



**SOCIEDADE
CRISE E RECONFIGURAÇÕES**

VII CONGRESSO PORTUGUÊS DE SOCIOLOGIA

19 a 22 Junho 2012

Universidade do Porto - Faculdade de Letras - Faculdade de Psicologia e Ciências da Educação

ÁREA TEMÁTICA: Teorias e Metodologias

MAPA DE ALERTA: INSTRUMENTO DE APOIO À QUALIDADE DO CENSOS 2011

REIS, Elizabeth
Doutora em Social Statistics,
Instituto Universitário de Lisboa (ISCTE-IUL)
elizabeth.reis@iscte.pt

VICENTE, Paula
Doutora em Métodos Quantitativos,
Instituto Universitário de Lisboa (ISCTE-IUL)
paula.vicente@iscte.pt

ROSA, Álvaro
Doutor em Gestão,
Instituto Universitário de Lisboa (ISCTE-IUL)
alvaro.rosa@iscte.pt

Resumo

O Recenseamento da População e Habitação é realizado em Portugal de 10 em 10 anos pelo Instituto Nacional de Estatística. Esta operação envolve múltiplos processos com múltiplas fontes de erro capazes de afectar a qualidade da informação produzida. No Censos 2011, o Instituto Nacional de Estatística implementou, no âmbito da sua estratégia de qualidade, mecanismos específicos de garantia de qualidade para detectar e corrigir atempadamente falhas humanas ou técnicas. Um desses mecanismos foi o Mapa de Alerta, uma mapa onde cada freguesia é tipificada com uma cor – verde, laranja ou vermelho – conforme o grau de risco – baixo, médio ou elevado - de que os níveis de qualidade desejados para o trabalho de campo possam não ser atingidos. Nesta comunicação descreve-se a construção do Mapa de Alerta e discute-se o seu contributo na garantia de qualidade do trabalho de campo no Censos 2011.

Abstract

The Portuguese Population and Housing Census is carried out every ten years by Statistics Portugal. This operation is a multi-stage process with multiple sources of error that could directly affect the quality of Census output. In the Census 2011, the Statistics Portugal had decided, under the quality assurance strategy, to create specific mechanisms in order to diminish the uncertainties and to cope with possible human and technical failures such that the operation might undergo adequately and smoothly. One of those mechanisms was the Map of Alert, a three colored map – green, orange and red – of all civil parishes of Portugal whose objective was to provide information on the potential risk – reduced, medium and high – of non-conformity of quality standards defined for the fieldwork process of the Census operation. In this paper we describe how the Map of Alert was built and discuss its contribution for the quality assurance in the fieldwork operation of the Portuguese Census 2011.

Palavras-chave: Censos; Qualidade; Mapa de Alerta; Análise de Clusters
Keywords: Census; Quality Assurance; Map of Alert; Cluster Analysis

[PAP0906]

1. Introdução

Os recenseamentos da habitação e da população (censos) constituem a maior e mais dispendiosa operação estatística que um país pode levar a cabo e têm como principal objectivo fornecer informação estatística de elevada qualidade que responda às necessidades dos seus utilizadores. Os seus resultados permitem obter um retrato no tempo e no espaço sobre a estrutura e as características dos alojamentos e da população, importante não só para um conhecimento aprofundado sobre a realidade do país (relevante ao nível da construção, implementação e avaliação de estruturas, equipamentos e políticas públicas), mas também enquanto fonte de informação privilegiada para diferentes tipos de utilizadores (administração central, autoridades locais, investigadores, jornalistas, público em geral...).

No quadro da produção das estatísticas oficiais, indicadores de precisão (erros, desvios ou variâncias) não constituem a única dimensão a ter em conta, é preciso a leitura de outros indicadores que permitam aferir sobre a qualidade dos procedimentos associados às diferentes fases de uma operação estatística. Neste âmbito, a perspectiva de controlo de processos permite desenvolver e implementar uma lógica de monitorização *on-going*, com base em métodos, ferramentas e indicadores de aferição da qualidade (INE, 2007).

