. *Latin American Politics and Society* 46(2), pp. 101–132.

Freedom House (2010) *Freedom of in the World 2010: the annual survey of political rights and civil liberties*. Washington: Freedom House. Disponível em <http://www.freedomhouse.org>. (Fevereiro de 2012).

Frey, B. (2009). How can business cope with terrorism? *Journal of Policy Modeling* 31, pp. 779–787.

Frey, B. e Luechinger, S. (2003). How to fight terrorism: alternatives to deterrence. *Defence and Peace Economics* 14(4), pp.237–249.

Frey, B., Luechinger, S. e Stutzer, A. (2009). The life satisfaction approach to valuing public goods: the case of terrorism. *Public Choice*, 138, pp. 317–345.

Freytag, A.; Krüger, J.; Meierrieks, D. & Schneider, F. (2008). The origins of terrorism: Cross-country estimates on socio-economic determinants of terrorism. *CIE Working Paper 2009-01, University of Paderborn*.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

- Freytag, A.; Kruger, J.; Meierrieks, D. e Schneider, F. (2011). The origins of terrorism: Cross-country estimates of socio-economic determinants of terrorism. *European Journal of Political Economy* 27, pp. 5-16.
- Furtado, V. (2002). *Tecnologia e gestão da informação na segurança pública*. Editora Garamond.
- Gaibulloev, K. e Sandler, T. (2009). The impact of terrorism and conflict on growth in Asia. *Economics e Politics* 21(3), pp. 359–383.
- Gaibulloev, K. e Sandler, T. (2011). The adverse effect of transnational and domestic terrorism on growth in Africa. *Journal of Peace Research* 48(3), pp. 355–371.
- Gakidou, E., Cowling, K., Lozano, R. e Murray, C. (2010). Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: a systematic analysis. *Lancet*, 376: pp. 959-974.
- Gassebner, M., Jong-A-Pin, R., e Mierau, J. (2008). Terrorism and electoral accountability: one strike, you're out! *Economic Letters*, 100, pp. 126–129.
- Gassebner, M. e Luechinger, S. (2011). Lock, stock, and barrel: a comprehensive assessment of the determinants of terror. *Public Choice* (149), pp. 235–261.
- Goldstein, K. (2005). Unemployment, inequality and terrorism: another look at the relationship between economics and terrorism. *Undergraduate Economic Review*, 1(1), Article 6.
- Gries, T., Krieger, T. e Meierrieks, D., (2011). Causal linkages between domestic terrorism and economic growth. *Defence and Peace Economics* 22(5), pp. 493–508.
- Han, J. e Kamber. M. (2001). *Data Mining, Concepts and Techniques*. Academic Press.
- Heston, A.; Summers, R. e Aten, B. (2011). *Penn World Table Version 7.0*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, June.
- Hoffman, B. (1998). *Inside Terrorism*, New York: Columbia University Press.
- Hoffman, B. (2006). *Inside Terrorism*, New York: Columbia University Press, revised and expanded edition.
- Imprensa Nacional Casa da Moeda (2003a). *Lei n.º52/2003*, publicada em 22 de Agosto.
- Imprensa Nacional Casa da Moeda (2003b). *Declaração de rectificação n.º16/2003*, publicada em 29 de Outubro.
- Imprensa Nacional Casa da Moeda (2008). *Lei n.º25/2008*, publicada em 5 de Junho.
- Imprensa Nacional Casa da Moeda (2011). *Lei n.º 17/2011*, publicada em 3 de Maio.
- Kass, G. (1980). An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics*, Vol. 29, N. 2 (1980), pp. 119–127.
- Kis-Katos, K.; Liebert, H. e Schulze, G. (2011). On the origin of domestic and international terrorism. *European Journal of Political Economy*, 27, pp. 17-36.
- Koch, M. e Cranmer, S. (2007). Testing the “Dick Cheney” hypothesis: do governments of the left attract more terrorism than governments of the right? *Conflict Management and Peace Science*, 24(3), pp. 311–326.
- Krieger, T. e Meierrieks, D. (2011). What causes terrorism? *Public Choice* 147, pp. 3–27.
- Krueger, A. (2007). *What Makes a Terrorist*. Princeton: Princeton University Press.



## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

- Krueger, A. (2008). What makes a homegrown terrorist? Human capital and participation in domestic Islamic terrorist groups in the U.S.A. *Economics Letters*, 101, pp.293–296.
- Krueger, A. e Laitin, D. (2008). Kto kogo? A cross-country study of the origins and targets of terrorism. In P. Keefer e N. Loayza (Eds.), *Terrorism, economic development, and political openness*, pp. 148-173. Cambridge: Cambridge University Press.
- Krueger, A. e Maleckova, J. (2003). Education, poverty and terrorism: is there a causal connection? *Journal of Economic Perspectives* 17(4), pp.119–144.
- Krueger, A. e Maleckova, J. (2009). Attitudes and action: public opinion and the occurrence of international terrorism. *Science*, 325(5947), pp. 1534–1536.
- Kurrild-Klitgaard, P., Justesen, M. e Klemmensen, R. (2006). The political economy of freedom, democracy and transnational terrorism. *Public Choice* 128, pp. 289–315.
- La Porta, R.; López-de-Silanes, F.; Shleifer, A. e Vishny, R. (1999). The Quality of Government. *Journal of Law, Economics and Organization*, 15(1): pp. 222-279.
- LaFree, G. (2010). The Global Terrorism Database (GTD): Accomplishments and Challenges. *Perspectives on Terrorism*, vol.4, Issue 1, March, pp.24 - 46.
- Lai, B. (2007). “Draining the swamp”: an empirical examination of the production of international terrorism 1968–1998. *Conflict Management and Peace Science*, 24(4), pp. 297–310.
- Laquer, W. (2003). *No end to war: Terrorism in the twenty-first century*. The Continuum International Publishing Group Inc..
- Laquer, W. (2009). *A history of terrorism*. The Continuum International Publishing Group Inc..
- Larocque, D., Lincourt, G. e Normandin, M. (2010). Macroeconomic effects of terrorist shocks in Israel. *Defence and Peace Economics* 21(4), pp. 317–336.
- Le Billon, P. (2001). The political ecology of war: natural resources and armed conflicts. *Political Geography* 20 (3), pp. 561-584.
- Li, Q. (2005). Does democracy promote or reduce transnational terrorist incidents? *Journal of Conflict Resolution* 49(2), pp. 278–297.
- Li, Q. e Schaub, D. (2004). Economic globalization and transnational terrorism: a pooled time-series analysis. *Journal of Conflict Resolution*, 48(2), pp. 230–258.
- Llorca-Vivero, R. (2008). Terrorism and international tourism: new evidence. *Defence and Peace Economics* 19(2), pp. 169–188.
- Magalhães, D. e Nasser, M. (2011). As relações do terrorismo moderno com a religião. *Revista Nures* n.º 17, Janeiro / Abril, Núcleo de Estudos Religião e Sociedade - Pontifícia Universidade Católica - SP ISSN 1981-156X, pp.1 - 9.
- Marshall, M. e Jagers, K. (2011). Polity IV Project: Political Regime Characteristics and Transitions, 1800-2010. *University of Maryland*, disponível em: <http://www.systemicpeace.org/polity/polity4.htm> (Fevereiro de 2012).
- Meierrieks, D. e Gries, T. (2012). Economic Performance and Terrorist Activity In Latin America. *Defence and Peace Economics*, pp. 1 - 24.
- Memon, N. (2007). *Investigative Data Mining: Mathematical Models for Analysing, Visualizing and Destabilizing Terrorist Networks*. Tese de Doutorado, Aalborg Universitet Esbjerg, Dinamarca.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

- Miranda, R. (2011). A Polícia Judiciária na prevenção, investigação e gestão de informações sobre os crimes de organizações terroristas e terrorismo: proposta de um (novo) modelo. *Revista de Investigação Criminal, Ciências Criminais e Forenses*, 2ª ed., 2011.
- Mirza, D. e Verdier, T. (2008). International trade, security and transnational terrorism: theory and a survey of empirics. *Journal of Comparative Economics*, 36, pp. 179–194.
- Mitchell, T. (1997). *Machine Learning*. McGraw Hill.
- Mourão, P. (2006). Contributo para o estudo económico dos indicadores regionais. *Revista Portuguesa de Estudos Regionais*, n.º12, pp.77 a 92.
- National Consortium for the Study of Terrorism and Responses to Terrorism - START (2011). *Global Terrorism Database*. Disponível em <http://www.start.umd.edu/gtd>.
- Nazareth, J. (2004). *Demografia - Ciência da População*. Presença, Lisboa.
- Neumayer, E. e Plümpert, T. (2009). International terrorism and the clash of civilizations. *British Journal of Political Science*, 39(4), pp. 711–734.
- Newman, E. (2006). Exploring the “Root Causes” of Terrorism. *Studies in Conflict e Terrorism*, 29-8, pp. 749-772.
- Nitsch, V. e Schumacher, D. (2004). Terrorism and international trade: an empirical investigation. *European Journal of Political Economy* 20, pp. 423–433.
- Oliveira, J. (2009). *Contextualização teórica das políticas de crescimento económico em Portugal*. Dissertação de Mestrado em Administração e Gestão Pública. Universidade de Aveiro, Secção Autónoma de Ciências Sociais, Jurídicas e Políticas.
- Pape, R.(2006). *Morir para Ganar - Las estratégia del terrorismo suicida*. Barcelona: Editora Paidós Ibérica.
- Patterson, D. (1996). *Artificial Neural Networks - Theory and applications*. Prentice Hall, Singapore.
- Petersen, R. e Wiil, U. (2011). CrimeFighter Investigator: A Novel Tool for Criminal Network Investigation. *IEEE Computer Society*.
- Piazza, J. (2006). Rooted in poverty? Terrorism, poor economic development, and social cleavages. *Terrorism and Political Violence*, 18(1), pp. 159–177.
- Piazza, J. (2007). Draining the swamp: democracy promotion, state failure, and terrorism in 19 Middle Eastern countries. *Studies in Conflict and Terrorism*, 30(6), pp. 521–539.
- Piazza, J. (2008a). Do democracy and free markets protect us from terrorism? *International Politics*, 45(1), pp. 72–91.
- Piazza, J. (2008b). Incubators of terror: do failed and failing states promote transnational terrorism? *International Studies Quarterly* 52, pp. 469–488.
- Piazza, J. (2009). Economic development, poorly managed political conflict and terrorism in India. *Studies in Conflict and Terrorism* 32, pp. 406–419.
- Piazza, J. (2011). Poverty, minority economic discrimination, and domestic terrorism. *Journal of Peace Research* 48 (3), pp. 339–353.
- Pletka, D. (2003). Iraq, Conflict, and Beyond. Disponível na internet via: <http://www.aei.org/article/foreign-and-defense-policy/regional/middle-east-and-north-africa/iraq-conflict-and-beyond/>, (AEI - American Enterprise Institute For Public Policy Research).

