

Estudo de viabilidade para os Censos 2021

Linhas gerais do novo modelo para os
Censos 2021, a testar em 2016
(Relatório QUAR*)

*Quadro de Avaliação e Responsabilização

4 de dezembro de 2015

*Users, funders and providers of official statistics want estimates that are “wider, deeper, quicker, better, cheaper”, (...) more relevant and less burdensome*¹

Sumário executivo

Este relatório faz uma síntese sobre o grau de execução do Programa de Trabalho 2013 – 2017: Estudo de viabilidade sobre a adoção de um novo modelo censitário.

O relatório aborda os seguintes pontos:

- *No Ponto 1: Principais linhas estratégicas para um novo modelo censitário em Portugal*; refere-se a estratégia de curto prazo para os Censos 2021, e a estratégia de médio/longo prazo, que contempla a mudança de paradigma no modelo de censos, que entre outros aspetos passa pela utilização de informação administrativa.
- *No Ponto 2: Construção de uma Base de População Residente (BPR) a partir de registos administrativos*; faz-se a descrição da metodologia de construção da BPR (Base de População Residente), assim como o modelo para o carregamento, integração e tratamento dos dados a partir de diferentes fontes. São realçados os aspetos que envolvem a ligação entre as diferentes bases de dados e as dificuldades que esse exercício representa, tendo em conta a ausência de atributos numéricos comuns em todos os ficheiros. Este primeiro protótipo da BPR permite estimar cerca de 8,6 milhões de indivíduos residentes em Portugal – destes cerca de 92% estavam representados nos Censos 2011. Esta base de dados constitui um repositório único de informação estatística em Portugal, que converge para 15 variáveis censitárias. Pela sua importância, este primeiro protótipo carece de um estudo mais aprofundado no sentido de um aperfeiçoamento metodológico com vista à melhoria dos resultados.

¹ Citro, C. 2014 (Director Committee on National Statistics, U.S. National Academy of Sciences/National Research Council)

- *No Ponto 3: Comparação de microdados entre ficheiros administrativos e Censos*; é feita uma comparação entre os microdados dos Censos 2011 e dos ficheiros administrativos para um conjunto de 18 variáveis demográficas e socioeconómicas provenientes de várias fontes. As taxas de correspondência dos conteúdos são, em geral, superiores a 90% no caso das variáveis demográficas, atingindo valores superiores a 80% em algumas variáveis relacionadas com a atividade económica. Para enquadrar os resultados foi feita também a comparação entre os microdados dos Censos 2011 e do 1º trimestre do Inquérito ao Emprego 2011: os resultados das taxas de correspondência são convergentes para os do comparativo Censos – ficheiros para variáveis análogas. Em termos de potencial de substituição das variáveis dos Censos por informação dos ficheiros, não obstante os resultados sugerirem uma boa qualidade da informação administrativa, equaciona-se como fundamental o estudo dos universos e complementaridade entre fontes em termos de cobertura.
- *No Ponto 4: Desenho do modelo para os Censos 2021*; são apresentadas as quatro grandes linhas estratégicas para o novo modelo censitário para 2021 evidenciando as mudanças face aos Censos 2011.
- *No Ponto 5: Conclusões e trabalho futuro*: Sintetizam-se os pontos-chave e os resultados apresentados no relatório.

Índice

Sumário executivo	3
Introdução	7
1. Principais linhas estratégicas para a transformação dos Censos em Portugal	8
2. O contributo dos ficheiros administrativos para os Censos: Etapas para a construção de uma Base de População Residente (BPR).....	10
2.1. Enquadramento.....	10
2.2. Criação do quadro legal e cooperação com as entidades detentoras dos ficheiros.....	12
2.3. Avaliação da qualidade dos ficheiros e identificação das variáveis provenientes dos diversos ficheiros	13
2.4. Interligação de bases de dados	21
2.4.1. Interligação de dados dos Censos e dos registos administrativos..	21
2.4.2. Interligação dos ficheiros administrativos e sinalização de presença de indivíduos	29
2.5. Estimação da população residente em 2011 em função de regras de residência.....	41
2.6. Introdução de variáveis socioeconómicas na BPR 2011	43
2.7. Análise dos principais resultados	48
3. Comparação de microdados entre Ficheiros Administrativos e Inquéritos Estatísticos.....	61
3.1. Enquadramento.....	61

3.2. Comparação de microdados entre os Censos 2011 e os ficheiros administrativos	61
3.2.1. Constituição dos universos para a realização do estudo.....	61
3.3. Comparação de microdados entre os Censos 2011 e o Inquérito ao Emprego	67
2.3.1. Seleção de variáveis e metodologia	67
3.4. Síntese de resultados e avaliação da qualidade das fontes	69
4. Desenho para os Censos 2021	77
4.1. Desenho para os Censos 2021: a caminho de um censo digital	77
4.1.1. Utilização de um ficheiro de alojamentos: alteração do modelo de distribuição de questionários.....	78
4.1.2. Internet como principal canal de resposta aos Censos 2021.....	80
4.1.3. Reforço das tecnologias de informação e comunicação no trabalho de campo.....	81
4.1.4. Utilização de informação administrativa	81
4.2. Principais alterações e riscos face aos Censos 2011.....	82
5. Conclusões e trabalho futuro	84
6. Bibliografia de referência.....	87

Introdução

Em 2013, o Conselho Diretivo do INE encarregou o Gabinete para os Censos 2021 de apresentar um Estudo de Viabilidade tendo em vista uma mudança do modelo censitário utilizado em Portugal.

A mudança de paradigma deverá permitir melhorias em quatro grandes áreas: redução dos custos das operações censitárias, redução da carga estatística, permitir uma maior periodicidade na disponibilização da informação de cariz censitário e possibilitar a integração de dados de natureza censitária.

No contexto internacional, diversos países da UNECE com modelos censitários semelhantes ao português (Reino Unido, Sérvia, Canadá, Estados Unidos, Austrália, Roménia, Itália, República Checa, Luxemburgo) têm em curso estudos de viabilidade idênticos ao que se encontra em desenvolvimento em Portugal.

O objetivo de estudar a viabilidade da utilização dos ficheiros administrativos para fins censitários prende-se por um lado, com a necessidade de permitir projetar uma visão de médio e longo prazo para o desenho do modelo censitário alinhado com as tendências internacionais, e, ao mesmo tempo, definir uma estratégia de curto prazo que responda de forma mais eficiente à realização dos Censos já em 2021.

Este relatório sintetiza os resultados alcançados até ao momento. No horizonte de médio e longo prazo foram feitos progressos importantes, no que se refere aos exercícios, de construção de um primeiro protótipo de uma Base de População Residente para Portugal a partir de informação administrativa, e, de comparação de microdados entre os Censos 2011 e os ficheiros administrativos. Apresentam-se ainda as principais linhas de inovação para os Censos 2021.

Em 2016, será realizado um inquérito teste cujos resultados servirão para afinar o desenho do modelo para os Censos 2021.

1. Principais linhas estratégicas para a transformação dos Censos em Portugal

Desde 2014 Portugal tem vindo a desenvolver um estudo de viabilidade que possibilite a transformação do modelo censitário. Foram analisadas as diferentes opções seguidas a nível internacional (conforme bibliografia de referência), tendo em conta as vantagens, os riscos e as condições de mudança.

Os trabalhos desenvolvidos até ao momento permitem afirmar que a mudança para um Censo administrativo não pode ser encarada no curto prazo, isto é, para os Censos 2021.

O processo de transformação de um Censo clássico (baseado num inquérito) para um Censo administrativo (baseado em registos administrativos) não é ainda seguro, nem claro. É necessário mais tempo para o desenvolvimento de estudos para fundamentar de forma segura opções metodológicas que permitam a substituição de um inquérito por informação administrativa. A maioria dos países que conduziu uma mudança desta natureza demorou vários anos.

Em Portugal o processo de mudança é porventura mais complexo do que em outros países que já conseguiram avançar com a substituição integral ou parcial do modelo de inquérito por informação administrativa. Na UE a maioria dos países já empreendeu essa mudança. Contudo, Portugal debate-se com limitações idênticas às verificadas em países como Reino Unido, Canadá, Austrália, Nova Zelândia, EUA. Nestes países não existe um número único, não dispõem de um registo de população residente e as fontes administrativas são insuficientes para responder às necessidades de informação proporcionada pelos Censos.

As principais linhas estratégicas para a transformação dos Censos em Portugal encontram-se alinhadas com a dos países referidos e podem ser estruturadas em duas grandes áreas:

- **Estratégia de curto prazo para os Censos 2021**

A estratégia de curto prazo para os Censos 2021 passa pela realização de um inquérito mais eficiente com quatro áreas de inovação:

- Reutilização de variáveis dos Censos 2011, na área da habitação;
- Substituição da distribuição de questionários porta-a-porta pelo envio postal de uma carta com códigos de acesso para resposta pela internet – principal modo de recolha;
- Utilização de plataformas móveis para realização do trabalho de campo;
- Utilização de informação administrativa para suporte à operação.

A implementação desta estratégia exige o desenvolvimento da investigação já em curso e anos de preparação.

- **Estratégia de médio e longo prazo**

O aprofundamento dos estudos em curso no quadro da utilização da informação administrativa deve continuar. Só uma metodologia baseada em informação administrativa permitirá alimentar a possibilidade de disponibilizar com maior frequência informação de cariz censitário.

A metodologia de construção do primeiro protótipo de Base de População Residente é o primeiro passo desta estratégia. Trata-se de um exercício complexo que carece de testes e de confirmação de resultados. A progressiva afinação desta base, passará a constituir o repositório da população residente em Portugal. A ligação entre as diferentes fontes de informação permite construir uma base com a informação de caracterização sobre os indivíduos.

Os trabalhos até ao momento permitiram a constituição de uma BPR com cerca de 8,6 milhões de indivíduos com informação estatística que deverá convergir para 15 variáveis sociodemográficas.

2. O contributo dos ficheiros administrativos para os Censos: Etapas para a construção de uma Base de População Residente (BPR)

2.1. Enquadramento

A avaliação do contributo dos ficheiros administrativos para a produção de estatísticas censitárias tem representado um aspeto central do Estudo de Viabilidade.

Pela primeira vez, o INE desenvolveu um trabalho de investigação aprofundado sobre o real contributo da informação administrativa disponível tendo em vista a substituição total ou parcial de informação recolhida através do inquérito censitário ou a sua utilização em complemento dessa informação.

Esse trabalho permitiu, pela primeira vez, a construção de um protótipo de uma Base de População Residente (BPR) para o nosso país, com recurso a informação administrativa. O exercício teve como período de referência o ano de 2011, de modo a utilizar os Censos 2011 como referencial para comparação dos resultados.

Os resultados apresentados possibilitam, pela primeira vez, avaliar o efetivo potencial de utilização da informação administrativa na produção de estatísticas sociais e de cariz censitário.

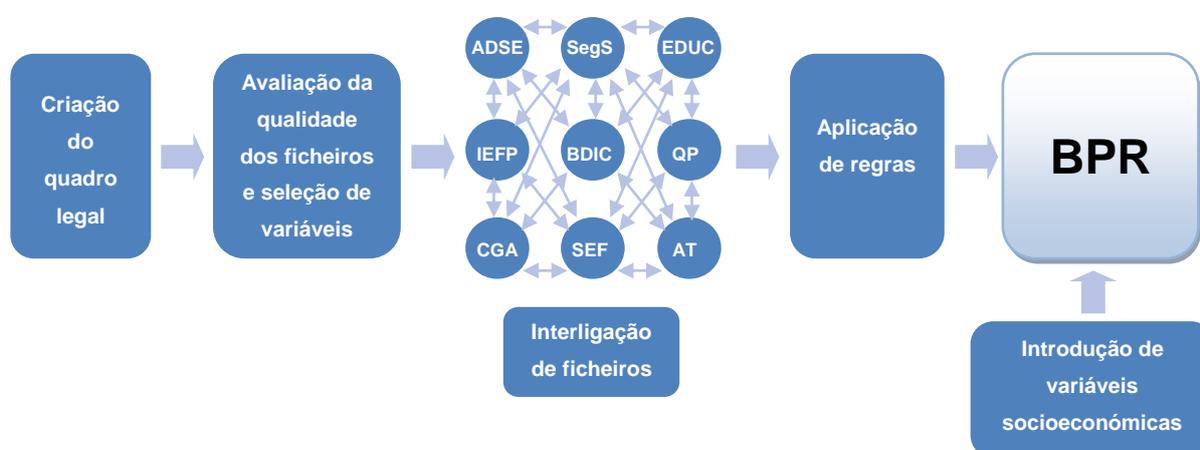
A BPR foi construída com informação proveniente de 9 ficheiros administrativos sendo a Base de Dados de Identificação Civil, BDIC, o ficheiro central na sua estrutura, uma vez que é o único ficheiro que reúne toda a população de nacionalidade portuguesa.

O desenvolvimento do primeiro protótipo da BPR, para 2011, através da interligação de diversos ficheiros administrativos, envolveu um conjunto de etapas, que se listam a seguir (e se ilustram no esquema da figura 1):

1. Criação do quadro legal e cooperação com as entidades detentoras dos ficheiros;

2. Avaliação da qualidade dos ficheiros e identificação das variáveis provenientes dos diversos ficheiros;
3. Interligação de bases de dados;
4. Estimação da população residente em 2011 em função de regras de residência;
5. Introdução de variáveis socioeconómicas.

Figura 1 – Etapas para a construção da BPR



Por se tratar de um primeiro exercício de construção de uma base de população residente para o país e pela sua complexidade, considera-se que o trabalho até agora desenvolvido tem ainda um caráter exploratório, sendo que os aspetos metodológicos estão ainda em fase de desenvolvimento e consolidação.

Por outro lado, a metodologia de atualização, a partir do ano zero, 2011, está ainda em desenvolvimento e depende da articulação com as entidades detentoras da informação. O próximo desafio será estimar a população para as freguesias do Teste de 2016. Os resultados dessa estimativa serão considerados no desenho final do modelo para os Censos 2021.

2.2. Criação do quadro legal e cooperação com as entidades detentoras dos ficheiros

A primeira etapa do projeto implicou a criação de condições legais de acesso aos registos individuais dos ficheiros administrativos e um intenso trabalho de cooperação com as 9 entidades detentoras da informação.

A Deliberação da Comissão Nacional de Proteção de Dados, CNPD, nº 929/2014, de 11 de junho, conjugada com a Lei do Sistema Estatístico Nacional, Lei nº 22/2008 de 13 de maio, conferiram a moldura legal necessária para esta fase do projeto. Sem este passo não teria sido possível o desenvolvimento do estudo, uma vez que existiam constrangimentos legais que impediam o acesso aos ficheiros administrativos.

Assim, até ao momento foram disponibilizados ao INE os ficheiros administrativos descritos no quadro 1 (foi solicitado à AT o ficheiro relativo ao IRS mas até ao momento o mesmo não foi disponibilizado).

Quadro 1 – Bases de dados disponíveis, por fonte e ano civil a que se referem os dados

Base de dados	2010	2011	2012	2013	2014
ADSE - Direção-Geral de Proteção Social aos Trabalhadores em Funções Públicas			X	X	
CGA - Caixa Geral de Aposentações	X			X	
Educação - Direção-Geral de Estatísticas da Educação e Ciência e Direção Educação da Madeira		X		X	
Quadros de Pessoal - Gabinete de Estratégia e Planeamento	X	X	X	X	
IEFP- Instituto do Emprego e Formação Profissional e Instituto de Emprego da Madeira		X	X	X	
Segurança Social - Instituto de Informática		X		X	
BDIC - Instituto dos Registos e do Notariado	X				X
SEF - Serviço de Estrangeiros e Fronteiras		X	X	X	
AT - Autoridade Tributária e Aduaneira – Cadastro					X
AT - Autoridade Tributária e Aduaneira – Imposto Municipal sobre Imóveis (IMI)					X

A Deliberação da CNPD estabelece contudo um conjunto de limitações para o acesso aos registos individuais. A ausência do endereço do individuo (apenas foi fornecido o código postal) e a ausência do nome (apenas foram fornecidas as três primeiras letras do 1º nome e as três últimas letras do último nome) são os aspetos mais críticos. Em particular, a ausência da morada inviabiliza a articulação entre a BPR e a base de alojamentos (via IMI ou FNA – Ficheiro Nacional de Alojamentos). Por sua vez, a restrição do nome dificulta a interligação entre os diferentes ficheiros.