A perspectiva de controlo de processos tem como premissa que dados fiáveis resultam de processos precisos e adequados (INE, 2010). Nas operações censitárias existe a possibilidade de erro a vários níveis e em diferentes etapas do processo, tornando-se, assim, importante desenvolver e implementar mecanismos e indicadores de qualidade associados a processos-chave, e seus sub-processos, passíveis de monitorização ao longo do processo de produção. Deste modo, o desenvolvimento de um programa de controlo e avaliação da qualidade constitui uma etapa importante e indispensável na preparação de uma operação censitária. Para os Censos 2011 o Instituto Nacional de Estatística (INE) implementou mecanismos que permitiram detectar falhas e aspectos que pudessem ser melhorados no decorrer da própria operação, bem como quantificar os diferentes tipos de erros que inevitavelmente têm lugar num recenseamento. Com esse objectivo foi criado um Sistema de Indicadores de Alerta que possibilitou obter informação sobre potenciais desvios da recolha e averiguar a sua provável causa. Foi, assim, possível monitorar em tempo real o processo de distribuição e recolha dos questionários e desencadear acções de ajustamento ou correcção com vista a garantir os princípios de qualidade que balizaram a realização desta operação censitária.

2. Sistema de Indicadores de Alerta

O Sistema de Indicadores de Alerta (SIA) desenvolvido incluiu duas componentes, operacionalizadas ao nível das freguesias:

- Mapa de Alerta: mapa onde as freguesias surgem tipificadas segundo três níveis de alerta correspondentes a diferentes graus de risco potencial de cada freguesia não garantir elevada qualidade do produto final. O nível de risco é identificado pela análise de um conjunto de indicadores relativos às freguesias. Aos três níveis de alerta corresponderam graus diferentes de acompanhamento da operação por parte do coordenador concelhio e dos coordenadores de freguesia.
- Intervalos de Alerta: intervalos de valores estimados para o total de alojamentos e de indivíduos em cada freguesia.

Nesta comunicação apresenta-se a metodologia de construção do Mapa de Alerta, o próprio mapa e os resultados que revelam em que medida se conseguiram assegurar os níveis de qualidade da operação censitária.

3. Metodologia

Para a construção do Mapa de Alerta foi necessário recolher dados que permitissem caracterizar alguns aspectos chave das freguesias e proceder depois à sua análise tendo por objectivo mapear as freguesias de Portugal em função dos níveis de risco de incumprimento na obtenção dos níveis de qualidade desejados.

3.1 Recolha dos dados

A recolha dos dados fez-se a partir de um inquérito aos Presidentes das Juntas de Freguesia (PJF), envolvendo um questionário estruturado e auto administrado, através do qual se recolheram as percepções dos Presidentes das Juntas sobre a população, as habitações e áreas/zonas da sua própria Freguesia. O questionário solicitava ainda informação sobre a experiência e habilitações do Presidente da Junta, bem como as competências e disponibilidade dos potenciais candidatos a recenseadores.

O questionário estava estruturado em três partes, para além da identificação do município e da freguesia. A primeira secção era facultativa e solicitava informação sobre a experiência e as competências pessoais dos PJF: idade, habilitações literárias, número de anos como presidente de junta, frequência de utilização do computador e da internet. A segunda secção era relativas às percepções do PJF sobre a existência, na freguesia de: população idosa, analfabeta, residente em bairros sociais, emigrante e imigrante, ou sem abrigo, sobre as habitações (em condomínio fechado) e sobre a existência nas áreas/zonas da freguesia de casas de segunda habitação ou veraneio, de construção recente, de difícil acesso, com habitações dispersas, ou de tipo dormitório. Na terceira secção questionava-se sobre o grau previsto de dificuldade/facilidade de execução e de recrutamento de recursos humanos/colaboradores para a operação Censos 2011.