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

- Plümper, T., e Neumayer, E. (2010). The friend of my enemy is my enemy: international alliances and international terrorism. *European Journal of Political Research*, 49(1), pp. 75–96.
- Powers, M. e Choi, S-W. (2012). Does Transnational Terrorism Reduce Foreign Direct Investment? Business-Related Versus non-Business-Related Terrorism. *Journal of Peace Research*, pp. 1-16.
- Quinlan, J. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers.
- Quivy, R. e Campenhoudt, L. (1997). *Manual de Investigação em Ciências Sociais*. Gradiva, Lisboa.
- Robison, K.; Crenshaw, E. e Jenkins, J. (2006). Ideologies of violence: the social origins of islamist and leftist transnational terrorism. *Social Forces*, 84(4), pp. 2009–2026.
- Rocha, M.; Cortez, P. e Neves, J. M. (2008). *Análise Inteligente de Dados- algoritmos e implementação em java*. FCA - Editora de Informática.
- Sambanis, N. (2008). Terrorism and civil war. In P. Keefer e N. Loayza (Eds.), *Terrorism, economic development, and political openness*. Cambridge: Cambridge University Press, pp. 174–208.
- Sandler, T. e Enders, W. (2008). Economic consequences of terrorism in developed and developing countries: an overview. In *Terrorism, Economic Development, and Political Openness*, edited by P. Keefer and N. Loayza. Cambridge: Cambridge University Press, pp. 17–47.
- Santos, F. e Azevedo, C. (2005). *Data Mining - Descoberta de Conhecimento em Bases de Dados*. FCA - Editora Informática.
- Sandroni, P. (1987). *Dicionário de economia*. Nova edição revista e ampliada. 7ªed., São Paulo.
- Sageman, M. (2004). *Understanding Terrorist Networks*. University of Pennsylvania Press.
- Savun, B. e Phillips, B. (2009). Democracy, foreign policy, and terrorism. *Journal of Conflict Resolution* 53(6), pp. 878–904.
- Siedenberg, D. (2003). Indicadores de desenvolvimento socioeconómico, uma síntese. *Desenvolvimento em questão*, Editora Unijuí, ano 1, n.º1, Janeiro/Junho, pp. 45 - 71.
- Tackrah, J. (1987). *Encyclopedia of Terrorism and Political Violence*. London. Routledge.
- Tavares, J. (2004). The open society assesses its enemies: shocks, disasters and terrorist attacks. *Journal of Monetary Economics* 51, pp. 1039–1070.
- Teorell, J.; Samanni, M.; Holmberg, S. e Rothstein, B. (2011). The Quality of Government Dataset, versão 6Abr11. University of Gothenburg: The Quality of Government Institute, <http://www.qog.pol.gu.se>.
- Testas, A. (2004). Determinants of terrorism in the Muslim world: an empirical cross-sectional analysis. *Terrorism and Political Violence*, 16(2), pp. 253–273.
- Torres, A. (1995). *Demografia e Desenvolvimento: Elementos Básicos*. Gradiva, Lisboa.
- UNRISD (United Nations Research Institute for Social Development - 1984). *Applicability of indicators of socioeconomic change for development planning*. Paris.
- Urdal, H. (2006). A clash of generations? Youth bulges and political violence. *International Studies Quarterly*, 50(3), pp. 607–629.
- US Department of State, *Patterns of Global Terrorism* (1998). Washington, DC, 1999.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Vegar, J. (2008). A Célula e as Fronteiras, Circulação e Posse de Informação no Terrorismo Jihadista, no Crime Organizado Contemporâneo e na Investigação de Segurança, *CIES-ISCTE e-WORKING PAPER* Nº 49/2008.

Ventura, J. P. e Nascimento, J. M. (2001). Violência, Terrorismo e Psicologia: uma abordagem exploratória. *Revista Portuguesa de Ciência Criminal*, 11 (4): 633-698.

Ventura, J. P. (2003). Terrorismo, a Criminalidade Organizada e a Segurança Doméstica. *Polícia e Justiça*, Loures, N.1 (Janeiro-Junho 2003), p.113-126.

Ventura, J. P. (2004). Terrorismo: da caracterização do fenómeno à reactividade proactiva. *Polícia e Justiça*, Loures, N.3 (Janeiro-Junho 2004), p.195-236.

Walsh, J. e Piazza, J. (2010). Why respecting physical integrity rights reduces terrorism. *Comparative Political Studies*, 43(5), pp. 551–577.

World Bank (2012). World Development Indicators. *The World Bank Group*. Washington DC, disponível em: <http://www.worldbank.org/> (Fevereiro de 2012).

#### Sites:

<http://data.worldbank.org/data-catalog> consultado em 01/02/2012;

<http://wits.nctc.gov/FederalDiscoverWITS/index.do?N=0> consultado em 10/01/2012;

[http://www.europa.eu/legislation\\_summaries/justice\\_freedom\\_security/fight\\_against\\_terrorism/l33275\\_pt.htm](http://www.europa.eu/legislation_summaries/justice_freedom_security/fight_against_terrorism/l33275_pt.htm) consultado em 20/12/2011

<http://www.europol.europa.eu> consultado em 20/12/2011;

<http://www.fbi.gov> consultado em 15/12/2011;

[http://www.fbi.gov/wanted/wanted\\_terrorists](http://www.fbi.gov/wanted/wanted_terrorists) consultado em 10/01/2012;

<http://www.globalization.kof.ethz.ch/> consultado em 01/03/2012

<http://www.healthmetricsandevaluation.org/ghdx> consultado em 01/02/2012;

<http://www.imf.org/external/index.htm> consultado em 01/08/2012;

<http://www.rand.org/nsrd/projects/terrorism-incidents.html> consultado em 10/01/2012;

<http://www.sis.pt> consultado em 21/12/2011;

<http://www.start.umd.edu/gtd/> consultado em 10/01/2012;

<http://www.start.umd.edu/start/> consultado em 10/01/2012;

<http://www.systemicpeace.org/polity/polity4.htm> consultado em 01/02/2012.

<http://www.unrisd.org/> consultado em 01/03/2012

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

## Anexos

### Anexo A: Listagem completa das variáveis de caracterização do terrorismo

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	NATUREZA	UNIDADE DE MEDIDA	DESCRIÇÃO
1	Hatent	Houve Atentados	-	Qualitativa	Escala Nominal	Se houve atentados no país nesse ano
2	IncTerror	Nº total incidentes terroristas	-	Quantitativa	Soma total v. discreta	Engloba o número total de incidentes terroristas ocorridos no país nesse ano
3	Sup24	Tem incidentes com duração superior a um dia	<i>Extended Incident</i>	Qualitativa	Escala Nominal	Se houve incidentes com duração superior a um dia nesse país nesse ano
4	NrSup24	N. Incidentes com duração superior a um dia	-	Quantitativa	Soma total v. discreta	Número total de incidentes com duração superior a um dia (+ de 24horas) nesse país nesse ano
5	NrInf24	Nº Incidentes com duração inferior a um dia	-	Quantitativa	Soma total v. discreta	Número total de incidentes com duração inferior a um dia (- de 24horas) nesse país nesse ano
6	Crit1	Nº Objectivo político, económico, religioso ou social	<i>Political, Economic, Religious, or Social Goal</i>	Quantitativa	Soma total v. discreta	Número total de incidentes registados nesse país nesse ano que se inserem dentro do critério de inclusão 1 (objectivo político, económico, religioso ou social)
7	Crit2	Nº Intenção de coagir, intimidar ou transmitir alguma outra mensagem a um público maior	<i>Intention to Coerce, Intimidate or Publicize to Larger Audience(s)</i>	Quantitativa	Soma total v. discreta	Número total de incidentes registados nesse país nesse ano que se inserem dentro do critério de inclusão 2 (Intenção de coagir, intimidar ou transmitir alguma outra mensagem a um público maior)
8	Crit3	Nº Têm que estar fora do contexto de actividades legítimas de guerra	<i>Outside International Humanitarian Law - The action must be outside the context of legitimate warfare activities</i>	Quantitativa	Soma total v. discreta	Número total de incidentes registados nesse país nesse ano que se inserem dentro do critério de inclusão 3 (Têm que estar fora do contexto de actividades legítimas de guerra - De acordo com os parâmetros permitidos pela lei humanitária internacional, como refletido no Protocolo Adicional das Convenções de Geneva de 12AGO1949
9	NrMult	Nº Ataques que fazem parte de um incidente múltiplo	<i>Part of Multiple Incident</i>	Quantitativa	Soma total v. discreta	Número de ataques que fazem parte de um incidente múltiplo - estando conectados apesar de terem ocorrido em horas ou locais diferentes
10	Sucesso	Nº ataques bem sucedidos (concretizados)	<i>Successful Attack</i>	Quantitativa	Soma total v. discreta	Número de ataques bem sucedidos, referindo-se à concretização do ataque planeado e não quanto aos objectivos que o motivaram
11	Suicidas	Nº ataques suicidas	<i>Suicide Attack</i>	Quantitativa	Soma total v. discreta	Número de ataques ocorridos nesse país nesse ano em que houve evidências de que o perpetrador não tinha intenção de escapar vivo do incidente
12	Assassinato	Nº ataques por assassinato	<i>Assassination</i>	Quantitativa	Soma total v. discreta	Número de ataques em que foi utilizado o método de assassinato (ataque em que o objectivo principal é matar uma pessoa pré-determinada, normalmente indivíduos proeminentes na sociedade)
13	AssArma do	Nº ataques por assalto à mão armada	<i>Armed Assault</i>	Quantitativa	Soma total v. discreta	Número de ataques em que foi utilizado o método de assalto à mão armada (ataque em que o principal objectivo é causar danos físicos ou morte a seres humanos - não individualizado - por outros meios que não através de explosivos
14	Explosiv	Nº ataques por deflagração ou detonação de explosivos	<i>Bombing/Explosion</i>	Quantitativa	Soma total v. discreta	Número de ataques em que foi utilizado o método de detonação de explosivos
15	Hijacking	Nº ataques por Hijacking	<i>Hijacking</i>	Quantitativa	Soma total v. discreta	Número de ataques em que foi utilizado o método de Hijacking (ato em que o objectivo é tomar o controlo de um veículo como um avião, um barco, um autocarro, ou outro, com o propósito de o desviar para um destino diferente, obter o pagamento de um resgate, forçar a libertação de prisioneiros, ou algum outro objectivo político)
16	RefBarr	Nº Tomada de reféns em barricada	<i>Hostage Taking (Barricade Incidente)</i>	Quantitativa	Soma total v. discreta	Número de incidentes em que o objectivo primário é a obtenção de concessões políticas ou outras em troca da libertação de reféns apanhados na sequência de barricada