PONTOS-CHAVE

- Dado que não há um identificador numérico único que permita a aplicação de um processo mais eficiente para interligar as diferentes bases de dados, é necessário avaliar com a CNPD a forma legal de ultrapassar as limitações decorrentes da deliberação, que não permitem aceder atualmente ao nome e à morada associados aos registos individuais.
- Tendo em vista a transição para um modelo de Censos baseado em registos administrativos e que se perspectiva, de acordo com as diretrizes internacionais, uma divulgação de base anual de variáveis de carácter censitário no pós-2021, é necessário criar os mecanismos legais de modo a formalizar as condições de cooperação com as entidades e o regular envio dos dados.

2.3. Avaliação da qualidade dos ficheiros e identificação das variáveis provenientes dos diversos ficheiros

O INE articulou com as diferentes entidades, no quadro de colaboração institucional, o conteúdo informacional necessário, as condições de tratamento e de transmissão da informação.

Para efeitos de transmissão de dados foi utilizada a aplicação de Codificação de Dados Administrativos (CDA), que permite a cada entidade realizar a encriptação,

abreviatura e transmissão de dados ao INE de forma segura e de acordo com os requisitos estabelecidos pela CNPD.

Após o carregamento dos dados provenientes das diversas fontes, cada ficheiro foi analisado separadamente. Foi produzido um relatório de carregamento onde consta o número de observações, bem como o desenho de registo, o conteúdo dos campos em termos de preenchimento, e quando necessário, a identificação das tabelas de descodificação, campos transformados e/ou criados a partir da informação original, frequentemente necessário para normalização dos registos.

No sentido de se compatibilizar a informação das diversas fontes, foi definido um conjunto de regras de uniformização de variáveis, cujo objetivo foi reduzir as diferenças nas codificações e nos formatos utilizados. Esta uniformização passou pela designação das variáveis e pela respetiva codificação, sendo que, sempre que possível, foram utilizadas as mesmas tabelas de descodificação nas variáveis disponibilizadas pelas várias fontes.

Na sequência da análise dos ficheiros, produziram-se, para cada fonte, fichas com uma caracterização dos universos, das variáveis, as respetivas percentagens de preenchimento, códigos inválidos, nulos, percentagem de ligação dos identificadores numéricos entre fontes (em anexo).

De seguida apresenta-se uma síntese dos resultados da análise das fontes administrativas e das suas potencialidades, tendo em conta a identificação das áreas para as quais existe ou não informação administrativa, associadas às respetivas limitações. Nesta fase do projeto, os processos de recolha de dados, processamento e análise foram, em grande parte, concluídos.

Consideram-se como referência as variáveis incluídas nos Censos 2011, decorrentes da regulamentação comunitária, sem prejuízo da visão mais global de âmbito nacional.

O conhecimento sobre a informação administrativa disponível que pode substituir variáveis censitárias encontra-se relativamente estabilizado e não é espetável que, a partir de informação administrativa, novas variáveis censitárias possam ser cobertas.

Nos quadros seguintes apresentam-se as variáveis provenientes de todos os ficheiros administrativos recebidos no INE com informação de cariz censitário. Para efeitos de construção da BPR, serão apenas consideradas as variáveis associadas ao indivíduo.

Quadro 2 – Variáveis obrigatórias no âmbito do regulamento comunitário, segundo a disponibilidade em ficheiros administrativos

Disponibilidade da informação em ficheiros administrativos	Número de variáveis	Variáveis
Totalmente disponível	12	Local de residência
		Sexo
		Idade
		Estado civil
		Nacionalidade
		Naturalidade
		Água
		Área útil
		Instalações sanitárias
		Época de construção
		Tipo de edifício
		Tipo de alojamento
Parcialmente disponível	6	Nível de ensino completo
		Condição perante a Atividade Económica
		Profissão
		Ramo de Atividade Económica
		Situação na profissão
		Local de trabalho
Não disponível	9	Eventual residência no estrangeiro e ano de chegada ao país
		Local de residência 1 ano antes
		Relações de parentesco
		Tipo de família
		Dimensão da família
		Forma de ocupação
		Banho ou duche
		Tipo de aquecimento
		Regime de ocupação
Total	27	

Da análise do quadro anterior destaca-se o facto de que a possibilidade de obter informação administrativa não ser uniforme, de acordo com as diferentes unidades estatísticas.

No seu conjunto, os ficheiros administrativos disponíveis permitem obter informação para 12 das 27 variáveis comunitárias e permitem uma cobertura parcial para 6 variáveis. Para as variáveis de âmbito nacional, existe uma cobertura total para 8 das 50 variáveis. Os quadros seguintes ilustram o potencial e as restrições das fontes administrativas inventariadas.

Quadro 3 – Número de variáveis (comunitárias e nacionais), de acordo com o grau de cobertura disponibilizado por informação administrativa

Unidade estatística	Censos 2011	Variáveis obrigatórias no âmbito do regulamento comunitário				Variáveis nacionais (Censos 2011)			
		Total	Cobertura total	Cobertura parcial	Sem cobertura	Total	Cobertura total	Cobertura parcial	Sem cobertura
Edifício	16	2	2	0	0	14	3	0	11
Alojamento	18	8	4	0	4	10	3	3	4
Família/Indivíduo	43	17	6	6	5	26	2	6	18
Total	77	27	12	6	9	50	8	9	33

Os quadros 4 e 5 permitem identificar o contributo dos vários ficheiros administrativos inventariados no âmbito do estudo de viabilidade para os Censos 2021. No quadro 4 incluem-se as variáveis obrigatórias por Regulamento comunitário, enquanto no quadro 5 se apresentam as restantes variáveis recolhidas nos Censos 2011.

Quadro 4 – Variáveis (comunitárias), com e sem fonte administrativa

Unidade estatística	Variáveis / Modalidades obrigatórias (União Europeia)	Cobertura Total / Parcial	Nº fontes admin	Fonte administrativa										
				SIOU (***)	IMI	IRS (****)	BDIC	SEF	EDUC	SS	CGA	QP	IEFP	ADSE
Variáveis obrigatórias, com 1 ou + fontes administrativas (assegurando diferentes níveis de cobertura do universo)														
Edifício_Alojamento	Tipo de edifício	T (*)	2	X	X									
	Época de construção	T (*)	2	X	X									
	Tipo de alojamento (familiar clássico)	T (*)	2	X	X									
	Área útil / Nº divisões (obrig 1 de 2)	T (*)	2	X	X									
	Água	T (*)	2	X	X									
	Instalações sanitárias	T (*)	1		X									
Família_Indivíduo	Sexo	T	9			X	X	X	X	X	X	X	X	X
	Idade	T	9			X	X	X	X	X	X	X	X	X
	Naturalidade (País, Município, Freguesia)	T	4			X	X	X		X				
	Nacionalidade (País)	T	6			X	X	X	X	X			X	
	Residência habitual	T (**)	5				X	X	X	X	X			
	Nível de ensino completo	P	3						X			X	X	
	Estado civil legal	T	5			X	X	X		X			X	X
	Local de trabalho (País, Município, Freguesia)	P	4							X	X	X		
	CPA-população empregada	P	3							X	X	X		
	CPA-estudantes	T	1						X					
	CPA-reformados	P	2							X	X			
	CPA-pop desemp procura 1º emprego	P	1										X	
	CPA-pop desemp procura novo emprego	P	2							X			X	
	Profissão / população empregada	P	2					X				X		
	Profissão / pop desemp proc novo emprego	P	1										X	
	Situação Prof-patrões / empregadores	P	1									X		
	Situação Prof-trab conta de outrem	P	3							X	X	X		
	Situação Prof-trab conta própria	P	1							X				
Situação Prof-outros (fam não remun, etc.)	P	1									X			
Actividade económica / pop empregada	P	4							X	X	X		X	
Activ econ / pop desemp proc novo emprego	P	1										X		
Variáveis obrigatórias, sem qualquer fonte administrativa														
Edifício_Alojamento	Forma de ocupação		0											
	Regime de ocupação		0											
	Tipo de aquecimento		0											
	Banho ou duche		0											
Família_Indivíduo	Dimensão da família		0											
	Relação de parentesco		0											
	Tipo de família		0											
	Residência 1 ano antes		0											
	Residência fora de Portugal (alg vez / ano)		0											
	Situação Prof / pop desemp proc novo emprego		0											
CPA-domésticos e outros		0												

(*): Dadas as características das unidades estatísticas “edifício” e “alojamento”, em termos de estabilidade, assume-se que os Censos 2011 constituem uma “fonte administrativa” para o parque habitacional construído até março de 2011.

(**): A variável “Residência habitual” não consta diretamente na principal fonte relacionada com a “População”, a BDIC. (***): Sistema de indicadores de operações urbanística. (****): Ficheiro ainda não recebido pelo INE.

Quadro 5 – Variáveis (nacionais), com e sem fonte administrativa

Unidade estatística	Variáveis nacionais (Censos 2011, não obrigatórias para a União Europeia)	Cobertura Total / Parcial	Nº fontes admin	Fonte administrativa										
				SIOU	IMI	IRS	BDIC	SEF	EDUC	SS	CGA	QP	IEFP	ADSE
Variáveis nacionais, com 1 ou + fontes administrativas (assegurando diferentes níveis de cobertura do universo)														
Edifício_Alojamento	Nº de pisos	T (*)	2	X	X									
	Elevador	T (*)	1		X									
	Tipo de utilização	T (*)	2	X	X									
	Esgotos	T (*)	2	X	X									
	Estacionamento	T (*)	2	X	X									
	Entidade proprietária	P	1		X									
	Encargos de aquisição	P	1			X								
	Escalão de renda	P	1			X								
Família_Indivíduo	Companheiro em união de facto	P	3			X				X				X
	Frequência de ensino	T	1						X					
	Ano de escolaridade	T	1						X					
	Fonte de rendimento (trabalho, reforma, etc)	P	1			X								
	Procura emprego - tempo	P	1										X	
	Disponibilidade p/ trabalhar	P	1										X	
	Nº horas trabalhadas	P	1									X		
Nº pessoas da empresa	P	1									X			
Variáveis nacionais, sem qualquer fonte administrativa														
Edifício_Alojamento	Material de revestimento exterior		0											
	Estrutura de construção		0											
	Tipo de cobertura		0											
	Configuração R/C		0											
	Edifício isolado / alto		0											
	Edifício gaveto / geminado		0											
	Edifício + ou - alto que adjacentes		0											
	Necessidade de reparações		0											
	Acessibilidade da entrada		0											
	Acessibilidade até entrada no alojamento		0											
	Recolha de resíduos sólidos urbanos		0											
	Ar condicionado		0											
	Fonte de energia		0											
	Tipo de contrato de arrendamento		0											
Época do contrato		0												
Família_Indivíduo	Presente momento censitário		0											
	Residência 5 anos antes		0											
	Resid fora de Portugal (país		0											
	Curso superior		0											
	Alfabetismo		0											
	Tempo p/ local trabalho / estudo		0											
	Meio de transporte 1 p/ local trabalho / estudo		0											
	Meio de transporte 2 p/ local trabalho / estudo		0											
	Trabalhou alguma vez		0											
	Tarefas na profissão		0											
	Bens e serviços produzidos		0											
	Dificuldade ver		0											
	Dificuldade ouvir		0											
	Dificuldade andar		0											
	Dificuldade memória		0											
	Dificuldade banho / vestir		0											
	Dificuldade compreender		0											
Religião		0												

(*):Dadas as características das unidades estatísticas “edifício” e “alojamento”, em termos de estabilidade, assume-se que os Censos 2011constituem uma “fonte administrativa” para o parque habitacional construído até março de 2011.

Como se referiu anteriormente, para uma parte das variáveis relativamente às quais existe informação nas fontes administrativas, verifica-se um problema de **cobertura** relativamente ao universo que se pretende quantificar ou caracterizar.

A título de exemplo, refira-se a não existência de dados administrativos que permitam caracterizar os funcionários públicos ao nível do enquadramento profissional. Também a nível geográfico existem limitações em termos de cobertura, nomeadamente a não existência de informação, ao nível do registo individual, dos alunos a frequentar o ensino básico e secundário para a Região Autónoma dos Açores.

No quadro 6 faz-se uma síntese de alguns dos problemas diagnosticados neste âmbito.

Quadro 6 – Síntese da análise de cobertura relativa ao universo de algumas variáveis existentes em fontes administrativas

Unidade estatística	Variável	Nº fontes admin	Problemas de cobertura em variáveis com 1 ou + fontes administrativas
Variáveis obrigatórias (União Europeia)			
Família_Individuo	Nível de ensino completo	3	Não há informação para o total da população, apenas para os alunos de um determinado ano e para os trabalhadores por conta de outrem inscritos nos QP.
	Local de trabalho (País, Município, Freguesia)	4	Em falta os dados para os trabalhadores por conta própria e os sem qualquer regime de segurança social.
	CPA-população empregada	3	Em falta os dados para outros regimes de segurança social, para além da SS e CGA. Em falta os indivíduos sem qualquer regime.
	CPA-reformados	2	Em falta os dados para outros regimes de segurança social, para além da SS e CGA. Em falta os indivíduos sem qualquer regime nacional.
	CPA-pop desempregada	2	Estudo de cobertura dos dados do IEFP ainda não efetuado.
	Profissão / população empregada	2	Trabalhadores da Administração Pública e trabalhadores por conta própria, sem fonte.
	Situação Prof-patrões / empregadores	1	Em falta os dados para outros regimes de segurança social, para além da SS e CGA. Em falta os indivíduos sem qualquer regime.
	Situação Prof-trab conta de outrem	3	Em falta os dados para outros regimes de segurança social, para além da SS e CGA. Em falta os indivíduos sem qualquer regime.
	Situação Prof-trab conta própria	1	Em falta os dados para outros regimes de segurança social, para além da SS e CGA. Em falta os indivíduos sem qualquer regime.
	Situação Prof-outros (fam não remun, etc.)	1	Estudo de cobertura ainda não efetuado.
Actividade económica / pop empregada	4	Em falta os dados para os trabalhadores por conta própria e os sem qualquer regime de segurança social. Limitações na desagregação da CAE para os trabalhadores da Administração Pública.	
Variáveis nacionais			
Edifício_Alojamento	Entidade proprietária	1	Estudo de classificação de NIF (no IMI) ainda não efetuado.
	Encargos de aquisição	1	Estudo de cobertura do IRS ainda não efetuado.
	Escalão de renda	1	Estudo de cobertura do IRS ainda não efetuado.
Família_Individuo	Companheiro em união de facto	3	Estudo de cobertura do IRS ainda não efetuado.
	Fonte de rendimento (trabalho, reforma, etc)	1	Estudo de cobertura do IRS ainda não efetuado.
	Procura emprego - tempo	1	Estudo de cobertura dos dados do IEFP ainda não efetuado.
	Disponibilidade p/ trabalhar	1	Estudo de cobertura dos dados do IEFP ainda não efetuado.
	Nº horas trabalhadas	1	Trabalhadores da Administração Pública sem fonte.
	Nº pessoas da empresa	1	Trabalhadores da Administração Pública sem fonte.

PONTOS-CHAVE

- A implementação de um modelo baseado exclusivamente em ficheiros administrativos está condicionada em Portugal, desde logo porque há lacunas nas variáveis obrigatórias a nível internacional. Os ficheiros administrativos permitem obter uma cobertura total para 12 das 27 variáveis comunitárias.
- Ausência de informação administrativa em domínios essenciais como a família.
- Em termos de cobertura, há variáveis disponibilizadas pelos ficheiros que permitem apenas uma cobertura e caracterização parcial da população residente. Os ficheiros administrativos fornecem informação parcial para 6 das variáveis comunitárias.
- Para as variáveis de utilização nacional, os ficheiros administrativos não permitem obter informação para 33 das 50 variáveis disponibilizadas nos Censos 2011.
- Os registos administrativos são provenientes de fontes independentes entre si, pelo que os dados apresentam, por vezes, inconsistências para a mesma variável e para registos associados a um mesmo indivíduo.
- As questões relativas à normalização dos diferentes ficheiros são fundamentais num processo de interligação de bases de dados e integração de informação.

2.4. Interligação de bases de dados

2.4.1. Interligação de dados dos Censos e dos registos administrativos

Um dos primeiros passos na construção da base de dados da população residente em Portugal, a partir de ficheiros administrativos, foi perceber se era possível localizar os indivíduos recenseados nos Censos 2011 nos diferentes ficheiros administrativos com a finalidade de reproduzir a população nesse ano via administrativa. Para tal, foi realizado um exercício de interligação de ficheiros, ou *record linkage* (como habitualmente designado na literatura internacional de referência), entre os registos dos Censos 2011 e os registos anonimizados disponíveis nos ficheiros administrativos cedidos ao INE para este fim.