Os questionários foram enviados por correio a todos os PJF (4260) no início de Outubro de 2010, tendo sido enviado um segundo *mailing* àqueles de quem não se tinha recebido resposta, em meados de Dezembro. A recolha deu-se por terminada a 27 de Janeiro de 2011.

Obtiveram-se 3855 respostas correspondendo a uma taxa de resposta de 90,5%.

3.2 Análise dos dados

Inicialmente, dado o elevado número de variáveis recolhidas com o inquérito aos PJF, foi aplicada uma Análise em Componentes Principais com o objectivo de reduzir a informação disponível e uniformizar as escalas de medida das variáveis originais. Esta análise foi realizada separadamente para os subconjuntos de variáveis relativas às características dos PJF e para as suas percepções sobre as características das respectivas Freguesias, por se tratarem de construtos teóricos distintos.

Em seguida, a partir das dimensões encontradas nas análises anteriores foram aplicados Modelos de Mistura Finita para identificação do número de grupos de freguesias (k) e dos seus centróides, resultados utilizados posteriormente numa Análise Não-hierárquica de Agrupamento/Clusters (k -means) para identificação das freguesias a incluir em cada grupo.

A validação dos clusters foi realizada através de uma Análise Discriminante com validação cruzada, em que cada freguesia é classificada com base em funções discriminantes estimadas a partir das restantes observações.

O passo seguinte consistiu na caracterização dos grupos de freguesias com base nas dimensões de agrupamento e a atribuição de um nível de risco potencial a cada freguesia.

Por fim, desenhou-se o Mapa de Alerta com a tipificação das freguesias segundo os níveis de risco: Reduzido (coloração verde), Médio (coloração laranja) e Elevado (coloração vermelha).

4. Resultados

4.1 As dimensões de risco potencial

Para estimar as dimensões do risco potencial das freguesias não cumprirem os objectivos previstos para a operação censitária, foram aplicadas duas análises em componentes principais aos dados do inquérito aos PJF. Na primeira análise, extraíram-se duas componentes principais a partir das cinco variáveis originais relativas às características dos PJF, e que foram designadas por “Competências do PJF” e “Experiência do PJF”. As duas componentes explicam 76,3% da variância inicial e evidenciaram uma razoável consistência

interna (Alfa de Cronbach de 0,85 e 0,64, respectivamente). Os dados revelaram-se razoáveis para a aplicação deste tipo de análise (KMO = 0,68, teste de esfericidade de Bartlett com $p\text{-value} < 0,001$).

Na segunda análise, realizada sobre catorze variáveis iniciais relativas às percepções dos PJF em relação às características da freguesia que poderiam dificultar a operação Censos 2011, foram extraídas cinco componentes principais designadas por “População de difícil contacto”, “Competências e disponibilidade dos recenseadores”, “População envelhecida”, “Zonas (interior) desertificadas” e “Zonas com população ausente (recentes ou de 2ª habitação)”. As cinco componentes explicam 61,3% da variância inicial. As quatro primeiras dimensões apresentam níveis de consistência interna médios a bons (Alfa de Cronbach de 0,60 a 0,86). Apenas a quinta componente teve um valor de Alfa de Cronbach abaixo do mínimo aceitável. Os dados revelaram-se razoáveis para a aplicação deste tipo de análise (KMO = 0,696, teste de esfericidade de Bartlett com $p\text{-value} < 0,001$).

4.2 Agrupamento das freguesias

O agrupamento das freguesias foi realizado através de uma análise não hierárquica de agrupamento (*k-means*) com pré-definição do número de grupos a formar. As variáveis de agrupamento foram as sete componentes principais obtidas das análises em componentes principais prévias. O número de grupos foi determinado pela aplicação de modelos de mistura finita que apontaram para a existência de 7 grupos. Como centros iniciais na análise de clusters usaram-se os centróides dos grupos resultantes da aplicação dos modelos de mistura finita.

O Quadro 1 mostra o perfil médio de cada grupo para as sete componentes principais, assim como o peso de cada grupo em termos do número de freguesias e da população total recenseada em 2001.