**Data Mining no Contra-terrorismo**  
**Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo**

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	NATUREZA	UNIDADE DE MEDIDA	DESCRIÇÃO
17	RefSequest	Nº Tomada de reféns por sequestro	<i>Hostage Taking (kidnapping)</i>	Quantitativa	Soma total v. discreta	Número de incidentes em que o objectivo primário é a obtenção de concessões políticas ou outras em troca da libertação de reféns/prisioneiros. Distingue-se da anterior por manter o refém durante mais tempo, numa localização clandestina e desconhecida. As vítimas usualmente são previamente escolhidas
18	AtInfraest	Nº Ataques dirigidos a infraestruturas	<i>Facility/Infrastructure attack</i>	Quantitativa	Soma total v. discreta	Número de incidentes dirigidos a infraestruturas, estando nesta variável excluídos os actos com recurso a explosivos. O principal objectivo é causar dano a um alvo não-humano, tal como, monumento, edifício, comboio, gasoduto, etc
19	SemArmas	Nº ataques sem armas	<i>Unarmed Assault</i>	Quantitativa	Soma total v. discreta	Número de ataques em que o objectivo principal é causar danos físicos ou morte directamente aos seres humanos por quaisquer outros meios que não explosivo, arma de fogo, incendiários, ou instrumento afiado (faca, etc)
20	TipoDesc	Nº ataques de modus desconhecido	<i>Unknown Attack Type</i>	Quantitativa	Soma total v. discreta	Número de ataques em que não foi possível determinar o tipo de procedimento por entre a informação disponível
21	AlvNegóc	Nº ataques com alvos ligados a negócios	<i>Target Type Business</i>	Quantitativa	Soma total v. discreta	As empresas são definidas como indivíduos ou organizações que exerçam actividade comercial ou mercantil como um meio de subsistência. Esta variável engloba todos os ataques que tenham como alvo um negócio/empresa ou particulares que patrocinem um negócio, como um restaurante, posto de gasolina, loja de música, bar, café, etc
22	AlvGover	Nº ataques com alvo no governo	<i>Target Government</i>	Quantitativa	Soma total v. discreta	Engloba qualquer ataque tendo como alvo um edifício governamental, membro do governo, ex-membros, incluindo membros de partidos políticos, as suas caravanas, ou eventos patrocinados por partidos políticos, movimentos políticos, ou de uma instituição patrocinada pelo governo, onde o ataque é expressamente realizado para prejudicar o governo.
23	AlvPolicia	Nº ataques dirigidos a alvos policiais e das forças de segurança	<i>Target Police</i>	Quantitativa	Soma total v. discreta	Inclui ataques a membros das forças policiais ou instalações policiais, sedes patrulhas, academias, carros, postos de controlo, etc.
24	AlvMilit	Nº ataques a alvos militares	<i>Target Military</i>	Quantitativa	Soma total v. discreta	Inclui ataques contra unidades do exército, patrulhas, quartéis, e comboios, jipes, etc. Também inclui ataques a locais de recrutamento, e os soldados envolvidos em funções internas de policiamento, como em postos de controlo e em antinarcóticos
25	AlvAbort	Nº ataques tendo alvos relacionados com aborto	<i>Target Abortion Related</i>	Quantitativa	Soma total v. discreta	Inclui ataques contra clínicas de aborto, empregados, patrões, ou pessoal de segurança posicionados em clínicas
26	AlvAerop	Nº ataques tendo como alvos aeroportos e linhas aéreas	<i>Target Airports &amp; Airlines</i>	Quantitativa	Soma total v. discreta	Inclui ataques contra aviões ou contra aeroportos. Os ataques contra funcionários de empresas aéreas, enquanto a bordo e nos escritórios também estão incluídos neste valor.
27	AlvDiplo	Nº ataques a representações diplomáticas do Governo	<i>Target Government Diplomatic</i>	Quantitativa	Soma total v. discreta	Os incidentes abrangidos por esta variável incluem os ataques realizados contra missões estrangeiras, incluindo embaixadas, consulados, etc
28	AlvEscol	Nº ataques dirigidos a instituições escolares	<i>Target Educational Institution</i>	Quantitativa	Soma total v. discreta	Ataques contra escolas, professores, ou guardas que protegem locais da escola. Inclui ataques a professores universitários, docentes e autocarros escolares. Além disso, inclui ataques contra escolas religiosas.
29	AlvReser	Nº ataques dirigidos a reservas de água e comida	<i>Target Food or Water Supplies or reserves</i>	Quantitativa	Soma total v. discreta	Esta variável inclui todos os ataques dirigidos a reservas de água ou comida
30	AlvMedia	Nº ataques dirigidos a Jornalistas e alvos relacionados com os media	<i>Target Journalists e Media</i>	Quantitativa	Soma total v. discreta	Inclui ataques a jornalistas, assistentes de notícias, fotógrafos, editores, bem como ataques a sedes de media e escritórios. Ataques em instalações de transmissão, tais como antenas ou torres de transmissão, estão incluídos neste valor
31	AlvMarim	Nº ataques a alvos marítimos	<i>Target Maritime</i>	Quantitativa	Soma total v. discreta	Implica marinha civil. Inclui ataques a navios de pesca, navios de petróleo, balsas, iates, etc

**Data Mining no Contra-terrorismo**  
**Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo**

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	NATUREZA	UNIDADE DE MEDIDA	DESCRIÇÃO
32	AlvONGs	Nº ataques a alvos ligados a entidades e funcionários de ONGs	<i>Target NGOs</i>	Quantitativa	Soma total v. discreta	Inclui ataques a escritórios e funcionários de organizações não-governamentais. As ONGs são definidas num âmbito de uma abrangência multinacional, incluindo organizações como a Cruz Vermelha ou os Médicos sem Fronteiras. As forças de paz também pertencem a este valor.
33	AlvOutros	Nº ataques dirigidos a outros alvos	<i>Target Others</i>	Quantitativa	Soma total v. discreta	Esta variável inclui actos de terrorismo cometidos contra alvos que não se encaixam nas outras categorias.
34	AlvPart	Nº ataques a cidadãos particulares e propriedades particulares	<i>Target Private Citizens e Property</i>	Quantitativa	Soma total v. discreta	Inclui ataques contra indivíduos, público em geral ou ataques em áreas públicas incluindo mercados, ruas comerciais, cruzamentos movimentados, shopping centers. Não inclui ataques que causem vítimas civis em empresas como restaurantes, cafés ou salas de cinema (entram na variável alvos ligados a negócios acima descrita)
35	AlvRelig	Nº ataques dirigidos a instituições ou representações religiosas	<i>Target Religious Figures/Institutions</i>	Quantitativa	Soma total v. discreta	Inclui ataques a líderes religiosos (imans, sacerdotes, bispos, etc), instituições religiosas (mesquitas, igrejas), lugares ou objectos religiosos (santuários, relíquias, etc.). Esta variável também inclui ataques a organizações que estejam associadas com as entidades religiosas que não são ONGs, empresas ou escolas
36	AlvTelecom	Nº ataques a alvos ligados às telecomunicações	<i>Target Telecommunication</i>	Quantitativa	Soma total v. discreta	Inclui ataques a instalações e infra-estruturas para a transmissão de informações. Mais especificamente, este valor inclui valores abrangendo torres de telefones móveis, cabines telefónicas, transmissores de televisão, rádio e torres de microondas.
37	AlvTerror	Nº ataques dirigidos a terroristas ou grupos terroristas	<i>Target Terrorists</i>	Quantitativa	Soma total v. discreta	Terroristas ou membros de grupos terroristas estão incluídos neste valor. Inclui informantes de grupos terroristas, mas exclui ex-terroristas. Abrange a focalização das milícias e guerrilheiros (como alvos).
38	AlvTurist	Nº ataques tendo como alvo Turistas	<i>Target Tourists</i>	Quantitativa	Soma total v. discreta	Este valor inclui ataques a autocarros turísticos, turistas, ou "excursões". Turistas são pessoas que viajam principalmente para fins de lazer ou diversão. Escritórios de turismo do governo estão incluídos neste valor. O ataque deve claramente visar os turistas, não apenas um assalto num sistema de negócios ou de transporte utilizado pelos turistas.
39	AlvTransp	Nº ataques dirigidos a transportes públicos, exceto aviação	<i>Target Transportation (Other than aviation)</i>	Quantitativa	Soma total v. discreta	Ataques a sistemas de transportes públicos estão incluídos neste valor. Isso pode incluir assaltos a autocarros, comboios, metro, rodovias (se a estrada em si é o alvo do ataque), pontes, estradas, etc.
40	AlvDesconh	Nº ataques com alvos desconhecidos	<i>Target Unknown</i>	Quantitativa	Soma total v. discreta	Esta variável destina-se a todas as situações em que o tipo de alvo não pode ser determinado a partir da informação disponível.
41	AlvUtilit	Nº ataques dirigidos a alvos utilitários/serviços públicos	<i>Target Utilities</i>	Quantitativa	Soma total v. discreta	Este valor diz respeito a ataques dirigidos a instalações para a transmissão ou a geração de energia. Por exemplo, linhas de energia, oleodutos, transformadores eléctricos, linhas de alta tensão, gás e subestações eléctricas, estão incluídos neste valor. Este valor também inclui postes ou luzes de rua.
42	AlvGrp PolViol	Nº ataques dirigidos a grupos políticos violentos	<i>Target Violent Political Parties</i>	Quantitativa	Soma total v. discreta	Esta variável refere-se a entidades que são simultaneamente partidos políticos (e, portanto, codificado como "governo" neste esquema de codificação) e terroristas. É definida operacionalmente como grupos que se envolvem em política eleitoral e aparecem como "autores" do GTD.
43	NrMortos	Nº total de mortos	<i>Total Number of Fatalities</i>	Quantitativa	Soma total v. discreta	Este campo armazena o número total de mortes confirmadas para todos os incidentes ocorridos nesse ano nesse país. O número inclui todas as vítimas e atacantes que morreram como resultado direto do incidente.
44	NrFeridos	Nº total de feridos	<i>Total Number of Injured</i>	Quantitativa	Soma total v. discreta	Refere-se ao número total de feridos não fatais confirmados.
45	DanProp	Nº incidentes em que resultaram danos de propriedade	<i>Incidents Property Damage</i>	Quantitativa	Soma total v. discreta	Estão incluídos nesta variável todos os incidentes em que resultaram danos de propriedade.

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

**Anexo B: Caracterização dos impactos do terrorismo**

Medidas descritivas do n.º total de incidentes terroristas por ano

<b>Nº total de incidentes terroristas</b>								
Ano	Observ.	Média	Desv. Padrão	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
1990	173	21,71	67,24	0	0	0	6	495
1991	173	25,79	82,25	0	0	2	9	659
1992	173	28,01	74,70	0	1	4	21	524
1994	173	18,57	44,71	0	1	2	13	299
1995	173	16,98	58,50	0	0	2	9	666
1996	173	15,88	43,78	0	0	2	11	409
1997	173	17,19	57,61	0	0	2	12	598
1998	173	4,95	15,75	0	0	0	3	147
1999	173	7,20	19,14	0	0	0	4	110
2000	173	7,76	22,31	0	0	0	4	128
2001	173	7,53	23,79	0	0	0	2	195
2002	173	5,64	19,04	0	0	0	2	143
2003	173	6,65	22,38	0	0	0	2	195
2004	173	6,34	27,34	0	0	0	1	305
2005	173	11,13	53,01	0	0	0	2	619
2006	173	15,43	73,51	0	0	0	2	836
2007	173	17,66	90,15	0	0	0	2	1041
2008	173	27,04	110,66	0	0	1	5	1103
2009	173	26,57	121,41	0	0	0	3	1134
2010	173	26,86	124,82	0	0	0	3	1176
<b>Total</b>	<b>3460</b>	<b>15,74</b>	<b>67,27</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>5</b>	<b>1176</b>