Colocação de NIC nos Censos 2011

Dado que os registos dos Censos 2011, não possuíam qualquer identificador numérico, designadamente, Número de Identificação Civil (NIC), Número de Identificação Fiscal (NIF) ou Número de Identificação da Segurança Social (NISS), foi usado um método de **interligação determinístico**, com recurso a chaves de ligação baseadas nas características dos indivíduos.

Foram utilizadas 12 chaves de ligação distintas, que combinaram de forma diferenciada a informação relativa às características dos indivíduos, nomeadamente, nome (primeiros 3 caracteres do primeiro nome; 3 primeiros ou últimos caracteres do último nome), sexo, data de nascimento, freguesia ou município de residência, naturalidade e estado civil.

No final deste processo, se o par de registos tinha correspondência no conjunto predefinido de características então o par foi considerado ligado, caso contrário não foi possível ligar o par de registos em comparação.

Refira-se que as chaves de ligação utilizadas foram ajustadas às comparações entre os dados dos Censos 2011 e cada um dos diferentes ficheiros administrativos, por motivos de diferenças estruturais da informação de cada fonte.

Para os registos dos Censos 2011 que não foi possível ligar aos ficheiros administrativos através das chaves de atributos, foi aplicado um método de **interligação probabilístico** recorrendo ao *software Quality Stage*. Este processo foi desenvolvido apenas na procura de registos com um ficheiro administrativo – BDIC 2010 – única fonte que tinha disponível o nome completo e a morada, variáveis fundamentais quando se recorre a este tipo de método.

Da ligação entre os Censos 2011 e os vários ficheiros administrativos foi possível colocar NIC (ou número de identificação do SEF para os indivíduos de nacionalidade estrangeira) a 9 944 756 registos censitários o que corresponde a 94,2% dos indivíduos recenseados em 2011.

A fonte BDIC foi, como seria natural, o ficheiro responsável pela colocação de NIC na esmagadora maioria (98%) dos registos dos Censos que foi possível ligar com pelo menos um ficheiro administrativo. Na atribuição de NIC, as restantes fontes tiveram contributos marginais.

Quadro 7 – Número de registos com NIC atribuído, segundo o ficheiro de origem

Ficheiro de Origem do NIC	Nº de registos	%
BDIC2010	9 032 686	90,8
BDIC2014	514 762	5,2
BDIC 2010 (<i>Quality Stage</i>)	190 567	1,9
DGEEC11	10 445	0,1
DGEEC13	25 063	0,3
SS13	70 186	0,7
SEF2011*	66 758	0,7
SEF2012*	5 776	0,1
SEF2013*	2 598	0,0
CGA10	3 364	0,0
CGA13	531	0,0
IEFP11	14 640	0,1
IEFP13	7 380	0,1
TOTAL	9 944 756	100,0

* O identificador corresponde ao título/cartão de residência

Dos registos dos Censos 2011, não foi possível obter correspondência, nas bases administrativas, para cerca de 617 mil indivíduos. Para estes indivíduos, não ligados nas fases automatizadas de ligação determinística ou probabilística, está em curso (aplicação ainda numa fase experimental) um plano de interligação manual de registos, ou *clerical matching*, que consiste na procura manual, de cada um dos registos, nos ficheiros administrativos, nomeadamente na BDIC 2010, ficheiro onde estão disponíveis as variáveis nome completo e morada.

Colocação de NIF e NISS nos Censos 2011

Tendo como objetivo potenciar um aumento no número de ligações entre ficheiros, através de identificadores numéricos, e na tentativa de ultrapassar a

limitação de em Portugal não existir um número único, para além do NIC (adicionado no processo inicial de interligação Censos - ficheiros administrativos), foram adicionados sempre que possível, o NIF e o NISS aos registos censitários.

O quadro 8 mostra a distribuição dos identificadores numéricos por fonte administrativa, associados à respetiva taxa de preenchimento. Note-se que a generalidade dos ficheiros administrativos possui, em regra, apenas um ou dois dos principais números de identificação utilizados em Portugal. Por outro lado, em alguns casos, verificam-se taxas elevadas de não preenchimento destes identificadores.

Quadro 8 – Ficheiros administrativos e identificadores numéricos (taxas de preenchimento), por fonte

Ficheiro de origem	NIC	NIF	NISS
IRN2010	100,0%	-	-
IRN2014	100,0%	-	-
SEF2011*	99,9%	57,8%	49,9%
SEF2012*	99,9%	59,9%	50,6%
SEF2013*	100,0%	61,4%	51,1%
SS2011	73,9%	96,4%	100,0%
SS2013	79,3%	97,6%	100,0%
QP2010	-	-	98,8%
QP2011	-	-	98,9%
QP2012	-	-	98,9%
QP2013	-	-	99,0%
IEFP2011	100,0%	97,9%	97,2%
IEFP2012	100,0%	98,8%	98,1%
IEFP2013	100,0%	99,3%	98,6%
CGA2010	71,4%	84,6%	-
CGA2013	77,0%	81,9%	-
EDUC2011	98,6%	-	-
EDUC2013	100,0%	-	62,5%
AT2014 – Cadastro	-	100,0%	-

* O identificador NIC corresponde ao título/cartão de residência

Nos quadros 9 a 11, mostram-se, por sua vez, os resultados da colocação dos identificadores numéricos NIF e NISS na tabela gerada pelo processo de interligação de ficheiros, com a contribuição de cada fonte administrativa.

Quadro 9 - Nº de registos com NIF atribuído segundo o ficheiro de origem

Ficheiro de origem	Nº de registos	%
SS2011	4 553 386	45,8
SS2013	1 260 477	12,7
SEF2011	56 216	0,6
SEF2013	5 981	0,1
CGA2010	574 856	5,8
CGA2013	21 430	0,2
IEFP2011	146 808	1,5
IEFP2013	117 801	1,2
AT2014 – Cadastro	2 103 762	21,2
Sem NIF atribuído	1 104 039	11,1
TOTAL	9 944 756	100,0
Total com NIF	8 840 717	88,9

Assim, através da interligação dos ficheiros foi possível colocar NIF a cerca de 8,8 milhões de registos dos Censos 2011. As fontes com maior contributo na colocação de NIF foram a Segurança Social e o ficheiro do Cadastro dos Contribuintes da Autoridade Tributária.

Relativamente ao Identificador NISS, a taxa de preenchimento é mais baixa que a obtida para o NIF. Foi possível atribuir NISS a cerca de 7,0 milhões de registos censitários. Como seria expectável a fonte com maior contributo para a colocação deste identificador numérico foi a Segurança Social.

Quadro 10 – Nº de registos com NISS atribuído segundo o ficheiro de origem

Ficheiro de origem	Nº de registos	%
SS2011	4 576 313	46,0
SS2013	1 280 345	12,9
SEF2011	39 709	0,4
SEF2013	3 955	0,0
QP2011	1 085	0,0
QP2013	408	0,0
IEFP2011	144 793	1,5
IEFP2013	120 432	1,2
EDUC2011	17 304	0,2
EDUC2013	821 969	8,3
Sem NISS atribuído	2 938 443	29,5
TOTAL	9 944 756	100,0
Total com NISS	7 006 313	70,5

Analisando a conjugação dos diferentes identificadores NIC, NIF e NISS (cujos resultados se ilustram no quadro 11) constatamos que, do total de registos censitários para os quais foi possível estabelecer ligação com ficheiros administrativos, cerca de 7 milhões têm os três identificadores numéricos preenchidos, o que representa cerca de 70% dos registos individuais dos Censos 2011.

Por sua vez, o número de registos com dois identificadores, nomeadamente NIC e NIF são cerca de 2 milhões, sendo que, com apenas os identificadores NIC e NISS em simultâneo são pouco mais de 200 mil registos.

Quadro 11 – Nº de registos segundo o preenchimento dos identificadores numéricos NIC, NISS e NIF

Identificadores	Nº de registos	%
NIC, NISS, NIF	6 806 151	68,4
NIC, NIF	2 034 566	20,5
NIC, NISS	200 162	2,0
NISS, NIF	0	0,0
NIC	903 877	9,1
NISS	0	0,0
NIF	0	0,0
Total	9 944 756	100,0

Considera-se que este é um importante avanço, na medida em que permite que as ligações entre os vários ficheiros administrativos possam ser realizadas via atributo numérico, logo de forma mais eficaz.

Refira-se ainda que, não obstante o esforço de enriquecimento dos identificadores numéricos através do cruzamento dos vários ficheiros administrativos, há ainda cerca de 900 mil registos dos Censos 2011 para os quais apenas está preenchido o NIC.

PONTOS-CHAVE

- O processo de interligação entre os Censos 2011 e os vários ficheiros administrativos permitiu colocar NIC a 94,2% dos indivíduos recenseados em 2011.
- Cerca de 70% dos registos individuais dos Censos 2011, com ligação a pelo menos um ficheiro administrativo, têm os três identificadores numéricos preenchidos e 22% têm pelo menos dois identificadores.
- Para cerca de 900 mil indivíduos dos Censos 2011 apenas foi possível atribuir o identificador NIC.
- Não foi possível encontrar correspondência nas bases administrativas para cerca de 617 mil registos individuais dos Censos 2011.

Qualidade das ligações

Como referido no ponto anterior, foram utilizados, no processo de interligação de ficheiros, métodos determinísticos e estocásticos.

No caso do método determinístico, com recurso a chaves de ligação construídas à base de um conjunto de características, considera-se que a ligação estabelecida, ao nível do indivíduo, é uma ligação exata ou verdadeira, se os conteúdos do par de registos ligado for igual.

Já para o método probabilístico, como se ligam pares de registos com base não na total correspondência do conjunto das características comuns mas apenas através de alguns pontos de contacto, a possibilidade de identificar como ligação verdadeira, uma ligação que afinal não diz exatamente respeito ao mesmo indivíduo aumenta largamente, face ao método determinístico.

Assim, os indicadores de qualidade das ligações baseiam-se essencialmente na análise, considerando o total de pares ligados, dos:

- falsos positivos, isto é, pares identificados como ligações verdadeiras mas que afinal se revelam falsos pares;
- falsos negativos, isto é, pares de registos que não foram ligados mas que afinal estão associados a um mesmo indivíduo.

Neste âmbito, a exatidão das ligações foi estimada em cerca de 94% (o que corresponde a uma taxa de falsos positivos de aproximadamente 6%), isto é, considera-se que 94% dos pares ligados correspondem a ligações verdadeiras no que se refere ao indivíduo ligado. Estes valores resultam da atualização – face ao último relatório QUAR – do número de pares ligados entre a BDIC 2010 e os Censos 2011 através do *software IBM Quality Stage*). O método utilizado para estimar as ligações verdadeiras está descrito no relatório QUAR sobre a interligação das diferentes bases de dados provenientes de fontes administrativas, de março de 2015.

Nos casos em que há inconsistências nos dados, dados omissos ou contradições, apenas através da verificação manual – *clerical matching* na literatura de referência – é possível validar as ligações. Este método de validação e verificação das ligações é também utilizado para detetar pares de registos que não foram ligados mas que correspondem a um mesmo indivíduo (falsos negativos). O procedimento para a implementação deste método está a ser atualmente ensaiado para as freguesias do Teste de 2016.

PONTOS-CHAVE

- A avaliação da qualidade das ligações entre os Censos 2011 e os ficheiros administrativos estima que cerca de 94% dos pares ligados correspondem a ligações verdadeiras, ou seja, correspondem ao mesmo indivíduo.
- A taxa de falsos positivos, pares ligados mas que não correspondem ao mesmo indivíduo, é de aproximadamente 6%.

2.4.2. Interligação dos ficheiros administrativos e sinalização de presença de indivíduos

Esta linha de trabalho tinha como principal objetivo identificar a população constante nos diversos ficheiros administrativos e que deveria ser considerada residente em Portugal. Para tal, era importante verificar em que fontes administrativas os indivíduos recenseados nos Censos 2011 estavam representados de modo a avaliar a representatividade da população censitária nos ficheiros administrativos.

Nesse sentido, foi construída uma base de dados, onde foi assinalada a presença dos registos censitários que foi possível localizar nos diferentes ficheiros administrativos.

No processo de sinalização de presença privilegiou-se, sempre que possível, a ligação através de pelo menos um identificador numérico (NIC, NIF ou NISS), sendo que apenas se recorreu à ligação por chave de atributos para os registos cuja ligação não foi possível obter via identificador.

No quadro 12, ilustra-se a forma como foram obtidas as ligações. Verifica-se que para a maior parte dos ficheiros, a maioria dos registos foram ligados por identificador numérico.

Quadro 12 – Ligações obtidas de acordo com o tipo de ligação (%), por fonte administrativa

Fonte	Ligação por Identificador numérico (%)	Ligação por chave de atributos (%)
BDIC2010	100,0	0,0
BDIC2014	100,0	0,0
SEGS2011	100,0	0,0
SEGS2013	89,1	10,9
SEF2011	58,8	41,2
SEF2013	57,7	42,3
CGA2010	84,2	15,8
CGA2013	97,3	2,7
QP2011	99,1	0,9
QP2013	99,0	1,0
IEFP2011	91,3	8,7
IEFP2013	88,7	11,3
EDUC2011	90,7	9,3
EDUC2013	97,0	3,0

Por sua vez, no quadro 13, indicam-se os registos dos Censos 2011 para os quais foi encontrada correspondência em pelo menos um ficheiro administrativo. De salientar que 93,6% estavam presentes na BDIC 2010, valor que ascende a 97,5% para a BDIC 2014. A maior percentagem de presenças registada na BDIC 2014 deve-se ao facto de algumas crianças recenseadas nos Censos não estarem ainda inscritas na BDIC 2010.

De referir também que, nos registos localizados na BDIC 2014 estão incluídos cerca de 332 695 indivíduos com óbito assinalado, o que significa que são indivíduos que foram recenseados nos Censos 2011 e que entretanto faleceram (3,4%).

Quadro 13 – Indivíduos dos Censos com correspondência em pelo menos uma fonte administrativa, segundo a presença nos diversos ficheiros administrativos

Fonte	Nº de indivíduos	%
BDIC2010	9 308 384	93,6
BDIC2014	9 693 570	97,5
SEGS2011	5 840 238	58,7
SEGS2013	5 863 075	59,0
SEF2011	225 337	2,3
SEF2013	191 332	1,9
CGA2010	979 264	9,8
CGA2013	943 847	9,5
QP2011	2 374 907	23,9
QP2013	2 236 777	22,5
IEFP2011	624 019	6,3
IEFP2013	833 439	8,4
EDUC2011	1 678 629	16,9
EDUC2013	1 495 729	15,0

Ainda a este propósito, note-se que a Segurança Social é a segunda fonte com maior número de indivíduos encontrados, com valores de 58,7% no ano 2011 e 59,0% no ano 2013.

Refira-se ainda que, cerca de 23,9% dos indivíduos dos Censos 2011 foram encontrados nos Quadros de Pessoal 2011 e para 16,9% dos registos foi estabelecida correspondência com os ficheiros da Educação 2011.

As restantes fontes utilizadas para verificar a presença dos registos dos Censos nos ficheiros administrativos têm valores mais baixos, na ordem dos 10% para a CGA 2011, 6,3% no IEFP 2011 e de apenas 2,3% para o ficheiro do SEF 2011.

Se atendermos ao número de fontes em que os indivíduos estão representados, o resultado do cruzamento entre os Censos 2011 e os ficheiros administrativos dá conta de um significativo número de registos, que podem ter a sua residência validada em Portugal, se for considerada a sua presença em mais de uma fonte administrativa.

De facto, mais de 85% dos indivíduos dos Censos 2011 para os quais foi possível encontrar correspondência em pelo menos um ficheiro administrativo, estão representados no mínimo em duas fontes administrativas. Estas presenças foram contabilizadas por fonte (BDIC, SEGS, SEF, CGA,...) e não por ano, ou seja, mesmo que um indivíduo seja encontrado nos dois anos em análise, apenas é contabilizada numa fonte.

No quadro 14, ilustra-se a distribuição dos registos segundo o número de fontes onde o mesmo foi localizado. Refira-se que dos 1 457 089 registos que foram encontrados apenas numa fonte administrativa, a esmagadora maioria referem-se a registos que apenas foi possível identificar na BDIC (1 405 205 registos).