Ao grupo 1 foi associado o nível máximo de alerta devido ao valor médio máximo atingido para a população de contacto difícil. Aos grupos 3, 6 e 7 foi associado o nível de alerta laranja: o grupo 3 pelas menores competências dos PJF, o grupo 6 pela existência de maior proporção de jovens, e o grupo 7 pela previsão de falta de competência e disponibilidade por parte dos recenseadores. Aos grupos 2, 4 e 5 foi associado o nível de alerta verde.

Quadro 1 - Perfil médio e pesos dos grupos em número de freguesias e população de 2001

Componentes principais	1	2	3	4	5	6	7
Competências do PJF	0,4	0,3	-1,7	0,3	0,0	0,4	0,0
Experiência do PJF	0,0	0,7	0,0	-0,2	-0,3	-0,2	-0,1
População de contacto difícil	2,2	-0,4	-0,2	-0,1	0,0	0,2	-0,4
Competências e disponibilidade dos recenseadores	0,1	0,8	-0,1	0,5	-0,1	-0,3	-1,1
População envelhecida	0,5	-0,2	-0,1	0,5	0,4	-1,0	0,7
Zonas (interior) desertificadas	-0,5	-0,3	-0,2	-0,1	1,9	-0,1	-0,5
Zonas com população ausente	-0,6	-0,5	0,0	1,3	-0,3	0,0	-0,2
Freguesias	260	736	494	609	424	782	550
	6,7%	19,1%	12,8%	15,8%	11,0%	20,3%	14,3%
População (2001)	2924656	1425530	735823	804602	773574	2467481	575237
	30,1%	14,7%	7,6%	8,3%	8,0%	25,4%	5,9%

A aplicação de uma análise discriminante às componentes principais usando a partição de 7 grupos permitiu estimar seis funções para prever a pertença de cada freguesia a cada grupo, tendo-se obtido uma percentagem global de 92,7% de casos igualmente classificados pela análise de agrupamento *k-means* e pela análise discriminante. Todas as componentes principais têm poder para discriminar os 7 grupos de freguesias, embora as três variáveis com maior poder discriminante estejam associadas de forma negativa ao risco potencial (Quadro 2).

Quadro 2 - Poder discriminante das componentes principais

Ordem de entrada e componente	Lambda de Wilks	F	Sig.
1 ^a . Competências do PJJ	,446	796,42	0,000
2 ^a . Zonas (interior) desertificadas	,240	668,73	0,000
3 ^a . População envelhecida	,150	578,37	0,000
4 ^a . População de contacto difícil	,094	543,12	0,000
5 ^a . Zonas com população ausente	,061	518,17	0,000
6 ^a . Competências e disponibilidade dos recenseadores	,040	504,90	0,000
7 ^a . Experiência do PJJ	,033	457,01	0,000

A percentagem de casos igualmente classificados pelos dois métodos variou entre 81,5% (grupo 1) e 97,8% (grupos 3 e 6) (Quadro 3). Os resultados com diferenças mais evidentes dizem respeito aos grupos 1 e 7, mas enquanto no grupo 7 as diferenças são parcialmente absorvidas por grupos com o mesmo nível de alerta (laranja), o mesmo não ocorre no caso do grupo 1, único grupo associado a um nível de alerta vermelho. Das 260 freguesias classificadas pela análise de agrupamento no grupo 1 (alerta vermelho), 26 (10%) são classificadas pela análise discriminante em grupos de alerta laranja, enquanto que 22 (8,5%) são classificadas em grupos de nível mínimo de alerta (verde).