Medidas descritivas do n.º total de mortos por ano

<b>Nº total de mortos</b>								
Ano	Observ.	Média	Desv. Padrão	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
1990	173	40,54	175,74	0	0	0	3	1505
1991	173	47,86	177,30	0	0	0	8	1429
1992	173	54,86	179,53	0	0	1	21	1247
1994	173	45,86	163,31	0	0	1	17	1571
1995	173	34,65	121,45	0	0	0	8	1020
1996	173	38,69	148,26	0	0	0	10	1371
1997	173	62,15	345,54	0	0	0	9	4145
1998	173	27,18	110,81	0	0	0	0	882
1999	173	18,73	67,08	0	0	0	2	448
2000	173	19,39	72,24	0	0	0	1	557
2001	173	36,72	241,07	0	0	0	0	3003
2002	173	19,26	79,52	0	0	0	0	529
2003	173	17,17	63,78	0	0	0	0	452
2004	173	32,50	177,18	0	0	0	0	2090
2005	173	35,77	258,44	0	0	0	0	3337
2006	173	53,52	366,51	0	0	0	1	4591
2007	173	72,87	518,48	0	0	0	1	6534
2008	173	51,06	257,71	0	0	0	2	2841
2009	173	49,54	248,78	0	0	0	2	2573
2010	173	42,38	221,87	0	0	0	1	2041
<b>Total</b>	<b>3460</b>	<b>40,04</b>	<b>228,62</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>3</b>	<b>6534</b>



**Data Mining no Contra-terrorismo**  
 Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Medidas descritivas do n.º total de feridos por ano

<b>Nº total de feridos</b>								
Ano	Observ.	Média	Desv. Padrão	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
1990	173	34,87	18,38	0	0	0	3	1042
1991	173	42,97	15,28	0	0	0	4	1326
1992	173	55,73	17,72	0	0	1	23	917
1994	173	41,38	107,59	0	0	1	21	657
1995	173	81,42	492,75	0	0	1	16	6225
1996	173	60,05	239,87	0	0	1	11	2216
1997	173	52,42	196,03	0	0	0	16	1834
1998	173	46,97	318,40	0	0	0	1	4000
1999	173	28,33	92,73	0	0	0	2	676
2000	173	28,25	105,56	0	0	0	3	805
2001	173	27,84	105,45	0	0	0	1	886
2002	173	32,51	146,57	0	0	0	0	1405
2003	173	35,23	158,24	0	0	0	1	1261
2004	173	66,82	355,08	0	0	0	0	3961
2005	173	71,93	470,52	0	0	0	0	5974
2006	173	88,77	658,47	0	0	0	1	8256
2007	173	120,56	936,48	0	0	0	1	1965
2008	173	101,08	554,35	0	0	0	4	6637
2009	173	119,25	780,35	0	0	0	4	9373
2010	173	87,20	573,15	0	0	0	4	6745
<b>Total</b>	<b>3460</b>	<b>61,18</b>	<b>417,71</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>4</b>	<b>11965</b>

Ano	Média de incidentes terroristas	Média de mortos	Média de feridos
1990	22	41	35
1991	26	48	43
1992	28	55	56
1993	4	0	0
1994	19	46	41
1995	17	35	81
1996	16	39	60
1997	17	62	52
1998	5	27	47
1999	7	19	28
2000	8	19	28
2001	8	37	28
2002	6	19	33
2003	7	17	35
2004	6	33	67
2005	11	36	72
2006	15	54	89
2007	18	73	121
2008	27	51	101
2009	27	50	119
2010	27	42	87

Ano	Total de incidentes terroristas	Total de mortos	Total de feridos
1990	3756	7013	6032
1991	4462	8280	7434
1992	4845	9490	9641
1993	691	0	0
1994	3213	7934	7159
1995	2937	5995	14085
1996	2747	6693	10389
1997	2974	10752	9068
1998	856	4702	8125
1999	1245	3241	4901
2000	1343	3355	4888
2001	1302	6352	4816
2002	976	3332	5624
2003	1151	2971	6095
2004	1096	5623	11560
2005	1925	6189	12444
2006	2670	9259	15357
2007	3055	12607	20857
2008	4678	8833	17487
2009	4597	8571	20631
2010	4646	7332	15086

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

**Anexo C: Caracterização dos países**

**Listagem das variáveis de caracterização económica**

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	FONTE	NATUREZA	UNIDADE DE MEDIDA	PERÍODO	N.º OBSERV.	ENTRADA NO ESTUDO
1	CombExp_n_1	Combustíveis, Exportações de (n-1)	<i>Fuel exports (% of merchandise exports)</i>	<i>World Bank</i>	Quantitativa Contínua	% das exportações de mercadorias 0 a 99,74	1989 a 2009	2460	SIM
2	CombImp_n_1	Combustíveis, Importações de (n-1)	<i>Fuel imports (% of merchandise imports)</i>	<i>World Bank</i>	Quantitativa Contínua	% das importações de mercadorias 0 a 55,63	1989 a 2009	2536	SIM
3	Comerc_n_1	Comércio (n-1)	<i>Trade (% of GDP)</i>	<i>World Bank</i>	Quantitativa Contínua	% do PIB 11,09 a 445,91	1989 a 2009	3232	SIM
4	Cred_FMI_n_1	Utilização de crédito do FMI (DOD, corrente EUA \$) (n-1)	<i>Use of IMF credit (DOD, current US\$)</i>	<i>World Bank</i>	Quantitativa Discreta	DOD, corrente EUA \$ 0 a 28317356000	1989 a 2009	2327	SIM
5	CrescPIB_n_1	Crescimento do PIB (n-1)	<i>GDP growth (annual %)</i>	<i>World Bank</i>	Quantitativa Contínua	% annual -51,03 a 106,28	1989 a 2009	3315	SIM
6	CrescPIBpc_n_1	Crescimento do PIB per capita (n-1)	<i>GDP per capita growth (annual %)</i>	<i>World Bank</i>	Quantitativa Contínua	% annual -50,29 a 92,59	1989 a 2009	3312	SIM
7	ConsumF_n_1	Despesa de consumo final (n-1)	<i>Final consumption expenditure, etc. (% of GDP)</i>	<i>World Bank</i>	Quantitativa Contínua	% do PIB 13,12 a 242,55	1989 a 2009	3171	SIM
8	ConsGov_n_1	Despesa geral de consumo final do Governo (n-1)	<i>General government final consumption expenditure (% of GDP)</i>	<i>World Bank</i>	Quantitativa Contínua	% do PIB 2,05 a 76,22	1989 a 2009	3142	SIM
9	DespNac_n_1	Despesa Nacional Bruta (n-1)	<i>Gross national expenditure (% of GDP)</i>	<i>World Bank</i>	Quantitativa Contínua	% do PIB 46,34 a 63368,50	1989 a 2009	3166	SIM
10	FormBrCap_n_1	Formação bruta de capital (n-1)	<i>Gross capital formation (% of GDP)</i>	<i>World Bank</i>	Quantitativa Contínua	% do PIB 1,53 a 113,58	1989 a 2009	3167	SIM
11	FormBrCapCr_n_1	Formação bruta de capital, cresc. (n-1)	<i>Gross capital formation (annual % growth)</i>	<i>World Bank</i>	Quantitativa Contínua	crescimento anual em % -81,77 a 723,20	1989 a 2009	2596	SIM
12	InfDefPIB_n_1	Inflação, deflator do PIB (n-1)	<i>Inflation, GDP deflator (annual %)</i>	<i>World Bank</i>	Quantitativa Contínua	% annual -33,58 a 26762,02	1989 a 2009	3307	SIM
13	InfPreCons_n_1	Inflação, preços de consumo (n-1)	<i>Inflation, consumer prices (annual %)</i>	<i>World Bank</i>	Quantitativa Contínua	% annual -16,12 a 24411,03	1989 a 2009	3003	SIM
14	InvEstrDir_n_1	Investimento estrangeiro directo, entradas líquidas (n-1)	<i>Foreign direct investment, net inflows (% of GDP)</i>	<i>World Bank</i>	Quantitativa Contínua	% do PIB -82,89 a 564,92	1989 a 2009	3202	SIM
15	Pib_PerCap_n_1	PIB per capita, PPP (n-1)	<i>GDP per capita, PPP (purchasing power parity)</i>	<i>World Bank</i>	Quantitativa Contínua	Constant. Internac. 2005 \$ 140,02 a 77108,22	1989 a 2009	3275	SIM
16	PoupIntBr_n_1	Poupanças internas brutas (n-1)	<i>Gross domestic savings (% of GDP)</i>	<i>World Bank</i>	Quantitativa Contínua	% do PIB -142,55 a 86,88	1989 a 2009	3171	SIM
17	ConsGov_Pib_n_1	Quota de consumo do Governo sobre o PIB per capita (n-1)	<i>Government Consumption Share of PPP Converted GDP Per Capita at 2005 constant prices</i>	<i>PENN World Table</i>	Quantitativa Contínua	% - preços constantes de 2005, convertido em PPP 0,90 a 65,89	1989 a 2009	3333	SIM

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	FONTE	NATUREZA	UNIDADE DE MEDIDA	PERÍODO	N.º OBSERV.	ENTRA NO ESTUDO
18	ConsPib_n_1	Quota de consumo do PIB per capita (n-1)	<i>Consumption Share of PPP Converted GDP Per Capita at 2005 constant prices</i>	<i>PENN World Table</i>	Quantitativa Contínua	% - preços constantes de 2005, convertido em PPP 7,57 a 221,09	1989 a 2009	3333	SIM
19	Invest_Pib_n_1	Quota de Investimento do PIB per capita (n-1)	<i>Investment Share of PPP Converted GDP Per Capita at 2005 constant prices</i>	<i>PENN World Table</i>	Quantitativa Contínua	% - preços constantes de 2005, convertido em PPP -33,14 a 86,34	1989 a 2009	3333	SIM
20	Rnb_CresPerCap_n_1	RNB, crescimento per capita de (n-1)	<i>GNI per capita growth (annual %)</i>	<i>World Bank</i>	Quantitativa Contínua	% anual -47,14 a 101,41	1989 a 2009	2593	SIM
21	Rnb_Cresc_n_1	RNB, crescimento de (n-1)	<i>GNI growth (annual %)</i>	<i>World Bank</i>	Quantitativa Contínua	% anual -50,11 a 99,810	1989 a 2009	2593	SIM
22	S_P_Ind_n_1	S&P Índices de acções globais (variação anual em%) (n-1)	<i>S&amp;P Global Equity Indices (annual % change)</i>	<i>World Bank</i>	Quantitativa Contínua	variação anual em% -84,23 a 912,28	1989 a 2009	<b>1086</b>	<b>NÃO</b> – Muitos missing values