Quadro 14 – Indivíduos dos Censos com correspondência em pelo menos uma fonte administrativa, segundo o número de fontes onde o registo foi encontrado

Nº de fontes com presença do registo	Nº de indivíduos	%
1	1 457 089	14,7
2	4 637 344	46,6
3	3 116 071	31,3
4	671 119	6,7
5 ou mais fontes	63 133	0,6
Total	9 944 756	100,0

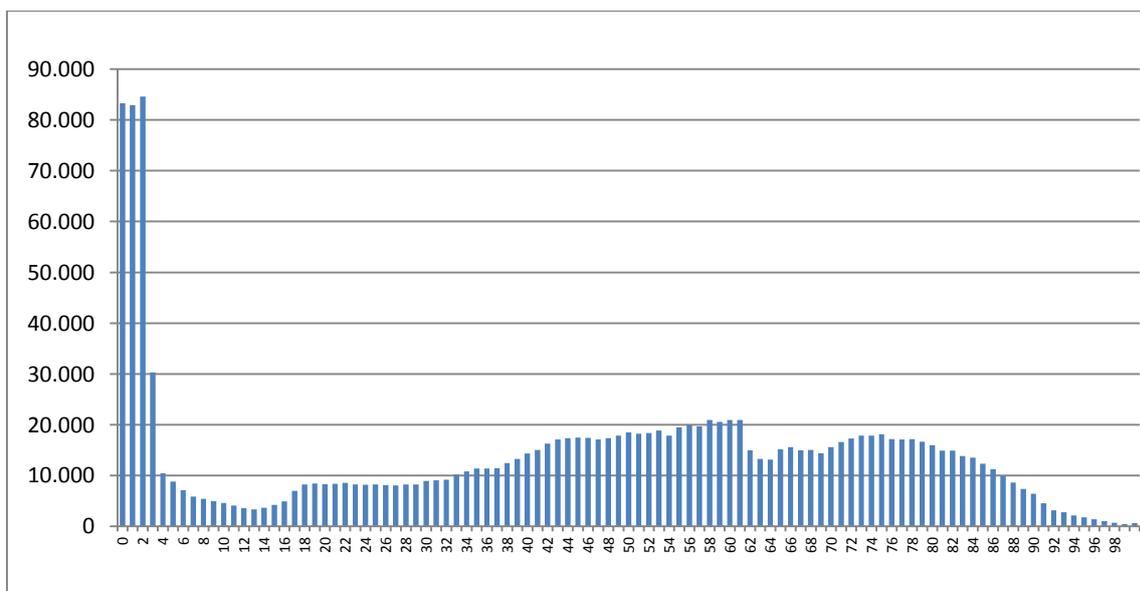
O facto de cerca de 14,1% da população dos Censos só estar representada na BDIC (ou só ter sido possível ligar à BDIC), pode representar uma limitação na aplicação da metodologia, tendo por base uma lógica de análise da residência em Portugal através de indícios de residência. Recorde-se que para efeitos de construção de uma base de população residente será necessário excluir indivíduos que estão na BDIC, e que efetivamente não são residentes em Portugal.

Caracterização dos indivíduos dos Censos 2011, para os quais só foi encontrada correspondência na fonte BDIC

A caracterização da população recenseada nos Censos 2011 e que apenas foi possível identificar na fonte BDIC, é importante para aferir se de facto essa população não se encontra representada noutros ficheiros administrativos, ou se a não associação a outro ficheiro administrativo se relaciona com dificuldades no processo de interligação.

Assim, no que se refere à idade, a população associada aos registos dos Censos 2011 que foi possível localizar apenas na BDIC (figura 2), evidencia uma maior representatividade das crianças mais novas, em particular com idades até aos 3 anos. Como era já expetável, este grupo populacional dificilmente aparece noutra fonte administrativa uma vez que ainda não frequenta o ensino e, por norma, não tem ligação aos restantes ficheiros administrativos utilizados nesta análise. Com outra ordem de grandeza, destacam-se também a maior expressão das idades compreendidas entre os 35 e os 60 anos, e ainda das idades entre os 70 e os 80 anos.

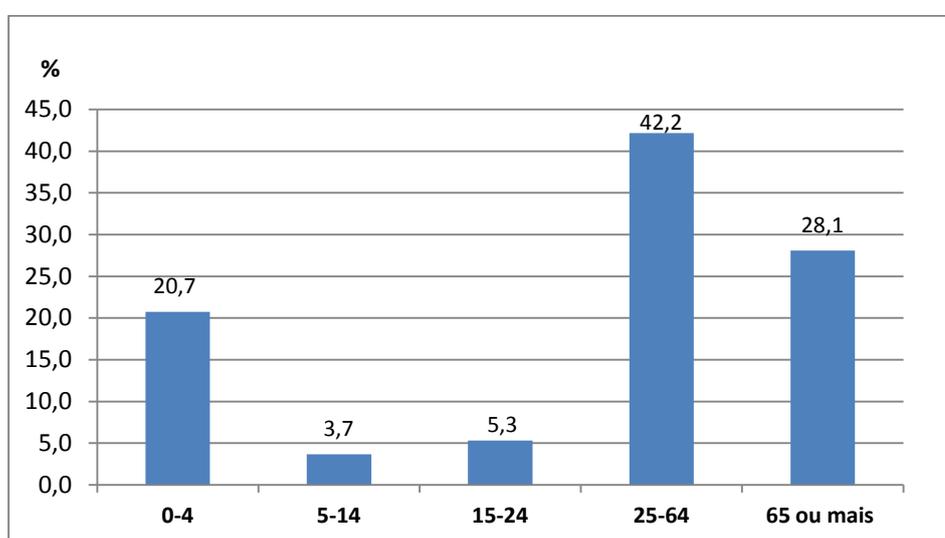
Figura 2 – Indivíduos dos Censos 2011 representados apenas na fonte BDIC, por idade ano a ano



A análise por grupo etário (figura 3) mostra que dos 1,4 milhões de indivíduos dos Censos 2011 com correspondência apenas na BDIC, 42,2% têm idades compreendidas entre os 25 e os 64 anos (o que corresponde a cerca de 590 mil indivíduos) e ainda que 28,1% têm 65 ou mais anos (394 mil indivíduos).

Como referido anteriormente, destaca-se também o grupo dos 0-4 anos, que representa cerca de 1/5 dos indivíduos que fizeram correspondência apenas com a BDIC.

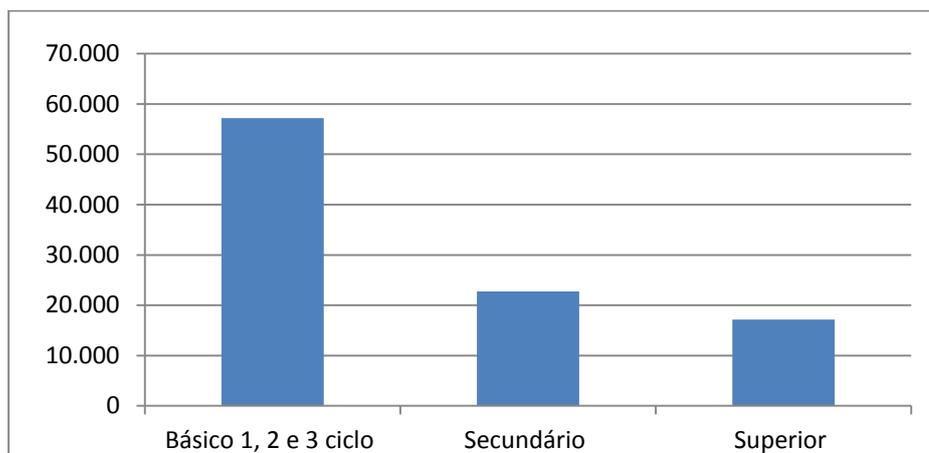
Figura 3 – Indivíduos dos Censos 2011 representados apenas na fonte BDIC, por grandes grupos etários



Cerca de 97 mil indivíduos localizados apenas na fonte BDIC estavam, segundo os Censos 2011, a frequentar um nível de ensino a partir do 1º Ciclo do Ensino Básico (não foi considerado o ensino pré-escolar, uma vez que só se dispõe desta informação para a Região Autónoma da Madeira e não para a totalidade do país. Também não existe informação disponível, ao nível do registo individual dos alunos do Ensino Básico e Secundário para a Região Autónoma dos Açores).

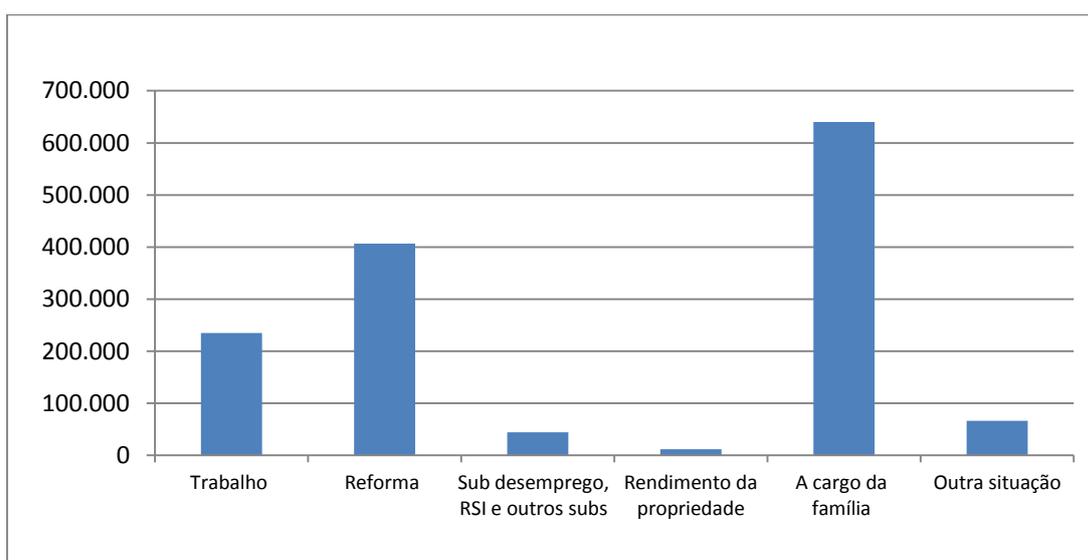
Neste contexto, com exceção dos indivíduos dos Açores a frequentar o Ensino Básico e Secundário, seria exetável localizar os indivíduos, que frequentavam os níveis de ensino a partir do 1º Ciclo do Ensino Básico nos ficheiros da Educação. As dificuldades de ligação aos ficheiros da Educação ou, eventualmente, algum desajustamento na referência temporal da informação podem justificar os resultados obtidos.

Figura 4 – Indivíduos dos Censos 2011, representados apenas na fonte BDIC, a frequentar o ensino, por nível de ensino



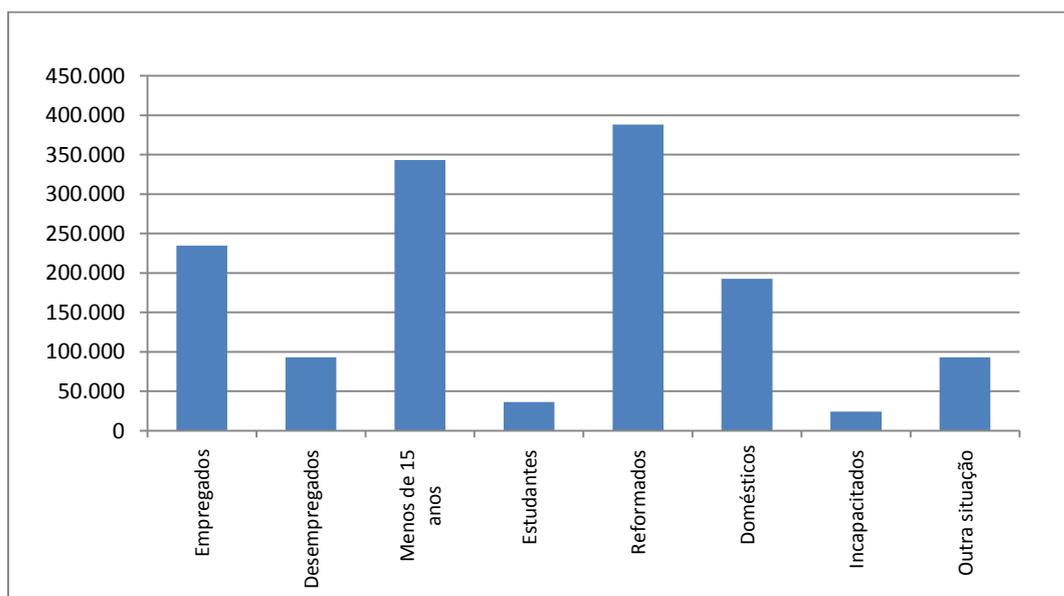
Considerando as variáveis de natureza económica, a figura 5 mostra a distribuição dos indivíduos com ligação apenas à BDIC segundo o principal meio de vida. Os resultados mostram que a maior parte dos indivíduos estão a “Cargo da Família”. Destaca-se igualmente o meio de vida “Reforma” com valores na ordem dos 400 mil indivíduos, o que, pelo menos no plano teórico, implicaria uma representação nos ficheiros da Segurança Social ou na CGA.

Figura 5 – Indivíduos dos Censos 2011 representados apenas na fonte BDIC, por principal meio de vida



Conforme se ilustra na figura 6, analisando a condição perante o trabalho dos indivíduos recenseados nos Censos 2011 com ligação apenas à BDIC, verifica-se que o grupo mais representado é o dos “Reformados”, com cerca de 400 mil indivíduos. Também neste sentido, seria expectável que, parte destes indivíduos, reformados, estivessem representados nos ficheiros da Segurança Social ou da CGA. O facto de isso não acontecer pode indiciar algumas dificuldades na ligação e identificação dos indivíduos nos referidos ficheiros. Obviamente que também existem reformados portugueses que trabalharam no estrangeiro, regressaram ao país e recebem reformas dos países onde trabalharam, não estando por isso representados nos ficheiros analisados.

Figura 6 – Indivíduos dos Censos 2011 representados apenas na fonte BDIC, por condição perante o trabalho



Também o grupo dos “Empregados”, com cerca de 230 mil indivíduos, tem uma elevada representação entre os recenseados em 2011 que apenas foi possível localizar na BDIC. Também aqui surge alguma incerteza, na medida em que deveriam ser indivíduos inscritos na Segurança Social ou na CGA mas para os quais não foi possível encontrar ligação a essas fontes (pode equacionar-se serem indivíduos que eventualmente trabalham na economia informal).

No quadro 15, ilustra-se a distribuição das principais profissões da população que foi possível ligar apenas à BDIC. Verifica-se que os “Trabalhadores da limpeza em casas particulares, hotéis, ...”, os “Trabalhadores qualificados da construção civil” e os “Vendedores em lojas” são as profissões mais representadas. Se, para as duas primeiras profissões, a importância da “economia não registada” poderá ser relevante na explicação dos números, já para o grupo dos “Vendedores em lojas” não será expectável que esse possa ser o principal motivo para estes indivíduos não serem localizados, pelo menos, no ficheiro da Segurança Social.

De salientar, ainda, o grupo dos “Advogados e Solicitadores” como a quarta profissão mais representada, o que estará relacionado com a existência da Caixa de Previdência dos Advogados e Solicitadores, não estando, por esse motivo, estes indivíduos nos ficheiros analisados.

Quadro 15 – Indivíduos dos Censos 2011 representados apenas na BDIC, segundo as 10 principais profissões

Profissão	Nº de indivíduos
Trab. da limpeza em casas particulares, hotéis...	26 769
Trab. qualificados da construção civil	25 318
Vendedores em lojas	24 523
Advogados e Solicitadores	12 429
Empregados escritório	12 286
Trab. não qualificados industria	9 272
Trab. da confeção e vestuário	7 682
Empregados de mesa e bar	6 834
Cozinheiro	6 566
Motoristas de veículos pesados e autocarros	6 220

PONTOS-CHAVE

- Na população recenseada nos Censos 2011 e que apenas foi possível identificar na fonte BDIC (1, 4 milhões), destacam-se as crianças entre os 0-4 anos que representam cerca de 20%, e os indivíduos com 65 ou mais anos com um peso de 28% face ao total.
- A análise das características económicas dos indivíduos dos Censos 2011, com representação apenas na BDIC, revela algumas dificuldades no processo de interligação dos ficheiros, nomeadamente a existência de cerca de 400 mil indivíduos reformados e de 230 mil empregados que não foi possível localizar na Segurança Social ou na CGA.
- A identificação da população apenas na BDIC, deve-se por um lado ao facto de determinados universos populacionais não estarem representados noutros ficheiros administrativos, como é o caso das crianças que ainda não frequentam o 1º ciclo do ensino básico; e por outro das dificuldades de interligação de ficheiros.
- A articulação e complexidade de todas estas dimensões e hipóteses, dificulta o processo metodológico de construção da BPR, tendo por base um método de indícios de residência.