Quadro 3 - Classificação pelo método *k-means* e pela análise discriminante (%)

Agrupamentos <i>k-means</i>	Grupo	Classificação pela análise discriminante							Total
		1	2	3	4	5	6	7	
	1	81,5	3,8	2,3	2,7	1,9	5,8	1,9	100,0
	2		95,5	1,2		,4	2,9		100,0
	3		,6	97,8	,6		1,0		100,0
	4		3,4	2,1	88,5	1,8	3,8	,3	100,0
	5			3,8		92,2	2,6	1,4	100,0
	6		1,2	,3	,3	,4	97,8	,1	100,0
	7		3,6	3,3	2,5	,7	2,7	87,1	100,0

O Quadro 4 resume as diferenças de classificação pelos dois métodos em termos de nível de alerta e de risco potencial:

- Ligeira maior probabilidade de a análise de clusters *k-means* classificar uma freguesia no nível de alerta vermelho;
- Todas as freguesias classificadas em alerta vermelho pela análise discriminante são incluídas no mesmo nível de alerta pela análise de clusters;
- Freguesias classificadas pela análise de clusters em alerta laranja ou verde, também o são pela análise discriminante e vice-versa.

Quadro 4 - Classificação pelo método *k-means* e pela análise discriminante segundo os níveis de alerta (%)

Agrupamentos <i>k-means</i>	Classificação pela análise discriminante			Total
	Vermelho	Laranja	Verde	
Vermelho	81,5%	10,0%	8,5%	100,0%
Laranja	-	96,8%	3,2%	100,0%
Verde	-	5,7%	94,3%	100,0%

4.3 Mapa de Alerta

Os resultados para cada nível de risco potencial e de alerta estão sumariados no Quadro 5. A Figura 1 apresenta o Mapa de Alerta para o Continente e Arquipélagos de Madeira e Açores. As freguesias com coloração negra correspondem às não-respostas do inquérito aos PJJF.

Quadro 5 – Distribuição das freguesias e da população de 2001 por níveis de alerta/risco potencial

	Nível de Alerta		
	Vermelho (elevado)	Laranja (médio)	Verde (reduzido)
Nº de freguesias	260	1826	1769
% de freguesias	6,7%	47,4%	45,9%
População em 2001	2.924.656	3.778.541	3.003.706
% da população em 2001	30,1%	38,9%	30,9%