**Listagem das variáveis de caracterização sociodemográfica**

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	FONTE	NATUREZA	UNIDADE DE MEDIDA	PERÍODO	N.º OBSERV.	ENTRA NO ESTUDO
1	ContSoc_n_1	Contribuições sociais (n-1)	<i>Social contributions (% of revenue)</i>	<i>World Bank</i>	Quantitativa Contínua	% da receita 0 a 60,22	1989 a 2009	<b>1050</b>	<b>NÃO</b> – Muitos missing values
2	CrescPop_n_1	Crescim. populacional (n-1)	<i>Population growth (annual %)</i>	<i>World Bank</i>	Quantitativa Contínua	% anual -7,53 a 18,59	1989 a 2009	3456	SIM
3	DensPop_n_1	Densidade populacional (n-1)	<i>Population density (people per sq. km of land area)</i>	<i>World Bank</i>	Quantitativa Contínua	Pessoas por km <sup>2</sup> de área de terra 1,65 a 7125,14	1989 a 2009	3437	SIM
4	DesempJov_n_1	Desemprego, total jovens(n-1)	<i>Unemployment, youth total (% of total labor force ages 15-24)</i>	<i>World Bank</i>	Quantitativa Contínua	% da força laboral total idades 15-24 0,7 a 70,9	1989 a 2009	<b>1352</b>	<b>NÃO</b> – Muitos missing values
5	DespSau_n_1	Despesa total em saúde (n-1)	<i>Health expenditure, total (% of GDP)</i>	<i>World Bank</i>	Quantitativa Contínua	% do PIB 0 a 19,31	1995 a 2009	2553	SIM
6	DtsEconM_n_1	Direitos Económicos da Mulher (n-1)	<i>Women's Economic Rights</i>	<i>The Cingranelli Richards Human Rights Dataset</i>	Qualitativa Ordinal	0=Nenhum 1=Há alguns na lei, mas nenhuns na prática 2=Há alguns na lei e na prática 3=Todos na lei e na prática	1989 a 2009	3055	SIM
7	DtsTrab_n_1	Direitos dos Trabalhadores (n-1)	<i>Worker's Rights</i>	<i>The Cingranelli-Richards Human Rights Dataset</i>	Qualitativa Ordinal	0=Extremamente restritos 1=Parcialmente restritos 2=Totalmente protegidos	1989 a 2009	3081	SIM

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	FONTE	NATUREZA	UNIDADE DE MEDIDA	PERÍODO	N.º OBSERV.	ENTRA NO ESTUDO
8	EducHom_n_1	Média de Anos de Educação (Homens) (n-1)	<i>Average Years of Education (Male)</i>	<i>Institute for Health Metrics &amp; Evaluation - University of Washington (Gakidou et al. 2010)</i>	Quantitativa Contínua	Média de anos 0,6 a 14,2	1989 a 2009	3240	SIM
9	EducMulh_n_1	Média de Anos Educação (Mulheres) (n-1)	<i>Average Years of Education (Female)</i>	<i>Institute for Health Metrics &amp; Evaluation - University of Washington (Gakidou et al. 2010)</i>	Quantitativa Contínua	Média de anos 0,1 a 14,2	1989 a 2009	3240	SIM
10	EscPrim_n_1	Escolarização, ensino primário (n-1)	<i>School enrollment, primary (% gross)</i>	<i>World Bank</i>	Quantitativa Contínua	% bruto/total matrículas para essa faixa 19,32 a 158,75	1989 a 2009	2807	SIM
11	EscSecun_n_1	Escolarização, ensino secundário (n-1)	<i>School enrollment, secondary (% gross)</i>	<i>World Bank</i>	Quantitativa Contínua	% bruto/total matrículas para essa faixa 4,47 a 162,35	1989 a 2009	2457	SIM
12	EscTerc_n_1	Escolarização, ensino terciário (n-1)	<i>School enrollment, tertiary (% gross)</i>	<i>World Bank</i>	Quantitativa Contínua	% bruto/total matrículas para essa faixa 0 a 118,10	1989 a 2009	2165	SIM
13	ExpVida_n_1	Expectativa de vida ao nascer, total (n-1)	<i>Life expectancy at birth, total (years)</i>	<i>World Bank</i>	Quantitativa Contínua	Anos 26,82 a 82,93	1989 a 2009	3367	SIM
14	FragEtnic_n_1	Fragmentação étnica (n-1)	<i>Ethnic fractionalization</i>	<i>Alesina et al. (2003)</i>	Quantitativa Contínua	Valor de probabilidade entre 0 e 1 0 a 0,93	1989 a 2009	3380	SIM
15	FragLing_n_1	Fragmentação linguística (n-1)	<i>Linguistic fractionalization</i>	<i>Alesina et al. (2003)</i>	Quantitativa Contínua	Valor de probabilidade entre 0 e 1 0 a 0,92	1989 a 2009	3277	SIM
16	GastEduc_n_1	Gastos públicos com educação, total (n-1)	<i>Public spending on education, total (% of GDP)</i>	<i>World Bank</i>	Quantitativa Contínua	% do PIB 0a 16,06	1989 a 2009	1247	NAO – Muitos missing values
17	GiniIndex_n_1	GINI Index (n-1)	<i>GINI Index</i>	<i>World Bank</i>	Quantitativa Contínua	Valor em [0,100] 19,49 a 67,40	1989 a 2009	696	NAO – Muitos missing values
18	HiatoPob_n_1	Hiato de pobreza na linha de pobreza nacional (n-1)	<i>Poverty gap at national poverty line (%)</i>	<i>World Bank</i>	Quantitativa Contínua	% 0,4 a 62,2	1989 a 2009	186	NAO – Muitos missing values
19	LinhTelef_n_1	Linhas telefónicas (n-1)	<i>Telephone lines (per 100 people)</i>	<i>World Bank</i>	Quantitativa Contínua	por 100 pessoas 0,01 a 74,46	1989 a 2009	3418	SIM
20	Pop15_64_n_1	População idade em 15-64 (n-1)	<i>Population ages 15-64 (% of total)</i>	<i>World Bank</i>	Quantitativa Contínua	% do total 44,79 a 84,99	1989 a 2009	3359	SIM
21	Pop_mais65_n_1	População com 65 anos e acima (n-1)	<i>Population ages 65 and above (% of total)</i>	<i>World Bank</i>	Quantitativa Contínua	% do total 0,45 a 22,05	1989 a 2009	3359	SIM
22	Pop0_14_n_1	População de 0-14 (n-1)	<i>Population ages 0-14 (% of total)</i>	<i>World Bank</i>	Quantitativa Contínua	% do total 13,44 a 53,03	1989 a 2009	3359	SIM
23	Pop_Urb_n_1	População urbana (n-1)	<i>Urban population (% of total)</i>	<i>World Bank</i>	Quantitativa Contínua	% do total 5,34 a 100	1989 a 2009	3459	SIM
24	Pop_Total_n_1	População, total (n-1)	<i>Population, total</i>	<i>World Bank</i>	Quantitativa Discreta	Pessoas 40130 a 1331380000	1989 a 2009	3457	SIM

**Data Mining no Contra-terrorismo**  
 Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	FONTE	NATUREZA	UNIDADE DE MEDIDA	PERÍODO	N.º OBSERV.	ENTRA NO ESTUDO
25	MortMaterna_n_1	Rácio de Mortalidade Materna (por 100.000 nascidos vivos) (n-1)	<i>Maternal Mortality Ratio (per 100,000 Live Births)</i>	<i>Institute for Heal&amp; Evaluation - University of Washington</i>	Quantitativa Contínua	Pessoas 3,9 a 2110,2	1989 a 2009	3192	SIM
26	Sub5_TxMort_n_1	Sub-5 Taxa de Mortalidade (por 1.000 nascidos vivos) (n-1)	<i>Under-5 Mortality Rate (per 1,000 Live Births)</i>	<i>Institute for Health Metrics &amp; Evaluation - University of Washington</i>	Quantitativa Contínua	% 2,66 a 301,46	1989 a 2009	3440	SIM
27	Alfab_Adult_n_1	Taxa de alfabetização, total de adultos (% de pessoas com idades entre 15 e acima) (n-1)	<i>Literacy rate, adult total (% of people ages 15 and above)</i>	<i>World Bank</i>	Quantitativa Contínua	% de pessoas com idades entre 15 e acima 9,39 a 99,99	1989 a 2009	385	<b>NÃO</b> – Muitos missing values
28	MortNeonat_n_1	Taxa de Mortalidade Neonatal (por 1.000 nascimentos) (n-1)	<i>Neonatal Mortality Rate (per 1,000 Births)</i>	<i>Institute for Health Metrics &amp; Evaluation - University of Washington</i>	Quantitativa Contínua	% 1,37 a 70,81	1989 a 2009	3440	SIM
29	MortPosNeo_n_1	Taxa de mortalidade pós-neonatal (por 1.000 nascimentos) (n-1)	<i>Postneonatal Mortality Rate (per 1,000 Births)</i>	<i>Institute for Health Metrics &amp; Evaluation - University of Washington</i>	Quantitativa Contínua	% 0,56 a 96,81	1989 a 2009	3440	SIM
30	PartLabMulh_n_1	Taxa de participação laboral, mulheres (% da população do sexo feminino 15+) (n-1)	<i>Labor participation rate, female (% of female population ages 15+)</i>	<i>World Bank</i>	Quantitativa Contínua	% 8,90 a 90,70	1989 a 2009	3159	SIM
31	Tot_Desemp_n_1	Total de desemprego (% da força de trabalho total) (n-1)	<i>Unemployment, total (% of total labor force)</i>	<i>World Bank</i>	Quantitativa Contínua	% 0,3 a 59,5	1990 a 2009	1765	<b>NÃO</b> – Muitos missing values
32	Util_Internet_n_1	Utilizadores de Internet (por cada 100 pessoas) (n-1)	<i>Internet users (per 100 people)</i>	<i>World Bank</i>	Quantitativa Contínua	por cada 100 pessoas 0 a 92,18	1989 a 2009	2824	SIM

**Listagem das variáveis de caracterização política**

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	FONTE	NATUREZA	UNIDADE DE MEDIDA	PERÍODO	N.º OBSERV.	ENTRA NO ESTUDO
1	AutInst_n_1	Autocracia institucionalizada(n-1)	<i>Institutionalized Autocracy</i>	<i>Polity IV Project</i>	Qualitativa Ordinal	0= Baixo, a 10= Elevado	1989 a 2009	2885	SIM
2	AutodElec_n_1	Autodeterminação Eleitoral (n-1)	<i>Electoral Self-Determination</i>	<i>The Cingranelli-Richards Human Rights Dataset</i>	Qualitativa Ordinal	0= Não existe na lei, nem na prática 1= Existe na lei, mas não na prática 2= Existe na lei e na prática	1989 a 2009	3080	SIM