Distribuição dos registos das fontes administrativas em função da ligação aos Censos 2011

O quadro16 mostra o número de registos ligados e não ligados aos Censos 2011, considerando os universos de cada fonte administrativa. De facto, se analisarmos o nível de interceção, entre os Censos 2011 e os diferentes ficheiros administrativos, verifica-se que em algumas situações os valores estão abaixo do que seria expectável.

Se para algumas fontes, dadas as características intrínsecas, não era possível, à partida, prever um nível de ligação com os Censos na totalidade – caso da BDIC

–, já para outras, era esperado um nível de ligação mais elevado – caso dos ficheiros da Educação ou da Segurança Social. Neste sentido, os resultados ficam abaixo do expectável, já que a percentagem de registos ligados foi de 81% para a Segurança Social 2011 e de 85,4% no caso da Educação 2011.

O caso particular do SEF, onde a percentagem de registos encontrados nos Censos 2011 foi de apenas 51,8%, está relacionado com o facto de não ter sido possível estabelecer qualquer ligação a cerca de 1/3 dos estrangeiros recenseados em 2011. Dos cerca de 600 mil indivíduos dos Censos para os quais não foi possível estabelecer qualquer ligação a uma fonte administrativa cerca de 143 mil eram de nacionalidade estrangeira.

Quadro 16 – Registos administrativos encontrados e não encontrados nos Censos 2011

Ficheiros Administrativos	Nº de registos do ficheiro	Encontrados nos Censos 2011	Não encontrados nos Censos 2011	% de registos encontrados nos Censos 2011	% de registos não encontrados nos Censos 2011
BDIC2010	11 565 714	9 308 384	2 257 330	80,5	19,5
BDIC2014	11 884 913	9 693 570	2 191 343	81,6	18,4
SEGS2011	7 209 027	5 840 238	1 368 789	81,0	19,0
SEGS2013	7 124 233	5 863 075	1 261 158	82,3	17,7
SEF2011	434 708	225 337	209 371	51,8	48,2
SEF2013	398 268	191 332	206 936	48,0	52,0
CGA2010	1 103 980	979 264	124 716	88,7	11,3
CGA2013	1 085 872	943 847	142 025	86,9	13,1
QP2011	2 736 659	2 374 907	361 752	86,8	13,2
QP2013	2 554 641	2 236 777	317 864	87,6	12,4
IEFP2011	702 215	624 019	78 196	88,9	11,1
IEFP2013	897 258	833 439	63 819	92,9	7,1
EDUC2011	1 965 842	1 678 629	287 213	85,4	14,6
EDUC2013	1 706 779	1 495 729	211 050	87,6	12,4
ADSE2012	1 377 353	886 188	491 165	64,3	35,7
ADSE2013	1 290 784	980 148	310 636	75,9	24,1

De modo a quantificar e caracterizar os indivíduos constantes dos registos administrativos para os quais não foi possível estabelecer ligação aos Censos 2011, foi construída uma tabela de dados com a totalidade dos registos administrativos não encontrados. O objetivo foi identificar os registos únicos, mas também aferir a eventual representação desses registos nas diversas fontes. Esta tabela tem cerca de 4,4 milhões de registos, a maioria dos quais consta apenas

numa única fonte. Destes registos administrativos que não tiveram correspondência com os Censos 2011, estão cerca de 2 milhões de registos que estão apenas na fonte BDIC, podendo representar os indivíduos que estando inscritos com uma morada em Portugal não residem efetivamente no país.

Em suma, através dos registos administrativos foi possível identificar como únicos cerca de 14,3 milhões de registos individuais, dos quais 9,9 milhões possuíam correspondência com os Censos 2011 e 4,4 milhões não estavam representados na operação censitária.

PONTOS-CHAVE

- O facto de Portugal não possuir uma base de população residente e não ter um identificador numérico único para cada cidadão, que seja comum nas várias fontes administrativas, dificulta o processo de interligação de ficheiros; não obstante, é de salientar os avanços conseguidos no enriquecimento dos registos com identificadores numéricos.
- A restrição da CNPD, no que refere aos nomes e moradas, dificulta o processo de interligação, fragiliza o processo de estimar as principais medidas da qualidade das ligações e impede a implementação de um processo de verificação manual de registos – prática essencial em processos de interligação de ficheiros que dizem respeito a registos associados a indivíduos.
- O processo de interligação de ficheiros, requer técnicas de programação avançadas e *software* especializado que, por se tratar de um processo pioneiro no INE, não está ainda otimizado.
- A dimensão e as especificidades dos ficheiros envolvidos neste processo – no seu conjunto, os registos das várias fontes administrativas são mais de 80 milhões – torna-o complexo e moroso.
- Os resultados apresentados indiciam dificuldades na identificação dos indivíduos em todos os ficheiros em que os mesmos eventualmente possam estar representados.

2.5. Estimação da população residente em 2011 em função de regras de residência

Após o processo de interligação dos vários ficheiros administrativos disponíveis e da sinalização da existência dos registos nas diferentes fontes, segue-se a construção de um primeiro protótipo da BPR. Nesta fase, o objetivo é estimar a população que reside em Portugal e para a qual há informação consistente nos ficheiros administrativos.

Para uma primeira aproximação à estimação da população residente com base em ficheiros administrativos foi perspectivada uma metodologia simplificada, que tem como objetivo incluir apenas registos para os quais existem fortes evidências administrativas de que o indivíduo é residente em Portugal.

Em termos muito gerais, para que um indivíduo seja incluído no protótipo da BPR é necessário que esteja presente em pelo menos duas fontes administrativas, sendo uma delas necessariamente a BDIC ou o SEF.

O desenvolvimento desta metodologia, tem como objetivo ultrapassar o facto da BDIC não ser um ficheiro de população residente, e como tal a presença neste ficheiro, ainda que com uma morada associada ao território nacional, não significar a residência efetiva no país.

Assim, numa tentativa de ultrapassar as limitações associadas à BDIC, apenas é considerado residente em Portugal um cidadão com registo da BDIC se o mesmo existir pelo menos em mais um ficheiro administrativo como seja a Segurança Social, Educação, Quadros de Pessoal, etc. Neste sentido foi definido um conjunto de regras de residência, a aplicar aos registos da BDIC e do SEF e que definem a inclusão dos indivíduos na BPR (quadro 17).

Quadro 17 – Regras para a inclusão de registos na BPR

Regra	Especificações
Regra 1	Registo existente na BDIC 2010 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro) e pelo menos na Segurança Social 2011
Regra 2	Registo existente na BDIC 2010 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro) e pelo menos na Educação 2011
Regra 3	Registo existente na BDIC 2010 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro) e pelo menos nos QP 2011
Regra 4	Registo existente na BDIC 2010 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro) e pelo menos na CGA 2010
Regra 5	Registo existente na BDIC 2010 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro) e pelo menos no IEF2011
Regra 6	Registo existente na BDIC 2010 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro) e pelo menos na ADSE 2012
Regra 7	Registo existente na BDIC 2014 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro, Estado V) e pelo menos na Segurança Social 2011
Regra 8	Registo existente na BDIC 2014 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro, Estado V) e pelo menos na Educação 2011
Regra 9	Registo existente na BDIC 2014 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro, Estado V) e pelo menos nos QP 2011
Regra 10	Registo existente na BDIC 2014 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro, Estado V) e pelo menos na CGA 2010
Regra 11	Registo existente na BDIC 2014 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro, Estado V) e pelo menos no IEF2011
Regra 12	Registo existente na BDIC 2014 (com NIC <i>not null</i> ; COD_RESIDENCIA_CONV diferente dos códigos estrangeiro, Estado V) e pelo menos na ADSE 2012, desde que data de nascimento anterior ou igual a 20111231
Regra 13	Registo existente no SEF 2011 (com SEF <i>not null</i>) e pelo menos na Segurança Social 2011
Regra 14	Registo existente no SEF 2011 (com SEF <i>not null</i>) e pelo menos na Educação 2011
Regra 15	Registo existente no SEF 2011 (com SEF <i>not null</i>) e pelo menos nos QP 2011
Regra 16	Registo existente no SEF 2011 (com SEF <i>not null</i>) e pelo menos na CGA 2010
Regra 17	Registo existente no SEF 2011 (com SEF <i>not null</i>) e pelo menos no IEF2011
Regra 18	Registo existente no SEF 2011 (com SEF <i>not null</i>) e pelo menos na ADSE 2012

Para a construção deste primeiro protótipo da BPR 2011 foram utilizados os ficheiros administrativos com a referência temporal mais próxima de 2011. A este propósito, muito embora ocorram diferenças entre as referências temporais exatas das diferentes fontes, considera-se como data de referência deste protótipo da BPR o final daquele ano.

Neste ponto da investigação, identifica-se a necessidade de definir regras adicionais às agora implementadas. Por diversos motivos, há indivíduos que, por exemplo, estão apenas presentes na BDIC, como seja o caso das crianças antes da entrada no sistema de ensino, ou pessoas que se ocupam das tarefas do lar, entre outras situações. Estas e outras limitações implicam um estudo e uma investigação mais aprofundada, que terá que ser devidamente programada e calendarizada e que não foi ainda possível iniciar.

PONTOS-CHAVE

- A BDIC é o ficheiro central na estrutura da BPR, complementado com os registos de cidadãos estrangeiros da responsabilidade do SEF.
- A metodologia de construção da BPR baseia-se, em linhas gerais, em considerar residentes os cidadãos portugueses que constem na BDIC e os cidadãos estrangeiros que constem no ficheiro do SEF, depois de verificar a existência de indícios de residência, isto é, a presença de registos associados ao indivíduo, noutras fontes administrativas.
- Devido às características da BDIC, neste primeiro protótipo foram excluídos todos os registos que apenas estão presentes nesta fonte administrativa.
- A metodologia aplicada tem um conjunto de limitações associadas, pelo que é fundamental desenvolver e aprofundar a investigação e o estudo realizado.

2.6. Introdução de variáveis socioeconómicas na BPR 2011

Após a identificação da população considerada residente e aplicação das regras que permitem decidir acerca da sua inclusão no protótipo da BPR 2011 (conforme

descrito no ponto anterior), procedeu-se à seleção das variáveis disponibilizadas pelas diferentes fontes administrativas que permitem caracterizar, em termos demográficos e socioeconómicos, essa população.

A BDIC é a fonte administrativa de referência para a contabilização da população residente em Portugal e respetiva caracterização das variáveis demográficas básicas. Assim, a informação de cariz demográfico é proveniente da fonte BDIC e necessariamente da fonte SEF para os indivíduos de nacionalidade estrangeira.

Ao nível das variáveis socioeconómicas, a definição da condição perante o trabalho e respetiva caracterização é condicionada pela existência de informação dispersa por diferentes fontes que contribuem para esta temática. A informação disponibilizada pelas várias fontes precisa de ser devidamente integrada e compatibilizada no sentido de produzir informação consistente a este nível. Nesta fase dos trabalhos, o protótipo da BPR dispõe apenas da informação original constante nos diferentes ficheiros, não estando ainda realizada a integração da informação das diferentes fontes.

Neste primeiro protótipo de BPR constam 47 variáveis com origem em 7 fontes administrativas. De referir que algumas das fontes utilizadas no processo de interligação dos ficheiros e que foram utilizadas nas regras de residência não contribuem com informação adicional ao nível das variáveis que caracterizam os indivíduos.

Após a constituição da tabela com os registos resultantes da aplicação das regras de residência, foram incluídas na BPR as variáveis dos ficheiros administrativos que se listam no quadro 18.

Quadro 18 – Variáveis dos ficheiros administrativos incluídas no protótipo da BPR

Designação da Variável	Código da Variável	Designação do campo no ficheiro de origem	Ficheiro de Origem
Ano de referência	ANO	2011	----
Código de identificação nos Censos 2011	IND	IND	Tabela interligação*
Número de Identificação Civil (encriptado)	NIC	NIC	Tabela interligação*
Número no SEF(encriptado)	Nº SEF	SEF	Tabela interligação*
Número de Identificação Fiscal (encriptado)	NIF	NIF	Tabela interligação*
Número de Identificação da Segurança Social (encriptado)	NISS	NISS	Tabela interligação*
Distrito de residência	DT_RESID_COD	DT_RESID_COD	BDIC 2010, BDIC 2014 ou SEF
Município de residência	MN_RESID_COD	MN_RESID_COD	BDIC 2010, BDIC 2014 ou SEF
Freguesia de residência	FR_RESID_COD	FR_RESID_COD	BDIC 2010, BDIC 2014 ou SEF
Código postal	CP4	CP4	BDIC 2010, BDIC 2014 ou SEF
Código postal auxiliar	CP3	CP3	BDIC 2010, BDIC 2014 ou SEF
Localidade Postal	LOCPOSTAL	CP_LOCAL_POSTAL (BDIC) LOCALIDADE (SEF)	BDIC 2010, BDIC 2014 ou SEF
Nome abreviado (3 primeiras letras do primeiro nome +3 últimas letras do último nome)	NOME	NOME	BDIC 2010, BDIC 2014 ou SEF
Sexo	SEXO	SEXO	BDIC 2010, BDIC 2014 ou SEF
Data de nascimento	DTNASC	DTNASC	BDIC 2010, BDIC 2014 ou SEF
Estado civil	ESTCIVIL	ESTCIVIL	BDIC 2010, BDIC 2014 BDIC 2010, BDIC 2014 ou SEF
País de nacionalidade	PAIS_NAC_COD	PAIS_NAC_COD	BDIC 2010, BDIC 2014 ou SEF
País de naturalidade	PAIS_NAT_COD	PAIS_NAT_COD	BDIC 2010, BDIC 2014 ou SEF
Município de naturalidade	MN_NAT_COD	MN_NAT_COD	BDIC 2010, BDIC 2014
Nível de ensino a frequentar	NENSINO_DGEEC	A10_ENSINO	DGEEC 2011
Qualificação na SS	TIPO_QLF_SS	TIPO_QLF	SS 2011
NIF da entidade empregadora	NIFEMP_SS	NIPC	SS 2011
NISS da entidade empregadora	NISSEMP_SS	NISS_EE	SS 2011
Nome da entidade empregadora	NOME_EMP_SS	FIRMA_DENOMINACAO	SS 2011
Freguesia da entidade empregadora	FREGEMP_SS	COD_FREGUESIA_EE	SS 2011
Município da entidade empregadora	MNEMP_SS	COD_CONCELHO_EE	SS 2011
Habilitações literárias	HABIL_QP	HABIL	QP 2011
Situação na profissão	SITPRO_QP	SITPRO	QP 2011
Profissão	PROF_QP	PROF	QP 2011
Número de horas trabalhadas	HORASTRAB_QP	PNT	QP 2011

Designação da Variável	Código da Variável	Designação do campo no ficheiro de origem	Ficheiro de Origem
NIF da entidade empregadora	NIFEMP_QP	NIF	QP 2011
Freguesia da entidade empregadora - empresa	FREGEMP_QP	FREGEMP	QP 2011
Freguesia da entidade empregadora - estabelecimento	FREGEST_QP	FREGEST	QP 2011
Freguesia da entidade empregadora - empresa	MNEMP_QP	IDEMP	QP 2011
Freguesia da entidade empregadora - estabelecimento	MNEST_QP	IDEST	QP 2011
Ramo de atividade económica da entidade empregadora - empresa	CAEEMP_QP	CAEEMP	QP 2011
Ramo de atividade económica da entidade empregadora - estabelecimento	CAEEST_QP	CAEEST	QP 2011
Nº de pessoas ao serviço da entidade empregadora - empresa	NPSEMP_QP	PEMP	QP 2011
Nº de pessoas ao serviço da entidade empregadora - estabelecimento	NPSEST_QP	PEST	QP 2011
Situação do indivíduo na CGA (Enquadramento na CGA)	SITUACAO_CGA	SITUACAO	CGA 2010
Origem (Administração Local / administração central)	ORIGEM_CGA	ORIGEM	CGA 2010
Nome da entidade empregadora	NOME_EMP_CGA	MINISTERIO	CGA 2010
Código postal da entidade empregadora	CP4_EMP_CGA	CP4_SERV	CGA 2010
Código postal auxiliar da entidade empregadora	CP3_EMP_CGA	CP3_SERV	CGA 2010
Habilitações literárias	HABIL_IEFP	HABILITACAO	IEFP 2011
Categoria de inscrição no IEFP	CATEGORIA_IEFP	CATEGORIA	IEFP 2011
Profissão anterior	PROF_ANT_IEFP	CNP_ANTERIOR	IEFP 2011
Ramo de atividade económica anterior	CAE_ANT_IEFP	CAE_ANTERIOR	IEFP 2011

*Tabela resultante do processo de interligação dos ficheiros administrativos

O nível de preenchimento de cada uma das variáveis descritas no quadro acima está naturalmente dependente da presença dos registos individuais nos respetivos ficheiros de origem, pelo que, para além das variáveis demográficas nenhuma das restantes variáveis se encontra preenchida para a totalidade dos registos do protótipo da BPR.