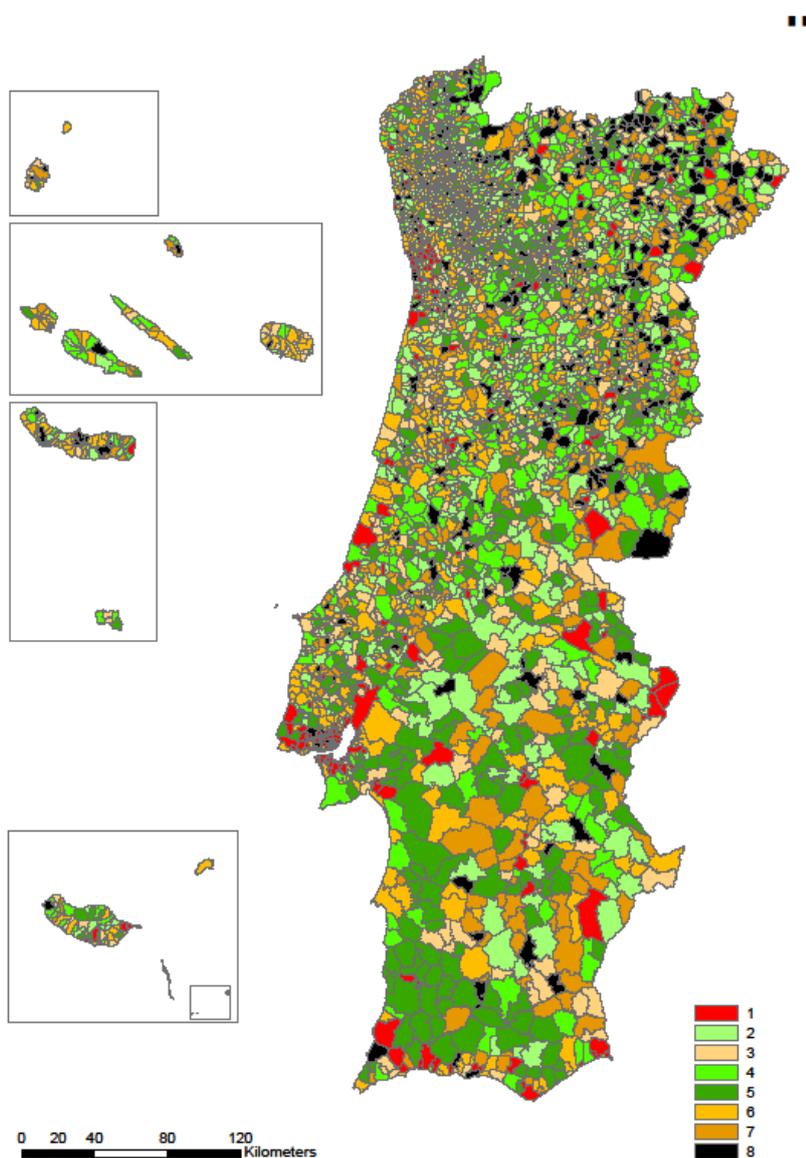


Figura 1: Mapa de alerta – Continente e Ilhas

A caracterização de cada grupo a partir dos resultados apresentados nos pontos anteriores é a seguinte:

- Grupo 1: 260 freguesias onde os presidentes da junta têm experiência média e admitem ter elevada competência; a execução da operação Censos 2011 não é considerada muito fácil devido à existência de população de contacto difícil (analfabeta, residente em bairros sociais, condomínios fechados ou tipo dormitório, estrangeiros, e zonas habitacionais recentes). Trata-se do grupo de maior risco potencial, correspondendo a 6,7% das 3855 freguesias que responderam ao Inquérito e a 30,1% da população.
- Grupo 2: grupo de 736 freguesias onde os presidentes da junta admitem ter elevada competência e experiência, consideram ser fácil a execução da operação Censos 2011 em especial por não existirem zonas de acesso ou população de contacto difíceis e ser fácil recrutar recenseadores competentes e disponíveis. É considerado um grupo de reduzido risco potencial e corresponde a 14,7% das freguesias respondentes (19,1% da população).
- Grupo 3: grupo de 494 freguesias onde os presidentes da junta admitem ter muito reduzidas competências (habilitações muito reduzidas, reduzida frequência de utilização do computador e da internet) e alguma experiência, mas ser fácil a execução da operação Censos 2011 devido sobretudo à não existência de população difícil de contactar, embora não exista população volátil. É considerado um grupo de risco potencial médio e representa 12,8% das freguesias respondentes (7,6% das freguesias).
- Grupo 4: grupo de 609 freguesias onde se destaca a elevada juventude e competência a par da reduzida experiência dos presidentes das juntas de freguesia. Trata-se de um grupo de freguesias que, na opinião dos seus presidentes das juntas de freguesia, apenas se destaca pela existência de construção habitacional recente a par da existência de população ausente (emigrantes, segundas habitações ou de veraneio). Admitem ser fácil recrutar recenseadores competentes e disponíveis, o que os leva a acreditar ser também fácil a execução da operação Censos 2011. É considerado um grupo de reduzido risco potencial e representa 15,8% das freguesias respondentes (8,3% da população).
- Grupo 5: 424 freguesias onde os presidentes das juntas são jovens mas admitem ter elevada competência mas menor experiência, ser relativamente difícil a execução da operação Censos 2011 devido à existência de zonas de interior desertificadas com difícil acesso e populações muito dispersas. É considerado um grupo de reduzido risco potencial e representa 11,0% das freguesias respondentes (8,0% da população).
- Grupo 6: 782 freguesias onde os presidentes da junta são jovens, admitem ter elevada competência mas reduzida experiência e não ser fácil a execução da operação Censos 2011. Este grupo distingue-se pelo menor envelhecimento da sua população. É considerado um grupo de médio risco potencial e representa 20,30% das freguesias respondentes (25,4% da população).
- Grupo 7: 550 freguesias onde os presidentes da junta admitem ter média competência e experiência, não ser fácil a execução da operação Censos 2011 devido à dificuldade em recrutar recenseadores com competência e disponibilidade, em conjunto com uma população envelhecida e analfabeta. É considerado um grupo de médio risco potencial e representa 14,3% das freguesias respondentes (5,9% da população).