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	FONTE	NATUREZA	UNIDADE DE MEDIDA	PERÍODO	N.º OBSERV.	ENTRA NO ESTUDO
3	CompPart_n_1	Competitividade da Participação Política (n-1)	<i>Competitive ness of Participation</i>	<i>Polity IV Project</i>	Qualitativa Nominal	0=Não aplicável, políticas não reguladas 1=Reprimida, a 5= Competitiva	1989 a 2009	2885	SIM
4	Democ_n_1	Democracia (n-1)	<i>Democracy</i>	<i>Cheibub et al. (2009)</i>	Qualitativa Nominal	0=Não 1=Sim	1989 a 2008	3208	SIM
5	Democ_Inst_n_1	Democracia institucionalizada «Implantação/institucionalização» (n-1)	<i>Institutionalized Democracy</i>	<i>Polity IV Project</i>	Qualitativa Ordinal	0= Baixo, a 10= Elevado	1989 a 2009	2885	SIM
6	Democ_Nív_n_1	Democracia - nível de (n-1)	<i>Democracy</i>	<i>Freedom House - Freedom in the World</i>	Quantitativa Contínua	0=Menos democrático 10=Mais democrático	1989 a 2009	3383	SIM
7	DtosIntFis_n_1	Índice de Direitos de Integridade física (n-1)	<i>Physical Integrity Rights Index</i>	<i>The Cingranelli-Richards Human Rights Dataset</i>	Qualitativa Ordinal	0=Nenhum respeito pelos direitos;a 8=Total respeito pelos direitos	1989 a 2009	3072	SIM
8	DtsPolit_n_1	Direitos Políticos (n-1)	<i>Political Rights</i>	<i>Freedom House - Freedom in the World</i>	Qualitativa Ordinal	1=Mais liberdade,a 7=Menos liberdade	1989 a 2009	3383	SIM
9	DtsPolM_n_1	Direitos Políticos da Mulher (n-1)	<i>Women's Political Rights</i>	<i>The Cingranelli-Richards Human Rights Dataset</i>	Qualitativa Ordinal	0=Não estão na lei 1=Estão na lei, mas não na prática 2=Estão na lei, mas parcialmente na prática 3=Estão na lei e na prática	1989 a 2009	3070	SIM
10	Durab_Reg_n_1	Durabilidade de regime (n-1)	<i>Durable</i>	<i>Polity IV Project</i>	Quantitativa Discreta	Anos 0 a 200	1989 a 2009	3041	SIM
11	FracPol_n_1	Fraccionamento político total (n-1)	<i>Total Fractionalization</i>	<i>Database of Political Institutions</i>	Quantitativa Contínua	Valor de probabilidade entre 0 e 1 0 a 1	1989 a 2009	2861	SIM
12	FragPol_n_1	Fragmentação política (n-1)	<i>Polity Fragmentation</i>	<i>Polity IV Project</i>	Qualitativa Ordinal	0=Nenhuma 1=Fraca 2=Moderada 3=Séria	2000 a 2009	1561	NÃO - Poucos anos para estudo
13	InstReg_n_1	Instituições em Regime (n-1)	<i>Regime Institutions</i>	<i>Cheibub et al. (2009)</i>	Qualitativa - Nominal	0 - Democracia parlamentar; 1 - Democracia semi-presidencial; 2 - Democracia presidencial; 3 - Ditadura civil; 4 - Ditadura militar; 5 - Ditadura real (monárquica).	1989 a 2009	3208	SIM
14	LiberAss_n_1	Liberdade de assembleia e associação (n-1)	<i>Freedom of Assembly and Association</i>	<i>The Cingranelli-Richards Human Rights Dataset</i>	Qualitativa - Ordinal	0=Extremamente restrita para todos 1=Limitada para todos ou extremamente restrita para alguns 2=Livre para todos	1989 a 2009	3081	SIM

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	FONTE	NATUREZA	UNIDADE DE MEDIDA	PERÍODO	N.º OBSERV.	ENTRA NO ESTUDO
15	LiberCivil_n_1	Liberdades Civas (n-1)	<i>Civil Liberties</i>	<i>Freedom House - Freedom in the World</i>	Qualitativa Ordinal	1 - Muita liberdade;a 7 - Pouca liberdade	1989 a 2009	3383	SIM
16	LiberExp_n_1	Liberdade de expressão (n-1)	<i>Freedom of Speech</i>	<i>The Cingranelli-Richards Human Rights Dataset</i>	Qualitativa Ordinal	0 - Completa; 1 - Alguma; 2 - Nenhuma	1989 a 2009	3081	SIM
17	Pris_Polit_n_1	Prisão por Motivos Políticos (n-1)	<i>Political Imprisonment</i>	<i>The Cingranelli-Richards Human Rights Dataset</i>	Qualitativa Ordinal	0 - Sim, muitas 1 - Sim, poucas 2 - Nenhuma	1989 a 2009	3079	SIM
18	RegPartPol_n_1	Regulamentação de Participação Política (n-1)	<i>Regulation of Participation</i>	<i>Polity IV Project</i>	Qualitativa Nominal	1=Não regulada (sem grupos políticos definidos) 2=Grupos políticos estáveis que competem 3=Sectária (postura intransigente entre grupos políticos) 4=Restrita (permitida mas de fraco sectarismo) 5=Regulada	1989 a 2009	2885	SIM
19	RestExecut_n_1	Restrições Executivas (As regras de decisão) (n-1)	<i>Executive Constraints (Decision Rules)</i>	<i>Polity IV Project</i>	Qualitativa Ordinal	1=Autoridade executiva ilimitada;a 7=Sem autoridade, ou subordinação aos outros grupos políticos	1989 a 2009	2885	SIM
20	Status_Lib_n_1	Status de Liberdade (n-1)	<i>Status</i>	<i>Freedom House - Freedom in the World</i>	Qualitativa Ordinal	1 - Liberdade 2 - Liberdade parcial 3 - Não há liberdade	1989 a 2009	3383	SIM
21	Tortura_n_1	Tortura (n-1)	<i>Torture</i>	<i>The Cingranelli-Richards Human Rights Dataset</i>	Qualitativa Ordinal	Escala Ordinal 0 - Praticada frequentemente; 1 - Praticada ocasionalmente; 2 - Nunca ocorreu	1989 a 2009	3081	SIM

Data Mining no Contra-terrorismo  
Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

**Listagem das variáveis de caracterização religiosa**

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	FONTE	NATUREZA	UNIDADE DE MEDIDA	PERÍODO	N.º OBSERV.	ENTRA NO ESTUDO
1	FragRel_n_1	Fragmentação religiosa (n-1)	<i>Religious fractionalization</i>	<i>Alesina et al. (2003)</i>	Quantitativa Contínua	Valor de probabilidade entre 0 e 1 0,00 a 0,86	1989 a 2009	3417	SIM
2	Relig_Catol_n_1	Religião Católica (n-1)	<i>Religion: Catholic</i>	<i>La Porta et al. (1999)</i>	Quantitativa Contínua	% da população total 0 a 99,10	1989 a 2009	3417	SIM
3	Relig_Mucul_n_1	Religião Muçulmana (n-1)	<i>Religion: Muslim</i>	<i>La Porta et al. (1999)</i>	Quantitativa Contínua	% da população total 0 a 99,90	1989 a 2009	3417	SIM
4	Relig_Outr_r_n_1	Outras religiões (percentagem da população total) (n-1)	<i>Religion: Other Denomination</i>	<i>La Porta et al. (1999)</i>	Quantitativa Contínua	% da população total 0 a 100	1989 a 2009	3377	SIM

**Listagem de outras variáveis**

	CÓDIGO	DESIGNAÇÃO	INGLÊS-ORIGINAL	FONTE	NATUREZA	UNIDADE DE MEDIDA	PERÍODO	N.º OBSERV.	ENTRA NO ESTUDO
1	AreaSup_n_1	Área de superfície (km <sup>2</sup> ) - área total geográfica do país (n-1)	<i>Surface area (sq. km)</i>	<i>World Bank</i>	Quantitativa Contínua	Km <sup>2</sup> 260 a 17098240	1989 a 2009	3440	SIM
2	AreaTerr_n_1	Área terrestre (km <sup>2</sup> ) (n-1)	<i>Land area (sq. km)</i>	<i>World Bank</i>	Quantitativa Contínua	Km <sup>2</sup> 260 a 16389950	1989 a 2009	3440	SIM
3	DespMilitt_n_1	Despesas militares (n-1)	<i>Military expenditure (% of GDP)</i>	<i>World Bank</i>	Quantitativa Contínua	em % do PIB 0,05 a 117,39	1989 a 2009	2678	SIM
4	Kof_Index_n_1	KOF - Índice de Globalização (n-1)	<i>KOF Index of Globalization</i>	<i>Dreher - KOF Index of Globalization</i>	Quantitativa Contínua	Intervalo de 1 a 100 12 a 93	1989 a 2009	3340	SIM
5	MortComb_n_1	Mortes relacionadas com Combate (número de pessoas) (n-1)	<i>Battle-related deaths (number of people)</i>	<i>World Bank</i>	Quantitativa Discreta	número de pessoas 0 a 50293	1989 a 2009	<b>595</b>	<b>NÃO</b> – Muitos missing values



**Data Mining no Contra-terrorismo**  
 Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

**Medidas descritivas das variáveis económicas**

A análise descritiva dos indicadores económicos permite perceber uma elevada dispersão nos dados, pelos valores de desvio padrão apresentados. O peso do desvio padrão em relação à média varia entre 19,5% e 1326%, sendo superior a 100% em mais de metade das variáveis, demonstrando assim que as médias não se afiguram boas medidas para caracterizar a sua distribuição. De facto, a variabilidade dos indicadores é muito elevada, como se comprova pelo diferencial entre os mínimos e máximos apurados e, também, entre os quartis 1 e 3. A situação económica dos países difere muito, havendo países mais ricos e que apresentam indicadores de crescimento positivos, e países com situação económica frágil.

Atributos explicativos	Observ.	Média	Desv. Padrão	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
Combustíveis, % das exportações de mercadorias (n-1)	2460	15,36	26,44	0	0,56	3,09	13,11	99,74
Combustíveis, % das importações de mercadorias (n-1)	2536	12,45	8,28	0	6,63	11,20	16,64	55,63
Comércio (em % do PIB) (n-1)	3232	84,03	46,85	11,09	52,56	74,51	104,40	445,91
Crescimento do PIB (% anual) (n-1)	3315	3,64	6,47	-51,03	1,51	3,94	6,13	106,28
Crescimento do PIB per capita (% anual) (n-1)	3312	2,01	6,22	-50,29	-0,08	2,25	4,50	92,59
Despesa de consumo final (% do PIB) (n-1)	3171	82,79	16,12	13,12	74,78	82,19	91,23	242,55
Despesa Nacional Bruta (% do PIB) (n-1)	3166	144,87	1567,06	46,34	97,88	103,94	111,87	63368,50
Despesas gerais de consumo final do Governo (% do PIB) (n-1)	3142	16,23	6,68	2,05	11,33	15,57	19,96	76,22
Formação bruta de capital (% do PIB) (n-1)	3167	22,69	8,45	1,53	17,91	21,70	26,34	113,58
Formação bruta de capital (crescimento anual em%) (n-1)	2596	6,02	25,14	-81,77	-3,15	5,30	13,25	723,20
Inflação, deflator do PIB (% anual) (n-1)	3307	53,77	616,13	-33,59	2,48	5,89	13,07	26762,02
Inflação, preços de consumo (% anual) (n-1)	3003	50,78	673,15	-16,12	2,31	5,31	11,36	24411,03
Investimento estrangeiro direto, entradas líquidas (% do PIB) (n-1)	3202	5,24	23,75	-82,89	0,68	2,17	5,02	564,92
PIB per capita, PPP (n-1)	3275	10601,84	12665,96	140,02	1770,76	5543,03	14980,69	77108,22
Poupanças internas brutas (% do PIB) (n-1)	3171	17,21	16,12	-142,55	8,77	17,81	25,22	86,88
Quota de consumo do Governo sobre o PIB per capita (n-1)	3333	11,88	7,80	0,90	6,99	9,68	13,99	65,89
Quota de consumo do PIB per capita (n-1)	3333	69,47	19,48	7,57	61,18	70,39	80,19	221,09
Quota de Investimento do PIB per capita (n-1)	3333	22,58	10,23	-33,14	16,40	21,40	26,89	86,34
RNB, crescimento per capita (% anual) (n-1)	2593	2,14	5,69	-47,14	-0,10	2,41	4,88	101,41
RNB,crescimento (% anual) (n-1)	2593	3,60	5,70	-50,11	1,36	3,90	6,37	99,81
Utilização de crédito do FMI (n-1)	2327	471.706.539,75	1.918.467.795,00	0,00	0,00	54.307.000,00	212.437.000,00	28.317.356.000,00

### **Medidas descritivas das variáveis sociodemográficas**

Os atributos de caracterização sociodemográfica mostram uma dispersão menos acentuada que os anteriores. Verificam-se coeficientes de variação superiores a 100% apenas nos atributos linhas telefónicas, sub-5 taxa de mortalidade, taxa de mortalidade pós-neonatal, densidade populacional, população total, rácio de mortalidade materna e utilizadores de internet, sendo mais acentuados os valores nos quatro últimos.