Sintetizando, a informação administrativa disponível atualmente na BPR concorre para 15 variáveis censitárias (12 comunitárias e 3 nacionais), que se apresentam no quadro 19, salvaguardando necessariamente as diferenças conceptuais existentes.

Quadro 19 – Variáveis censitárias para as quais existe informação na BPR, por fonte

	BDIC	SEF	SEG_S	QP	CGA	IEFP	EDU
Local de residência	X	X					
Sexo	X	X					
Idade	X	X					
Estado civil	X	X					
País/local de nascimento	X	X					
Nacionalidade	X	X					
Condição perante o trabalho			X	X	X	X	X
Profissão				X			
Ramo de atividade económica			X	X	X		
Situação na profissão			X	X			
Horas trabalhadas				X			
Nº trabalhadores empresa				X			
Local de trabalho			X	X	X		
Frequência de ensino							X
Nível de ensino completo				X		X	

PONTOS-CHAVE

- Foi construído, pela primeira vez um protótipo de uma base de população residente, assente exclusivamente em informação administrativa e que disponibilizará informação sobre 15 variáveis censitárias.
- A integração e a compatibilização da informação disponibilizada pelas várias fontes administrativas é um processo moroso e complexo, nomeadamente para a situação face ao mercado de trabalho e respetiva caracterização.
- O nível de preenchimento das variáveis depende da presença dos registos nos ficheiros de origem, pelo que a informação disponível não é uniforme para toda a população, nomeadamente ao nível das variáveis socioeconómicas.

2.7. Análise dos principais resultados

A população estimada através de ficheiros administrativos pela Base de População Residente, para o ano de 2011, é de 8 649 135 indivíduos. Este valor resulta da aplicação das regras de residência aos ficheiros da BDIC e do SEF com base nos indícios dados por outras fontes administrativas.

A BPR integrou parte significativa dos registos administrativos disponibilizados pelas diferentes fontes, no entanto não foi possível integrar todos os registos existentes nos ficheiros que traduzem indícios de residência. No quadro 20, apresenta-se o contributo dos ficheiros administrativos para a BPR, por fonte. Ao analisar a percentagem de registos dos diversos ficheiros integrados na BPR verifica-se que o IEFP e a CGA foram as fontes com valores mais elevados, 95,0% e 94,7% respetivamente.

Quadro 20 - Universo dos ficheiros administrativos integrados e não integrados no Protótipo da BPR 2011

Ficheiros Administrativos	Nº de registos do ficheiro	Integrados na BPR 2011	Não integrados na BPR 2011	% de registos integrados na BPR 2011	% de registos não integrados na BPR 2011
BDIC2010	11 565 714	8 182 243	3 383 471	70,7	29,3
BDIC2014	11 884 913	8 549 504	3 335 409	71,9	28,1
SEF2011	434 708	117 428	317 280	27,0	73,0
SS2011	7 209 027	6 277 127	931 900	87,1	12,9
CGA2010	1 103 980	1 045 052	58 928	94,7	5,3
EDUC2011	1 965 842	1 776 623	189 219	90,4	9,6
IEFP2011	702 215	667 373	34 842	95,0	5,0
QP2011	2 736 659	2 360 948	375 711	86,3	13,7
ADSE2012	1 377 353	787 473	589 880	57,2	42,8

Relativamente à BDIC, que na metodologia desenvolvida representa o pilar em que assenta a BPR, 70,7% dos registos individuais da BDIC 2010 foram transpostos para a base de dados de população residente. Para o ano de 2014, esse valor é de 71,9%: este aumento pode estar relacionado com o facto de

algumas crianças não estarem ainda integradas na BDIC 2010 (recorde-se que só a partir de 2007 é que as crianças são registadas à nascença na BDIC).

De salientar ainda a baixa taxa de integração verificada para o ficheiro do SEF, na ordem dos 27%. A interligação deste ficheiro administrativo com as restantes fontes proporcionou resultados baixos, decorrente em certa medida da não existência de um identificador numérico único e estável para a população de nacionalidade estrangeira em Portugal. Por outro lado, verifica-se que uma parte considerável dos indivíduos sinalizados no SEF 2011 estão simultaneamente na BDIC 2010, o que pode estar relacionado com alguma desatualização no ficheiro do SEF para os estrangeiros que entretanto adquiriram nacionalidade portuguesa. Neste âmbito, não se pode também deixar de referir que existem alguns problemas de ligação associados ao ficheiro do SEF.

A diferença entre as taxas de integração da BDIC e do SEF, na BPR, mostra que as especificidades da população estrangeira implicam uma abordagem diferenciada, pelo que se terá que definir uma metodologia capaz de proporcionar melhores resultados para essa população.

No que concerne ao ficheiro da Educação foi possível incluir no protótipo da BPR, 90,4% dos registos individuais constantes nesta fonte administrativa. Numa primeira análise, estes resultados traduzem algumas dificuldades no processo de interligação com o ficheiro da Educação, uma vez que, em princípio, a esmagadora maioria dos indivíduos que estavam a frequentar o sistema de ensino teriam que estar representados essencialmente na BDIC ou no SEF.

Relativamente ao ficheiro da Segurança Social 2011 a taxa de integração na BPR foi de apenas 87,1%, o que significa que para mais de 900 mil registos deste ficheiro, não foi possível estabelecer ligação com a BDIC ou com o SEF (condição para que os registos fossem incluídos no protótipo da BPR). Também no caso da fonte Quadros de Pessoal apenas foi possível integrar na BPR cerca de 86,3% dos registos constantes no respetivo ficheiro.

A taxa de integração de registos verificada para a ADSE, 57,2%, prende-se com alguns constrangimentos encontrados no identificador NIC, pelo que apenas foi possível utilizar os registos que tinham NIF preenchido.

PONTOS-CHAVE

- O facto de não ter sido possível integrar a totalidade dos registos individuais existentes em fontes que traduzem uma efetiva residência (como é o caso dos ficheiros da Educação ou da Segurança Social), e assim aumentar o número de registos a incluir neste primeiro protótipo, constitui uma limitação à metodologia desenvolvida.
- Esta limitação resulta de dificuldades na interligação dos ficheiros, decorrentes da não existência de um identificador único e universal nas diferentes fontes administrativas.

De seguida apresenta-se uma caracterização geral dos registos que integram a BPR 2011.

Cerca de 77% dos registos do protótipo da BPR 2011 têm os 3 principais identificadores numéricos preenchidos

Analisando o grau de preenchimento dos identificadores numéricos dos registos incluídos no protótipo da BPR 2011, verifica-se que 76,6% têm os 3 principais identificadores numéricos preenchidos NIC², NIF e NISS. Com dois identificadores numéricos associados estão 18,3% dos registos (16,2% NIC e NIF e 2,1% NIC e NISS), sendo que para 5,1% dos registos apenas está disponível o NIC.

² Os indivíduos de nacionalidade estrangeira têm número de SEF.

Quadro 21 – Registos do protótipo da BPR 2011, segundo o preenchimento dos identificadores numéricos NIC, NIF e NISS

Identificadores numéricos	Nº de registos	%
NIC, NIF, NISS	6 627 291	76,6
NIC, NIF	1 401 353	16,2
NIC, NISS	183 185	2,1
NIC	437 306	5,1
Total	8 649 135	100,0

Face ao exposto, considera-se que não obstante o enriquecimento efetuado ao nível construção do tríptico de identificadores numéricos, este aspeto continua a ser um fator limitativo no processo de interligação dos registos.

O primeiro protótipo da BPR permitiu apurar, com base exclusivamente em ficheiros administrativos, cerca de 82% da população recenseada nos Censos 2011

A aplicação desta metodologia, para estimar a população residente utilizando informação administrativa, permitiu apurar cerca de 82% do volume populacional dado pelos Censos 2011. Significa isto que o protótipo da BPR 2011 subestima em cerca de 18% a população considerada residente na operação censitária de 2011.

Em particular, a ligação entre a BPR 2011 e os Censos 2011 revela que 92,1% dos indivíduos considerados residentes do ponto de vista administrativo e incluídos neste protótipo estavam representados nos Censos 2011 (7 964 037). Note-se que estes valores decorrem da ligação ao nível do registo individual, pelo que se considera tratar do mesmo indivíduo nas duas bases de dados.

Para os restantes 7,9% de registos da BPR (685 098 indivíduos) que não foi possível encontrar na operação censitária de 2011, podem estar relacionados com:

- Dificuldades no processo de interligação de registos (recorde-se que não foi possível estabelecer qualquer ligação administrativa a cerca de 600 mil registos censitários);
- Integração de indivíduos não recenseados nos Censos 2011 (a taxa de subcobertura para os indivíduos dos Censos 2011 foi de 2,5%, o que significa que ficaram por recensear cerca de 250 mil indivíduos residentes em Portugal).

Registos não incluídos na BPR 2011 associados a indivíduos com potencial de residência: cerca de 1,4 milhões de indivíduos recenseados nos Censos 2011 só estavam representados, ou só foi possível estabelecer ligação à fonte BDIC

Dos indivíduos dos Censos 2011 não incluídos no protótipo da BPR 2011, destaca-se um conjunto considerável de indivíduos que está representado na fonte BDIC, mas que não acumula a presença num outro registo administrativo, pelo que, segundo a metodologia seguida, não têm indícios de residência suficientes em termos administrativos para serem incluídos na BPR.

O cruzamento dos dados dos Censos 2011 com as várias fontes administrativas mostrou que cerca de 1,4 milhões de indivíduos recenseados nos Censos 2011 só estavam representados, ou só foi possível estabelecer ligação à fonte BDIC. Como descrito no ponto relativo à interligação de ficheiros, algumas dessas situações resultam das dificuldades na interligação dos ficheiros, uma vez que dadas as características associadas a alguns indivíduos, no plano teórico, teriam que estar representados noutros ficheiros administrativos, como é o caso da população reformada.

Outra situação relaciona-se com indivíduos que efetivamente não estão representados noutros ficheiros para além da BDIC, como por exemplo as crianças com idade inferior à idade escolar, pessoas que se ocupam das tarefas do lar, indivíduos que desenvolvem a sua atividade económica sem qualquer enquadramento formal, etc. Estas situações colocam limitações reconhecidas à

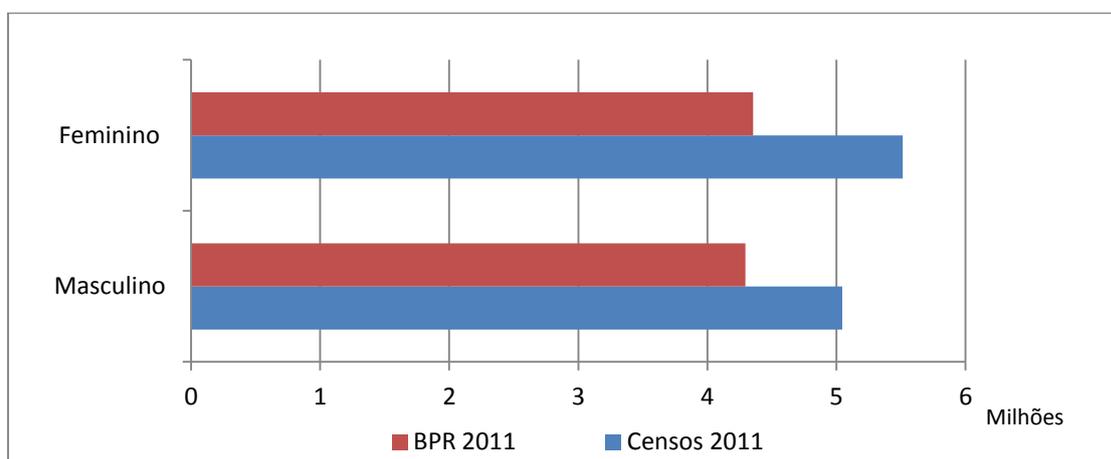
partida na metodologia testada no protótipo agora desenvolvido, pelo que será necessário proceder a ajustes ao modelo de estimação da população a incluir na BPR, nomeadamente no que se refere à atualização da mesma.

O grupo etário das crianças até à idade escolar é aquele que apresenta o maior nível de subestimação na BPR 2011

O protótipo da BPR 2011 estimou um total de 8 649 135 residentes do ponto de vista administrativo, dos quais 4 295 182 são do sexo masculino e 4 353 953 são do sexo feminino (figura 7).

Comparativamente com os Censos 2011, a população da BPR 2011 evidencia uma maior subestimação para as mulheres, 21,1%; nos homens esse valor é de 14,9%. Esta situação traduz-se também na respetiva relação de masculinidade (número de homens por cada 100 mulheres): no protótipo da BPR o valor deste indicador é de 98,7 enquanto na operação censitária de 2011 é de 91,5.

Figura 7 – População residente por sexo, BPR 2011 e Censos 2011



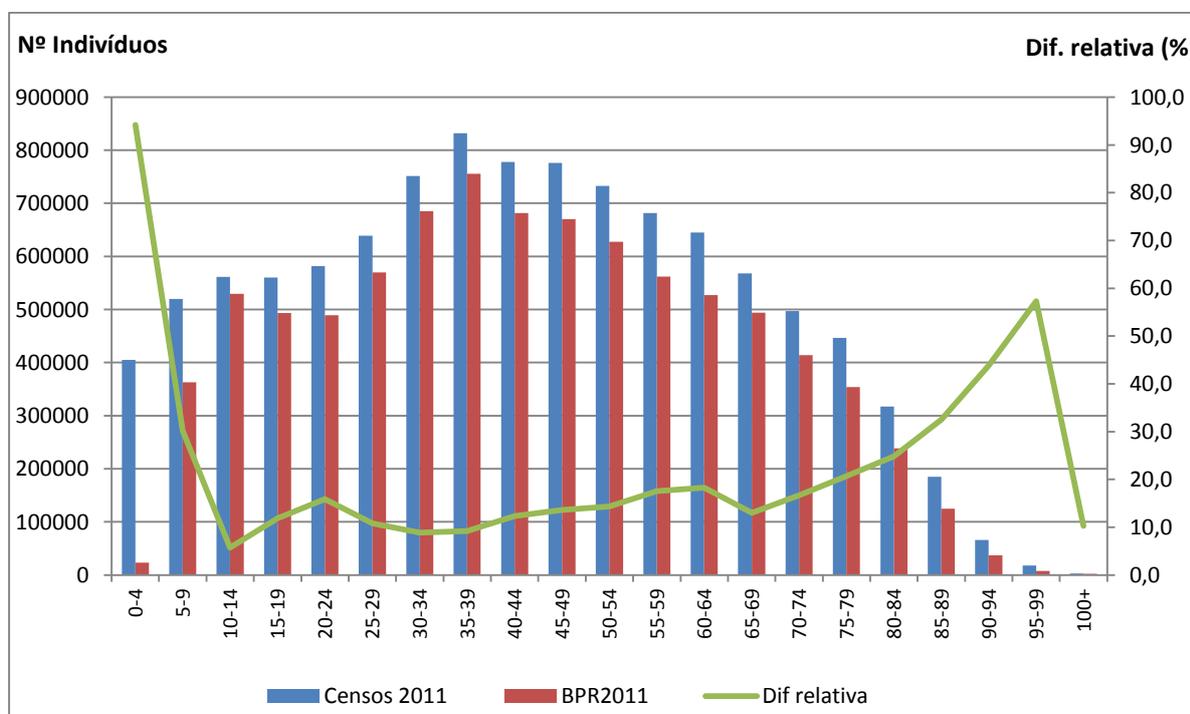
A comparação entre a população do protótipo da BPR 2011 e a recenseada nos Censos 2011 evidencia uma subestimação da primeira em todos os grupos

etários (figura 8). A sub-representação da população na BPR é mais evidente, por um lado, nas idades mais baixas, nomeadamente nas crianças entre 0-4 anos e 5-9 anos e, por outro, nas idades mais elevadas, em particular a partir dos 80 anos.

No que se refere especificamente ao grupo etário 0-4 anos a metodologia utilizada permitiu incluir apenas cerca de 5,8% da população recenseada em 2011, pelo terá que ser desenvolvido um conjunto de novas regras que permita abranger as crianças que ainda não estão representadas nos ficheiros da Educação.