4.4 Qualidade dos resultados finais

Terminada a operação Censos 2011 foi possível fazer uma avaliação da qualidade dos resultados finais, o que permite, por sua vez, uma avaliação da qualidade do SIA enquanto conjunto de mecanismos que permitiram monitorar em tempo real a execução da operação.

Nessa avaliação apurou-se que apenas 681 freguesias apresentaram desvios percentuais considerados “excessivos”, i.e., superiores a 10% face aos valores estimados (INE, 2011). Estes desvios foram considerados aceitáveis por serem difíceis de antecipar (por falta de informação). A sua justificação coube dentro de um conjunto de razões que incluem: a criação de novas freguesias ou a alteração dos limites de freguesias existentes, saída de população para o estrangeiro ou deslocação de população entre freguesias e perda de população por óbitos ou deslocação para lares.

5. Conclusões

As operações censitárias têm como principal objectivo recolher, produzir e fornecer informação estatística de elevada qualidade que responda às necessidades dos seus utilizadores. Na prossecução deste objectivo de qualidade é importante implementar mecanismos que permitiam detectar falhas e aspectos que possam ser melhorados no decorrer da própria operação, bem como quantificar os diferentes tipos de erros que inevitavelmente têm lugar num recenseamento.

Para aferir da qualidade dos procedimentos associados às diferentes fases da operação Censos 2011 numa lógica de monitorização em tempo real, com base em métodos, ferramentas e indicadores de aferição da qualidade foi desenvolvido um Sistema de Indicadores de Alerta, operacionalizado ao nível das freguesias através do qual foi possível obter informação sobre potenciais desvios da recolha e averiguar a sua provável causa.

O Mapa de Alerta foi uma das componentes deste sistema, construído a partir de uma análise não hierárquica de agrupamento (*k-means*) das freguesias que utilizou como variáveis de agrupamento um conjunto de indicadores relativos ao grau de risco potencial de cada freguesia não garantir elevada qualidade do produto final. Os dados foram recolhidos através da realização de um inquérito a nível nacional aos presidentes das juntas de freguesia.

Aos três níveis de alerta correspondem graus diferentes de acompanhamento da operação por parte do coordenador concelhio. O mapa identifica um conjunto de 260 freguesias a quem foi atribuído o nível de alerta vermelho pelo seu maior risco potencial. Como factor de risco identificou-se a existência, acima da média, de população de contacto difícil (analfabeta, residente em bairros sociais, condomínios fechados ou tipo dormitório, estrangeiros, e zonas habitacionais recentes). Acresce o facto de os presidentes da junta considerarem que a execução da operação Censos 2011 não seria muito fácil, embora admitam ter média experiência e elevada competência para a realização deste tipo de operação. Correspondem a 6,7% das 3855 freguesias que responderam ao Inquérito e a 30,1% da população recenseada em 2001. Foi este conjunto de freguesias que teve um maior acompanhamento da operação censitária por parte do coordenador concelhio e atenção redobrada pelos coordenadores e subcoordenadores de freguesia.

O balanço final da operação revelou bons resultados numa óptica da qualidade uma vez que apenas 681 freguesias (16% do total de freguesias) apresentaram contagens do número de indivíduos fora dos valores estimados.

Referências Bibliográficas

Instituto Nacional de Estatística (2007). *Programa de Acção para os Censos 2011*. Lisboa: Gabinete dos Censos, INE.

Instituto Nacional de Estatística (2010). *Censos 2011 Plano de Controlo e Avaliação da Qualidade-Controlo do Processo Produtivo*. Lisboa: Gabinete dos Censos, INE.

Instituto Nacional de Estatística (2011). *Avaliação do Sistema de Indicadores de Alerta Censos 2011 (SIA2011)*. Lisboa: Gabinete dos Censos, INE.