As três variáveis crescimento populacional, densidade populacional e população total concentram os valores mais elevados, e com registos acentuadamente altos, em apenas 25% da amostra. Na despesa total em saúde verifica-se que 25% dos casos apontam para um gasto máximo de 4,41% do PIB e 75% registam gastos em saúde num máximo de 7,54% do PIB. Constata-se um valor máximo de gastos em saúde de 19,31% do PIB e há casos em que o gasto é nulo.

Em referência à escolarização denotam-se diferenças acrescidas conforme aumenta o grau de ensino. Quanto ao ensino primário 25% dos casos têm, pelo menos, 95,02% de matrículas activas em referência à população de idade a que oficialmente corresponde esse nível de formação, 50% dos casos têm 101,50%, até a um máximo de 158,75%, o que demonstra que a grande maioria das crianças pertencentes aos países em estudo tem acesso ao ensino primário e que em alguns há indivíduos com idade mais avançada que estão a frequentar este nível de ensino. No ensino secundário os valores já não são tão generalizantes, mas pode-se apurar que 50% dos casos, tem, pelo menos, 81,36% da população que corresponde a esse nível de formação com matrícula activa. Quanto ao ensino superior (terciário), existem observações em que não há qualquer matrícula neste nível de ensino. Em 25% dos casos existem, no máximo, apenas 6,48% de matrículas activas em relação à população que corresponde à possível frequência de ensino universitário, em metade dos casos essa percentagem é 22,02%, havendo no último quarto da amostra um aumento de 43,90% para 118,10%, podendo-se concluir que este grau de formação se encontra mais generalizado em apenas 25% dos registos em análise.

As linhas telefónicas e os utilizadores de internet são indicadores que marcam o *status* de desenvolvimento e o nível de globalização de um país, e daí fazer algum sentido que as suas medidas descritivas tenham características aproximadas. Denotam uma grande disparidade entre o seu valor mínimo e máximo, bem como um desvio padrão superior ao seu valor médio. Em metade da amostra apenas 9,66 em cada 100 pessoas tem acesso a linhas telefónicas e 2,21 em cada 100 pessoas têm acesso à internet. Dos 75% da amostra até à sua totalidade existe uma variação de 27,75 para 74,46 pessoas em cada 100 com acesso a linhas telefónicas, enquanto na utilização de internet verifica-se um diferencial de 14,19 para 92,18 pessoas em cada 100.

Os valores referentes às taxas de mortalidade e principalmente ao rácio de mortalidade materna são assinalados com números altos de desvio padrão, o que determina a existência de fortes

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

disparidades no âmbito da amostra. Os índices mais elevados encontram-se no último quartil, consistindo numa diferença substancial relativamente aos outros 75% da amostra.

Quanto aos direitos dos trabalhadores constata-se que a maioria das observações remete para uma restrição dos direitos, de forma total na percentagem de 30,4% e parcialmente restritos com 44,8%. Apenas 24,8% da amostra refere uma total protecção dos direitos dos trabalhadores. Relativamente aos direitos económicos da mulher verifica-se que 65% das observações da amostra indica a nula efectividade de direitos económicos das mulheres. Em 31% há alguns direitos na lei e na prática e só em 4% existem na realidade todos os direitos na lei e na prática.

Atributos explicativos	Observ.	Média	Desv. Padrão	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
Crescimento populacional (% anual) (n-1)	3456	1,57	1,56	-7,53	0,54	1,55	2,50	18,59
Densidade populacional (pessoas por km <sup>2</sup> de área de terra) (n-1)	3437	150,33	469,54	1,65	23,30	60,97	126,34	7.125,14
Despesa total em saúde (% do PIB) (n-1)	2553	6,12	2,38	0,00	4,41	5,84	7,54	19,31
Escolarização, ensino primário (% bruto) (n-1)	2807	98,59	19,22	19,32	95,02	101,50	108,99	158,75
Escolarização, ensino secundário (% bruto) (n-1)	2457	71,14	32,87	4,47	43,82	81,36	95,53	162,35
Escolarização, ensino terciário (% bruto) (n-1)	2165	27,65	23,51	0,00	6,48	22,02	43,90	118,10
Expectativa de vida ao nascer, total (anos) (n-1)	3367	65,97	10,81	26,82	57,75	69,36	74,20	82,93
Fragmentação étnica (n-1)	3380	0,45	0,26	0,00	0,20	0,48	0,67	0,93
Fragmentação linguística (n-1)	3277	0,40	0,29	0,00	0,12	0,38	0,64	0,92
Média de Anos de Educação (Homens) (n-1)	3240	7,20	3,23	0,60	4,60	7,10	10,00	14,20
Média de Anos Educação (Mulheres) (n-1)	3240	6,12	3,81	0,10	2,60	5,90	9,60	14,20
Linhas telefónicas (por 100 pessoas) (n-1)	3418	17,04	18,56	0,01	1,35	9,66	27,75	74,46
População 15-64 anos (% do total) (n-1)	3359	60,60	6,85	44,79	54,10	61,25	66,73	84,99
População com 65 anos e acima (% do total) (n-1)	3359	6,82	4,74	0,45	3,21	4,51	10,77	22,05
População de 0-14 (% do total) (n-1)	3359	32,58	10,65	13,44	22,19	33,86	42,48	53,03
População urbana (% do total) (n-1)	3459	52,59	23,51	5,34	32,40	53,50	71,40	100,00
População, total (n-1)	3457	34.488.311,63	126.852.096,71	40.130,00	2.480.000,00	7.649.630,75	22.377.000,00	1.331.380.000,00
Rácio de Mortalidade Materna (por 100.000 nascidos vivos) (n-1)	3192	274,48	366,07	3,90	26,15	74,95	474,35	2110,20
Sub-5 Taxa de Mortalidade (por 1.000 nascidos vivos) (n-1)	3440	56,90	57,82	2,66	12,13	31,57	91,91	301,46
Taxa de Mortalidade Neonatal (por 1.000 nascimentos) (n-1)	3440	20,25	15,30	1,37	7,13	15,75	31,38	70,81
Taxa de mortalidade pós-neonatal (por 1.000 nascimentos) (n-1)	3440	18,87	19,04	0,56	3,58	11,04	29,12	96,81
Taxa de participação laboral, mulheres (% da população do sexo feminino 15+) (n-1)	3159	51,05	17,03	8,90	40,00	51,50	62,20	90,70
Utilizadores de Internet (por cada 100 pessoas) (n-1)	2824	12,06	20,00	0,00	0,16	2,21	14,19	92,18

**Data Mining no Contra-terrorismo**  
 Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Atributos explicativos		Observ.	%
Direitos dos Trabalhadores (n-1)	Extremamente restritos	938	30,4%
	Parcialmente restritos	1379	44,8%
	Totalmente protegidos	764	24,8%
	<b>Total</b>	<b>3081</b>	<b>100,0%</b>
Direitos Económicos da Mulher (n-1)	Nenhum	242	7,9%
	Há alguns na lei, mas nenhuns na prática	1743	57,1%
	Há alguns na lei e na prática	948	31,0%
	Todos na lei e na prática	122	4,0%
	<b>Total</b>	<b>3055</b>	<b>100,0%</b>

**Medidas descritivas das variáveis políticas**

Na dimensão política a maior parte das variáveis estão delimitadas em intervalos definidos, verificando-se na sua maioria inferiores níveis de dispersão comparativamente com os registados nas outras variáveis acima. A autocracia institucionalizada varia entre de 0 (baixo) e 10 (elevado) tendo para metade dos casos um nível máximo de 0 e para 75% dos casos um nível máximo de 5. Os últimos 25% da amostra tiveram como valor máximo 10, o que indica que os níveis mais altos de autocracia institucionalizada se concentram em apenas um quarto do conjunto total de observações. A democracia institucionalizada, em escala idêntica, de forma análoga indica que 25% da amostra tem registo de valor máximo 0, metade tem registo de máximo 6 e 75% das observações têm como máximo o valor 9. Os restantes 25% da amostra têm 10 como valor máximo. Assim, pelo menos metade da amostra apresenta valores de nível acima de metade da escala de medição da institucionalização da democracia.

Os valores do índice de direitos de integridade física, numa escala de 0 a 8, indicam que em 25% da amostra o nível máximo era de 3, enquanto que em 50% o valor máximo é de 5, um ponto acima de metade da escala. Em 75% da amostra registou o valor máximo de 7, perto do total respeito dos direitos, ficando no quartil seguinte o registo máximo da escala. No que se refere aos direitos políticos, avaliado numa escala de 1 (mais livre) a 7 (menos livre), metade das observações têm como valor máximo o valor 3, indicando elevada a média liberdade no exercício dos direitos políticos. Avaliando o resto da amostra, nos dois quartis seguintes registam-se os máximos de 6 e 7, pelo que se infere que nestes 50% de observações existem níveis de liberdade limitados no exercício dos direitos políticos.

A durabilidade do regime tem um desvio padrão em relação à média que se traduz num coeficiente de variação de 130%, muito elevado, não sendo por isso a média uma orientação como medida de distribuição das observações. Esta variável tem como valor mínimo 0, o que significa que houve regimes que nem 1 ano duraram. Em 25% da amostra os regimes duraram no máximo 4 anos, e em 50% os regimes duraram no máximo 12 anos. Em 75% das observações a duração máxima foi de 33 anos e na totalidade do conjunto atingiu os 200 anos.

## Data Mining no Contra-terrorismo

### Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

O fraccionamento político total indica que em 25% da amostra a probabilidade máxima de dois deputados escolhidos aleatoriamente serem de dois partidos diferentes é de 0,43 (de 0 a 1), em 50% a probabilidade máxima é de 0,60, em 75% é de 0,75 e na totalidade da amostra existe como probabilidade máxima o valor 1, significando que existe muita diversidade de partidos políticos na assembleia. As restrições executivas variam numa escala de 1 a 7, de total autoridade e liberdade para sem autoridade ou subordinação total. Em 25% do total de observações verifica-se o escalão 3 como nível máximo registado, correspondendo a uma leve a moderada limitação à autoridade executiva. Em metade da amostra já se regista como máximo o nível 5, ou seja, limitações substanciais à autoridade executiva. No terceiro quartil e no registo da totalidade da amostra efectiva-se como valor máximo o último escalão de 7, correspondendo à existência de elevado grau de subordinação dos governos e perda de autoridade em cerca de metade da amostra.