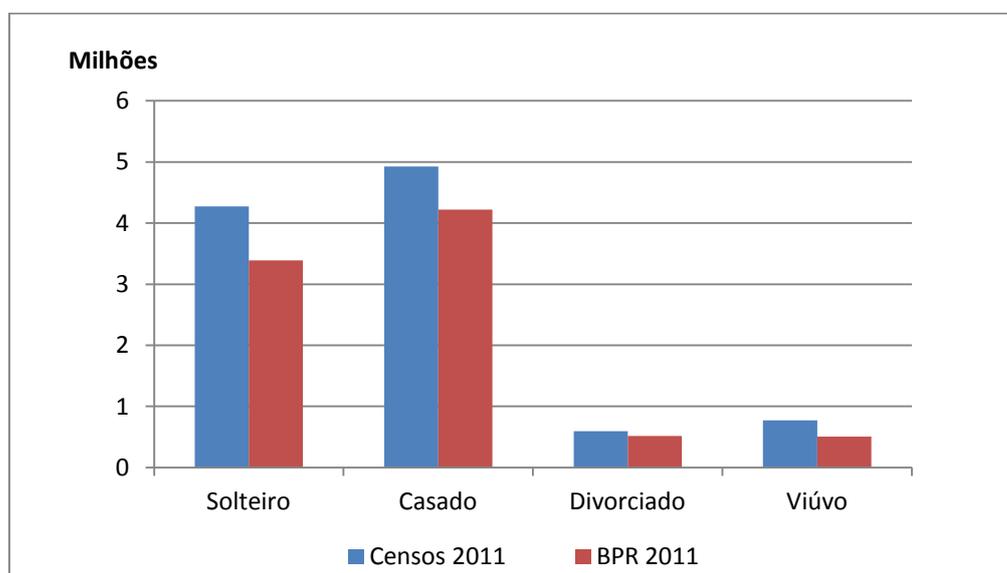
A título meramente ilustrativo, se a metodologia aplicada considerasse todas as crianças constantes na BDIC nascidas entre 2006 e 2011 (i.e., com idade inferior à entrada no ensino básico) e que não foram incluídas neste protótipo, o número de registos da BPR passaria para 9,2 milhões (mais 575 mil registos), o que reduziria a subestimação face aos Censos 2011 em cerca de 6 pp, passando dos atuais 18% para 12%.

Figura 8 – População residente por grupo etário, BPR 2011 e Censos 2011



A distribuição da população no protótipo da BPR 2011, por estado civil (figura 9), revela alguma consistência com os dados dos Censos 2011, apesar de apresentar níveis de subcobertura distintos para as diferentes categorias do estado civil. As modalidades “Viúvo” e “Solteiro” são aquelas onde as diferenças relativas são mais significativas, -34,8% e -20,6%, respetivamente. Esta situação está relacionada com a própria estrutura etária da população incluída no protótipo, nomeadamente o facto de as crianças e os idosos com idades mais elevadas serem os grupos populacionais menos representados no protótipo da BPR 2011.

Figura 9 – População estimada no protótipo de BPR 2011 e população residente nos Censos 2011, por estado civil

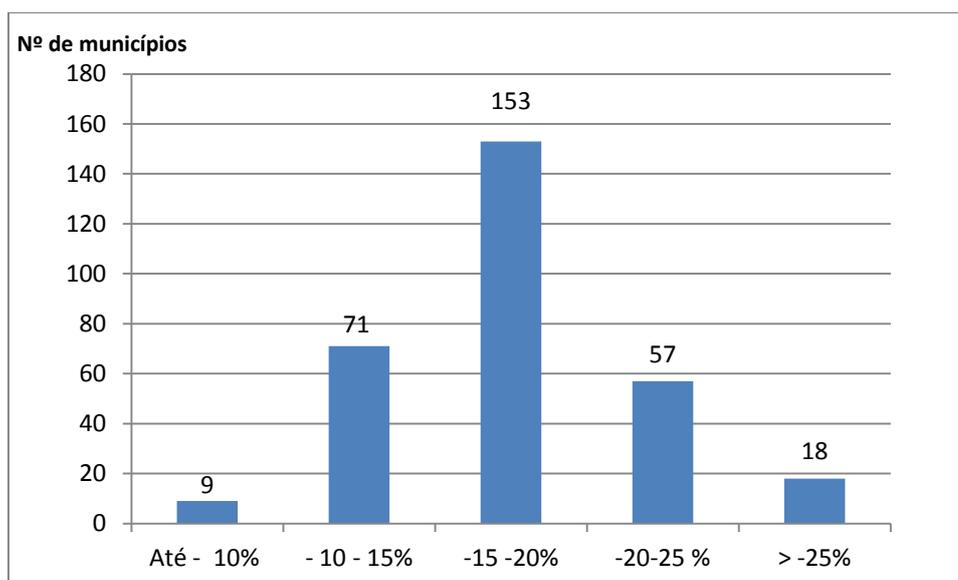


Cerca de metade dos municípios têm população, no protótipo da BPR 2011, subestimada entre os 15 e os 20%

A análise comparativa entre a população estimada no protótipo da BPR 2011 e a apurada nos Censos 2001 mostra que, com exceção do município de Porto Santo na Região Autónoma da Madeira onde a BPR sobrestimou a população, todos os municípios tem um subcobertura, que varia entre - 3,7% no município de Porto Moniz e os - 34,0% no município de Aljezur.

Do total de municípios portugueses cerca de metade (49,7%) tem uma população no protótipo da BPR 2011 subestimada entre os 15 e os 20%, sendo que em cerca de 23% dos municípios as diferenças são na ordem dos 10-15% face à população recenseada em 2011.

Figura 10 – Municípios segundo as diferenças entre a população estimada no protótipo de BPR 2011 e população residente nos Censos 2011

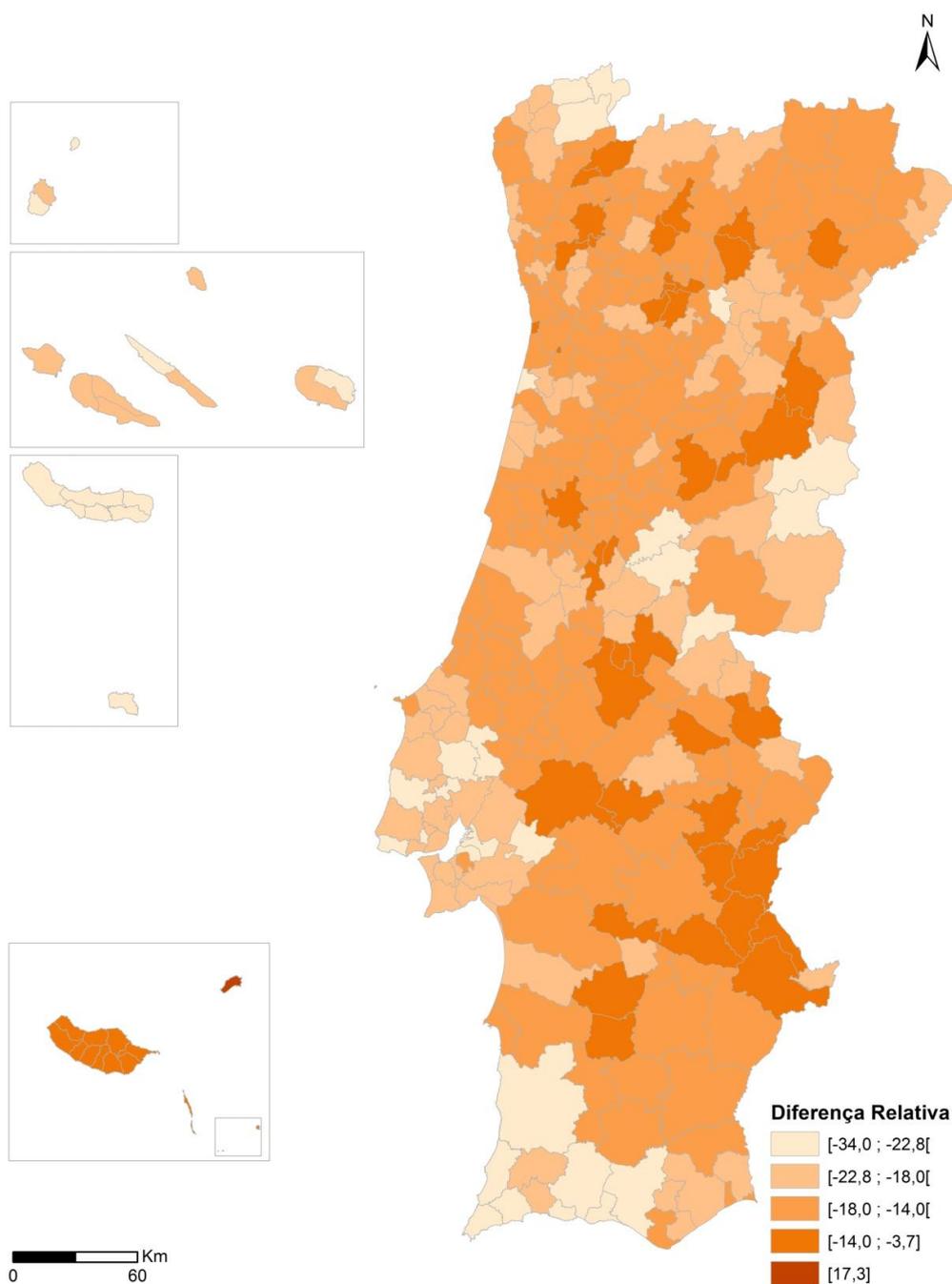


Em termos gerais, é na Região Autónoma da Madeira e no interior da região do Alentejo que se localizam os municípios onde as diferenças entre a população no protótipo da BPR 2011 e os Censos 2011 são mais baixas. Em oposição, as variações mais significativas situam-se em alguns municípios da Região Autónoma dos Açores, do Algarve, Centro interior e na área circundante à Grande Lisboa (figura 11).

Estas diferenças podem ser explicadas, em parte, pela própria informação disponibilizada pelas fontes administrativas. Recorde-se que no caso da Educação não foi possível obter para a Região Autónoma dos Açores dados individualizados para os alunos a frequentar o ensino básico e secundário,

enquanto na Região Autónoma da Madeira foram incluídos também registos de alunos do pré-escolar e de crianças a frequentar creches.

Figura 11 – Diferenças relativas (%) entre a população estimada no protótipo de BPR 2011, e população residente nos Censos 2011, por município



Por outro lado, as diferenças encontradas estão relacionadas com o peso da população estrangeira a nível municipal. Dada a fraca cobertura neste primeiro protótipo da BPR relativamente à população estrangeira (na BPR 2011 apenas 0,6% dos registos dizem respeito a indivíduos de nacionalidade estrangeira enquanto que nos Censos 2011 essa percentagem é de 3,7%) é natural que municípios onde essa população tenha mais peso, apresentem uma sub-representação mais significativa. Por exemplo, Aljezur é o município que apresenta a diferença BPR-Censos mais significativa (-34%) sendo também aquele onde, segundo os Censos 2011, o peso da população estrangeira é o mais elevado (cerca de 22%).

Finalmente, se a análise aos resultados da BPR 2011 for centrada nos 10 municípios mais populosos (quadro 22) verifica-se que as diferenças, entre a população recenseada em 2011 e a estimada pela BPR, se situam entre os -15,3% para o município do Porto e os -24,8% para o município de Cascais. Para Lisboa o protótipo da BPR 2011 subestimou a população em cerca de 100 mil indivíduos, correspondendo a uma diferença de -18,6% face à população recenseada em 2011. No anexo 2, apresentam-se estes indicadores para a totalidade dos municípios.

Quadro 22 – População residente nos Censos 2011 e população estimada no protótipo de BPR 2011 para os 10 municípios mais populosos

Município	Censos 2011	BPR 2011	Diferença absoluta BPR 2011- Censos 2011	Diferença relativa BPR 2011 – Censos 2011
TOTAL	10 562 178	8 649 135	-1 913 043	-18,1
Lisboa	552 700	449 644	- 103 056	-18,6
Sintra	377 835	292 939	- 84 896	-22,5
Vila Nova de Gaia	302 298	252 452	- 49 846	-16,5
Porto	237 591	201 223	- 36 368	-15,3
Cascais	206 479	155 361	- 51 118	-24,8
Loures	199 494	157 304	- 42 190	-21,1
Braga	181 494	152 953	- 28 541	-15,7
Matosinhos	175 478	144 084	- 31 394	-17,9
Amadora	175 136	133 504	-41 632	-23,8
Almada	174 030	138 356	-35 674	-20,5

Em termos de trabalho futuro considera-se, no âmbito da BPR, as seguintes linhas de investigação:

➤ ***Aperfeiçoamento da metodologia de estimação da população residente baseada em ficheiros administrativos***

Por se tratar de um primeiro protótipo há ainda um conjunto de regras por definir e outras que devem ser melhoradas no sentido de dar resposta aos problemas anteriormente identificados, nomeadamente a não inclusão de crianças até à idade de entrada no ensino básico, a baixa representação de população estrangeira, as pessoas que não estão associadas a outras fontes administrativas para além da BDIC, etc.

➤ ***Integração e compatibilização da informação disponibilizada pelas várias fontes***

A integração e a compatibilização da informação disponibilizada pelas várias fontes administrativas é um processo moroso e complexo, nomeadamente no que se refere à situação face ao mercado de trabalho e respetiva caracterização.

➤ ***Melhoria nos processos de interligação ente ficheiros***

O aperfeiçoamento deste protótipo de BPR passa também pela melhoria dos processos de interligação entre ficheiros no sentido de garantir que a ausência de determinadas ligações se devem à não representação dos indivíduos nessas fontes e não decorrem de problemas e dificuldades decorrentes da implementação das técnicas de interligação de ficheiros. Será necessário encontrar na BDIC todos os indivíduos de nacionalidade portuguesa presentes noutros ficheiros administrativos, como por exemplo, na Segurança Social ou na Educação. Este é um processo determinante de forma a validar se a metodologia baseada em indícios de residência, permite, de facto, estimar a população residente em Portugal.

PONTOS-CHAVE

- Os resultados obtidos neste 1º protótipo da BPR evidenciam a complexidade do método para a construção de uma base de população residente com base em ficheiros administrativos.
- Esta linha de investigação é essencial no sentido de encarar uma mudança progressiva do modelo censitário a médio e longo prazo, que poderá permitir alimentar séries censitárias de periodicidade anual que se perspetivam no Eurostat no âmbito do sistema das estatísticas demográficas pós 2021.
- O aperfeiçoamento da metodologia de estimação da população residente baseada em ficheiros administrativos é um exercício exigente, complexo e moroso, com potencial de melhoria e que necessita de investigação e estudo aprofundado.
- A melhoria deste protótipo de BPR passa também pela própria melhoria dos processos de interligação entre ficheiros, no sentido de garantir que os indivíduos são localizados em todos os ficheiros onde de facto estão representados.
- Estes primeiros resultados permitem já, tirar algumas conclusões, ajustar pressupostos e afinar a metodologia; considerando as fontes disponíveis, estes resultados devem ser articulados com os objetivos do Teste de 2016 e a preparação da BPR 2016.

3. Comparação de microdados entre Ficheiros Administrativos e Inquéritos Estatísticos

3.1. Enquadramento

Pretende-se comparar, para um conjunto de variáveis, os dados recolhidos através dos Censos 2011 com a informação proveniente de ficheiros administrativos, ao nível dos registos individuais. O objectivo principal desse estudo é avaliar a consistência entre a informação administrativa disponível nas diversas fontes e os dados recolhidos nos Censos. Adicionalmente considera-se que os resultados deste estudo podem contribuir para validar o processo de qualidade das ligações entre registos bem como os conteúdos dos registos ligados por fonte.

O universo objeto de análise é assim constituído pelos registos individuais que foi possível ligar, ou fizeram *matching*, entre os Censos 2011 e os ficheiros administrativos disponíveis.

Finalmente, para validar os resultados, no sentido de ter um referencial para esses resultados, em particular no que se refere a um conjunto selecionado de variáveis que caracterizam o mercado de trabalho e a escolaridade, faz-se também a comparação entre os microdados dos Censos 2011 e do IE 2011.