Em 50% das observações o valor máximo registado relativamente à liberdade de assembleia e associação é de 1, com liberdade medianamente limitada. Em 75% da amostra já se atinge o valor máximo de liberdade, 2. Relativamente às liberdades civis, para uma escala de 1 (mais livre) a 7 (menos livre), em 50% das observações já se atinge o escalão 5 como valor máximo, e em 75% o valor 7 e mais alto, indicando que cerca de 50% da amostra regista baixos valores de liberdades civis.

A variável autodeterminação eleitoral indica que em 54,5% da amostra não existe liberdade no exercício do direito de voto, estando nesse valor incluídos 28,6% em que o direito existe na lei, mas não na prática. No que se refere à democracia, numa escala de sim/não 54% das observações registam regimes democráticos. Em 3,5% da amostra não estão definidos na lei direitos à participação política das mulheres. Na restante amostra esses direitos estão consagrados na lei, contudo em 15,1% não existem na prática, em 72,4% apenas existem parcialmente na prática e apenas em 8,9% existe o exercício completo dos direitos políticos das mulheres.

A liberdade de expressão é um direito plenamente gozado em apenas 24,8% da amostra, parcialmente em 45,7% e de forma alguma exercido em 29,5% das observações, o que confirma a escassa liberdade de expressão na maioria da amostra. Em 55,3% do total de observações ainda é praticada a prisão por motivos políticos e em 44,7% a tortura é praticada frequentemente, o que segue em linha com a percentagem de regimes não democráticos. Em referência ao indicador composto *status* de liberdade, verifica-se que 40,2% da amostra está abrangida pelo exercício de liberdade, sendo parcial em 32,9% e nula em 26,9% das observações.

**Data Mining no Contra-terrorismo**  
 Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Atributos explicativos		Observ.	%
Autodeterminação Eleitoral (n-1)	Não existe na lei, nem na prática	798	25,9%
	Existe na lei, mas não na prática	880	28,6%
	Existe na lei e na prática	1402	45,5%
	<b>Total</b>	<b>3080</b>	<b>100,0%</b>
Competitividade da Participação Política (n-1)	Não aplicável, políticas não reguladas	56	1,9%
	Reprimida	377	13,1%
	Suprimida (alguma, mas sob pressões)	412	14,3%
	Faccionária (interesses individuais/facções)	673	23,3%
	Transitório (categoria intermédia)	738	25,6%
	Competitiva	629	21,8%
<b>Total</b>	<b>2885</b>	<b>100,0%</b>	
Democracia (n-1)	Não	1475	46,0%
	Sim	1733	54,0%
	<b>Total</b>	<b>3208</b>	<b>100,0%</b>
Direitos Políticos da Mulher (n-1)	Não estão na lei	108	3,5%
	Estão na lei, mas não na prática	464	15,1%
	Estão na lei, mas parcialmente na prática	2224	72,4%
	Estão na lei e na prática	274	8,9%
	<b>Total</b>	<b>3070</b>	<b>100,0%</b>
Instituições em Regime (n-1)	Democracia parlamentar	796	24,8%
	Democracia semi presidencial	379	11,8%
	Democracia presidencial	558	17,4%
	Ditadura civil	811	25,3%
	Ditadura militar	450	14,0%
	Ditadura real (monárquica)	214	6,7%
<b>Total</b>	<b>3208</b>	<b>100,0%</b>	
Liberdade de expressão (n-1)	Completa	765	24,8%
	Alguma	1408	45,7%
	Nenhuma	908	29,5%
	<b>Total</b>	<b>3081</b>	<b>100,0%</b>
Prisão por Motivos Políticos (n-1)	Sim, muitas	888	28,8%
	Sim, poucas	814	26,4%
	Nenhuma	1377	44,7%
	<b>Total</b>	<b>3079</b>	<b>100,0%</b>
Regulamentação de Participação Política (n-1)	Não regulada (sem grupos políticos definidos)	21	0,7%
	Grupos políticos estáveis que competem	976	33,8%
	Sectária (postura intransigente entre grupos políticos)	621	21,5%
	Restrita (permitida mas de fraco sectarismo)	638	22,1%
	Regulada	629	21,8%
<b>Total</b>	<b>2885</b>	<b>100,0%</b>	
Status de Liberdade (n-1)	Liberdade	1360	40,2%
	Liberdade parcial	1113	32,9%
	Não há liberdade	910	26,9%
	<b>Total</b>	<b>3383</b>	<b>100,0%</b>
Tortura (n-1)	Praticada frequentemente	1378	44,7%
	Praticada ocasionalmente	1269	41,2%
	Nunca ocorreu	434	14,1%
	<b>Total</b>	<b>3081</b>	<b>100,0%</b>

**Data Mining no Contra-terrorismo**  
 Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

<b>Atributos explicativos</b>	<b>Observ.</b>	<b>Média</b>	<b>Desv. Padrão</b>	<b>Mínimo</b>	<b>Percentil 25</b>	<b>Mediana</b>	<b>Percentil 75</b>	<b>Máximo</b>
Autocracia institucionalizada (n-1)	2885	2,27	3,07	0,00	0,00	0,00	5,00	10,00
Nível de democracia (0 a 10) (n-1)	3383	6,09	3,23	0,00	2,92	6,90	9,17	10,00
Democracia institucionalizada (0 a 10) (n-1)	2885	5,16	3,97	0,00	0,00	6,00	9,00	10,00
Índice de Direitos de Integridade física (n-1)	3072	4,82	2,29	0,00	3,00	5,00	7,00	8,00
Direitos Políticos (1 a 7) (n-1)	3383	3,60	2,19	1,00	1,00	3,00	6,00	7,00
Durabilidade de regime (n-1)	3041	23,47	30,41	0,00	4,00	12,00	33,00	200,00
Fraccionamento político total (n-1)	2861	0,55	0,27	0,00	0,43	0,60	0,75	1,00
Liberdade de assembleia e associação (n-1)	3081	1,16	0,82	0,00	0,00	1,00	2,00	2,00
Liberdades Cívicas (1 a 7) (n-1)	3383	3,60	1,83	1,00	2,00	4,00	5,00	7,00
Restrições Executivas (Às regras de decisão) (1 a 7) (n-1)	2885	4,71	2,15	1,00	3,00	5,00	7,00	7,00

**Medidas descritivas das variáveis religiosas**

Na avaliação relativa à fragmentação religiosa verifica-se que em 25% da amostra em cada duas pessoas escolhidas aleatoriamente existe a probabilidade máxima de 0,23 (de 0 a 1) de elas pertencerem a religiões diferentes. Em 50% das observações a probabilidade tem um máximo de 0,47 e em 75% da amostra há 0,65 de probabilidade máxima de elas pertencerem a religiões diferentes. Na totalidade da amostra regista-se um máximo de 0,86 de probabilidade, correspondendo ao mais alto índice de fragmentação da amostra.

Nas três variáveis referentes ao tipo de religião registam-se desvios padrão muito elevados, o que significa uma elevada dispersão nos dados. Todas têm como valor mínimo zero e de máximo praticamente cem. Nas outras religiões existe uma dispersão menor, atingindo em 25% da amostra o máximo de 8,9% de pessoas de religião diferente da católica e muçulmana em relação à população total. Em metade da amostra esse valor máximo sobe para 45%, em 75% das observações tem como máximo os 80,9% e na totalidade da amostra há observações que chegam a uma percentagem de 100%, indicando esse número a inexistência nessa população de elementos pertencentes à religião católica ou muçulmana. Na religião católica em metade da amostra há um registo máximo de 14% de população católica, aos 75% das observações atinge-se o máximo de 59,5%, até ao limite de 100% registado no âmbito da análise a toda a amostra. A religião muçulmana apresenta uma alta concentração de percentagem de muçulmanos em 25% das observações, já que em metade da amostra se verificou no máximo 1,5% de população muçulmana em relação à população total, em 75% do total das observações no máximo 39,4% da população era muçulmana e na avaliação do último quarto da amostra essa percentagem sobe para um máximo de 99,9% da população.

**Data Mining no Contra-terrorismo**  
 Uma Abordagem para a Compreensão dos Factores Determinantes do Terrorismo

Atributos explicativos	Observ.	Média	Desv. Padrão	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
Fragmentação religiosa (n-1)	3417	0,44	0,23	0,00	0,23	0,47	0,65	0,86
Religião Católica (percentagem da população total) (n-1)	3417	30,60	35,17	0,00	0,70	14,00	59,50	99,10
Religião Muçulmana (percentagem da população total) (n-1)	3417	24,15	36,00	0,00	0,00	1,50	39,40	99,90
Outras religiões (percentagem da população total) (n-1)	3377	45,27	35,45	0,00	8,90	45,00	80,90	100,00

**Medidas descritivas de outras variáveis de caracterização**

A análise descritiva a estas variáveis permite observar uma discrepância muito grande entre o valor mínimo (260km<sup>2</sup>) e o máximo (17.098.240km<sup>2</sup>), apresentando um coeficiente de variação entre 264% e 266%. Metade da amostra tem como valor máximo a área de 155.395km<sup>2</sup> para a área geográfica total e 149.355km<sup>2</sup> para a área terrestre. Em 75% das observações o máximo das áreas registadas atinge os 581.050km<sup>2</sup> e 567.935km<sup>2</sup> respectivamente. Nos últimos 25%, que totalizam o conjunto da amostra, constata-se um número máximo registado muito mais elevado em relação aos outros 75%, de 17.098.240km<sup>2</sup> e 16.389.950km<sup>2</sup>.

A variável despesas militares mostra que em apenas 25% das observações se verificaram os mais altos índices de gastos. Todas as observações apresentam gastos militares, com um valor mínimo de 0,05% do PIB. Metade da amostra tem como valor máximo 1,88% e 75% das observações o valor máximo de 3,11%. Dos 75% da amostra para a sua totalidade regista-se uma subida muito grande relativamente ao valor máximo de despesas militares, que atinge os 117,39% do PIB, número que remete para a existência de países em guerra.

O índice de globalização (KOF) apresenta uma distribuição mais normalizada, com um coeficiente de variação a rondar os 35%. Partindo de um valor mínimo registado de 12 (numa escala de 0 - mínimo, a 100 - máximo, em que valores mais altos significam mais globalização), 25% da amostra atinge um máximo de 38 pontos, metade das observações chega ao valor de 49 pontos e 75% tem como valor máximo registado 62 pontos. A diferença pontual para o máximo verificado nos restantes 25% atinge os 31 pontos.

Atributos explicativos	Observ.	Média	Desv. Padrão	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
Área de superfície (km <sup>2</sup> ) - área total geográfica do país (n-1)	3440	758.963,53	2.020.180,69	260,00	34.990,00	155.395,00	581.050,00	17.098.240,00
Área terrestre (km <sup>2</sup> ) (n-1) (excluindo área sob corpos de água)	3440	733.992,78	1.937.734,81	260,00	31.615,00	149.355,00	567.935,00	16.389.950,00
Despesas militares (em % do PIB) (n-1)	2678	2,61	3,51	0,05	1,21	1,88	3,11	117,39
KOF - Índice de Globalização (0 a 100) (n-1)	3340	51,42	17,99	12,00	38,00	49,00	62,00	93,00