3.2. Comparação de microdados entre os Censos 2011 e os ficheiros administrativos

3.2.1. Constituição dos universos para a realização do estudo

Neste exercício, comparam-se os registos que estão associados ao mesmo indivíduo. Por esse motivo, para cada variável, o número de registos a comparar é função do número de ligações encontradas no processo de interligação das bases de dados ou *matching*. Esses registos estão associados a 9 949 599 de indivíduos³ e representam cerca de 94% da população recenseada em Portugal

³ Este valor pode diferir do apresentado noutros pontos deste relatório, pois o trabalho desenvolvido não decorreu exatamente em simultâneo com todas as fases da construção da BPR.

em 2011 que foi possível ligar através de várias chaves de atributos construídas a partir de variáveis demográficas. Das variáveis demográficas, selecionámos, para este exercício, a freguesia e o município de residência, o sexo, a data de nascimento, o estado civil, o município e país de naturalidade e o país de nacionalidade.

No que se refere à variável condição perante a atividade económica considera-se, em particular, a população residente com 15 ou mais anos, 8 989 849 indivíduos. Para as restantes variáveis relacionadas com o mercado de trabalho, o universo é constituído pela população residente com 15 ou mais anos, com condição perante a atividade económica (em sentido restrito) ativa empregada, num total de 4 361 187 indivíduos.

Quanto às variáveis relativas a questões do ensino, excluem-se do universo de referência os indivíduos que ainda não tinham nascido em 31 de Dezembro de 2009, num total de 117 085 indivíduos.

As contagens dos registos iguais e diferentes foram feitas por variável, proveniente em simultâneo dos diversos pares de fontes: Censos 2011 – ficheiro administrativo. Sempre que possível foi usada a fonte administrativa com informação relativa a 2011; quando esta não estava disponível utilizou-se a do ano mais próximo.

A seleção das variáveis para a realização deste exercício, baseou-se no contributo da informação das diversas fontes administrativas, tendo como foco a eventual substituição dos dados recolhidos pelos Censos.

Foram usados os ficheiros provenientes de várias fontes administrativas, designadamente:

- **Instituto dos Registos e do Notariado (IRN 2010)**

O IRN disponibilizou, relativamente a 2010, a Base de Dados de Identificação Civil (BDIC) que contém os registos dos cidadãos de nacionalidade portuguesa residentes em Portugal, num total de 11 565 714 indivíduos.

- **Serviço de Estrangeiros e Fronteiras (SEF 2011)**

O SEF disponibilizou, para 2011, a base de dados dos indivíduos de nacionalidade estrangeira com título de residência válido, num total de 434 708 indivíduos; no que se refere à condição perante a atividade económica, foram considerados os registos relativos a pessoas com 15 ou mais anos, 379 965 indivíduos, e quanto à Profissão os registos com código de situação na profissão correspondente a “ativo”, 124 721 indivíduos.

- **Instituto de Segurança Social (ISS 2011)**

O ISS disponibilizou, para 2011, a base de dados dos indivíduos que têm uma qualificação ativa, no sentido de um enquadramento no sistema de Segurança Social, num total de 7 209 027 indivíduos; no que se refere à condição perante a atividade económica, foram considerados os registos relativos a pessoas com 15 ou mais anos, 7 066 838 indivíduos, e quanto às restantes variáveis do mercado de trabalho os registos com códigos de tipo de qualificação correspondentes a ativos, 4 107 425 indivíduos.

- **Quadros de Pessoal (QP 2011)**

O Gabinete de Estratégia e Planeamento do Ministério da Solidariedade, Emprego e Segurança Social disponibilizou, para 2011, a base de dados dos indivíduos empregados com contrato individual de trabalho em empresas com pelo menos uma pessoa ao serviço e abrangidas pelo Código de Trabalho, num total de 2 736 659 indivíduos.

- **Instituto de Emprego e Formação Profissional (IEFP 2011)**

O IEFP disponibilizou, para 2011, a base de dados dos indivíduos desempregados inscritos no Centro de Emprego, num total de 702 215 indivíduos. No que se refere à Profissão, foram considerados os registos preenchidos relativamente à profissão anterior, 629 748 indivíduos.

- **Direção Geral de Estatísticas da Educação e Ciência (DGEEC 2011)**

A DGEEC disponibilizou, para o ano letivo de 2011/2012, a base de dados dos alunos matriculados nos diferentes graus ou níveis de ensino, num total de 1 965 842 indivíduos.

- **Caixa Geral de Aposentações (CGA 2010)**

A CGA disponibilizou, para 2010, a base de dados dos funcionários públicos ou equiparados, reformados da função pública e pensionistas de sobrevivência de funcionários públicos ou equiparados, num total de 1 103 980 indivíduos.

3.2.2. Seleção das variáveis a comparar e metodologia

Considerando os ficheiros administrativos disponíveis, foram selecionadas 18 variáveis de interesse censitário: 8 de natureza demográfica e 10 de natureza socioeconómica. Nos quadros 23 e 24, apresentam-se as variáveis, bem como uma síntese do contributo de cada fonte para o processo de comparação com os microdados dos Censos 2011.

Quadro 23 – Contributo, em termos de variáveis, das várias fontes administrativas nas comparações com os microdados dos Censos 2011

BDIC 2010	SEF 2011	SS 2011	QP 2011	IEFP 2011	DGEEC 2011	CGA 2010
<ul style="list-style-type: none"> • FR residência • MN residência • Sexo • Data nascimento • Estado civil • MN naturalidade • País nacionalidade 	<ul style="list-style-type: none"> • País nacionalidade • País naturalidade • Condição perante atividade económica • Profissão 	<ul style="list-style-type: none"> • Condição perante atividade económica • Local trabalho • Situação profissão 	<ul style="list-style-type: none"> • Local trabalho • Profissão • Ramos atividade económica empresa • Ramo atividade económica estabelecimento • Situação profissão 	<ul style="list-style-type: none"> • Condição perante atividade económica • Profissão anterior • Ramo atividade económica anterior 	<ul style="list-style-type: none"> • Frequência de ensino • Nível de ensino 	<ul style="list-style-type: none"> • Condição perante atividade económica

Quadro 24 – Variáveis a comparar entre os Censos 2011 e os Ficheiros Administrativos

Designação da variável	Variável administrativa (designação na fonte administrativa)	Ficheiro de origem
Freguesia de residência	FR_RESID_COD	BDIC 2010
Município de residência	MN_RESID_COD	BDIC 2010
Sexo	SEXO	BDIC 2010
Data de nascimento	DTNASC	BDIC 2010
Estado civil	ESTCIVIL	BDIC 2010
Município de naturalidade	MN_NAT_COD	BDIC 2010
País de naturalidade	PAIS_NAT_COD	BDIC 2010
		SEF 2011
País de nacionalidade	PAIS_NAC_COD	BDIC 2010
		SEF 2011
Condição perante a atividade económica	TIPO_QLF	ISS 2011
	TIPO_OCUPACAO	SEF 2011
	SITUACAO	CGA 2010
	CATEGORIA	IEFP 2011
Local de trabalho (Município)	COD_CONCELHO_EE	ISS 2011
	IDEST (município estabelecimento)	QP 2011
Profissão (comparação a 4, 3, 2 e 1 dígito)	PROF	QP 2011
	CNP_ANTERIOR	IEFP 2011
	PROFISSAO_CNP	SEF 2011
Ramo de atividade económica (comparação a: Letra ->4, Grupo ->3, Divisão ->2 e Secção ->1 dígito)	CAEEMP (CAE Empresa)	QP 2011
	CAEEST (CAE estabelecimento)	QP 2011
	CAE_ANTERIOR	IEFP 2011
Situação na profissão	SITPRO	QP 2011
	TIPO_QLF	ISS 2011
Número de trabalhadores da empresa	PEMP	QP 2011
Número de horas de trabalho	PNT	QP 2011
Nível de ensino completo	HABIL	QP 2011
Frequência de ensino	A10_ENSINO	DGEEC 2011
Nível de ensino a frequentar	A10_CICLO	DGEEC 2011

Em termos metodológicos, consideram-se os registos dos Censos 2011 que fizeram *matching* com os ficheiros administrativos, os quais têm um número de indivíduo associado (atribuído nos Censos) que faz a ligação às variáveis de interesse dos microdados dos Censos 2011.

Partindo dessa tabela inicial de *matching*, fez-se a ligação a cada um dos ficheiros administrativos que continha a variável de interesse a comparar, através do NIC encriptado associado a cada registo.

No caso das variáveis provenientes de fontes administrativas cuja informação relativamente ao atributo em comparação com os microdados dos Censos 2011 não tem uma correspondência direta de fatores, recorreu-se a ligações adicionais e/ou a tabelas de codificação para o efeito de harmonização desses campos.

Para o efeito, foi necessário garantir a harmonização entre as categorias de cada variável, afim de estas poderem ser efetivamente comparáveis. Isto é, no caso dos ficheiros administrativos, por exemplo, não obstante o trabalho de normalização prévio, já efetuado aquando da receção dos ficheiros disponibilizados pelas entidades, para as variáveis que não foram utilizadas nas chaves de *matching* (caso das variáveis não demográficas), foi necessária uma recodificação já que, com frequência, a informação disponível nas diversas fontes, para um mesmo atributo que nos Censos, não está compilada da mesma forma, isto é, não tem os mesmos fatores. Nesta situação estão, em particular, as variáveis relacionadas com a condição perante a atividade económica. No caso da fonte Segurança Social, apesar de não existir uma variável com informação explícita neste campo, foi construída uma tabela que permite aferir, ainda que com limitações, a condição perante a atividade económica através do tipo de qualificação.

PONTOS-CHAVE

- Os resultados da comparação entre os microdados dos Censos 2011 e dos ficheiros administrativos permitem validar diretamente o conteúdo dos ficheiros administrativos e indiretamente a qualidade das ligações entre os ficheiros, independentemente do grau de cobertura dos universos associados às variáveis censitárias selecionadas para as quais há informação administrativa.
- Os resultados da comparação entre os microdados dos Censos 2011 e dos ficheiros administrativos permitem encarar, num cenário futuro, a utilização de informação administrativa em complemento ou substituição de informação estatística, no sentido de um uso integrado dessa informação.

3.3. Comparação de microdados entre os Censos 2011 e o Inquérito ao Emprego

2.3.1. Seleção de variáveis e metodologia

A comparação dos microdados dos Censos 2011 e do Inquérito ao Emprego de 2011 desenvolveu-se de modo análogo ao comparativo Censos – Ficheiros Administrativos, descrito em 2.2.. A metodologia diferiu apenas no que se refere à comparação entre as respetivas variáveis dos dois inquéritos (neste caso diferiu também o número de variáveis: foram selecionadas 5 variáveis de natureza demográfica e 8 de natureza socioeconómica).

Assim, com a finalidade de comparar os microdados dos Censos 2011 e do Inquérito ao Emprego (IE) 2011, foram selecionados os dados deste último inquérito referentes ao primeiro trimestre de 2011. O conjunto é constituído por 39 884 registos e um vasto conjunto de variáveis das quais foram selecionadas as variáveis demográficas:

- Nome
- Data de nascimento
- Sexo
- Estado civil
- Freguesia do alojamento

e as variáveis socioeconómicas:

- Condição perante o trabalho
- Profissão principal
- Classificação atividade económica
- Situação na profissão principal
- Escalão do número de trabalhadores ao serviço
- Escalão de horas semanais habitualmente trabalhadas no emprego principal
- Frequência de ensino
- Nível de ensino completo.

Para ser possível fazer a comparação entre os resultados dos dois inquéritos, foi necessário efetuar a ligação entre os dois conjuntos de registos (na prática pretende-se localizar os registos do IE nos Censos e verificar se o preenchimento das variáveis correspondentes, nas duas fontes, é igual, isto é, se um certo indivíduo tem as mesmas características nos Censos e IE).

Para o efeito, foi construída uma chave para fazer a ligação entre os registos das duas fontes, constituída pelos seguintes campos:

- Sexo (código numérico)
- Primeiras três letras do primeiro Nome
- Últimas três letras do último Nome
- Data de nascimento (formato: *aaaammdd*).
- Estado civil (código numérico)
- Freguesia de alojamento (código numérico)

No final do processo foram ligados 17 732 registos com correspondência única (que representam 44% dos 39 884 obtidos via IE). Foram esses os registos objeto de comparação por variável. As variáveis Sexo, Nome (3 primeiros caracteres do primeiro nome e três últimos do último), Data de nascimento, Estado civil e Freguesia de alojamento integraram a chave de ligação e por isso têm uma correspondência de 100% entre as duas fontes.

De referir que previamente foi necessário normalizar os nomes do IE:

- foram removidos todos os acentos e caracteres especiais
- todas as 'M^a' passaram a 'MAR'
- todas as letras minúsculas passaram a maiúsculas

Após a limpeza e normalização dos nomes foram criadas dois novos campos, construídos a partir do campo NOME, com as três primeiras letras do primeiro nome e as três últimas letras do último nome.

Os códigos numéricos dos restantes campos e os formatos das datas do IE foram normalizados de acordo com os utilizados nos Censos. Foi também necessário ajustar alguns dos escalões de variáveis cujos valores são intervalares, de modo a permitir uma comparação direta entre as duas fontes.

PONTOS-CHAVE

- Os resultados do comparativo Censos – IE confirmam a consistência do comparativo Censos – ficheiros, se as taxas de correspondência (não obstante a diferente cobertura dos universos associados às variáveis seleccionadas) forem convergentes.

3.4. Síntese de resultados e avaliação da qualidade das fontes

Apresenta-se, no quadro 25, uma síntese dos resultados obtidos neste exercício de comparação de microdados. Para as variáveis em que foi possível desagregar os dados e fazer o exercício de comparação por categoria ou nível, os valores apresentados dizem apenas respeito ao nível mais agregado de informação.

Na primeira coluna listam-se as variáveis objeto de estudo pela ordem correspondente à apresentação dos resultados. Nas colunas centrais (2^a a 4^a coluna) apresentam-se os números dos universos em comparação e o número de

registos individuais comparados, que correspondem aos pares de registos que foi possível ligar entre Censos/ficheiro administrativo.

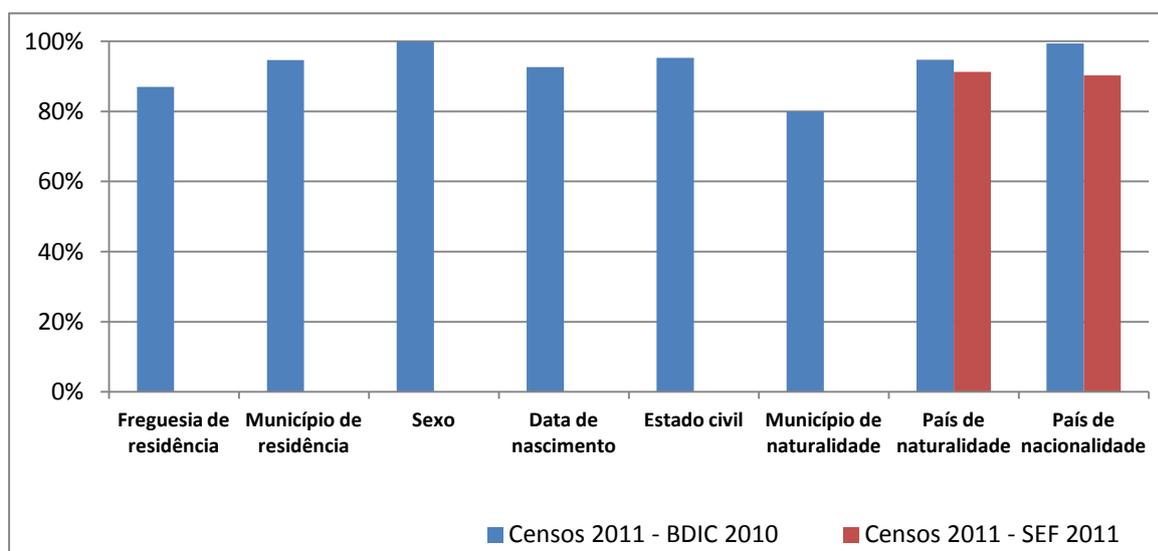
Em complemento à informação do comparativo entre Censos/ficheiros administrativos, acrescenta-se também a informação sobre o comparativo entre os Censos e o Inquérito ao Emprego para uma seleção de variáveis comuns aos dois questionários. Nas duas últimas colunas da direita apresentam-se as percentagens de igualdade obtidas no comparativo. Na coluna mais à direita, apresentam-se ainda os valores do principal indicador no que se refere à avaliação dos erros de conteúdo entre os dados dos Censos 2011 e o Inquérito de Qualidade dos Censos 2011, o Índice de consistência global, ICG (conforme descrito na publicação do INE *Inquérito de Qualidade dos Censos 2011 – Metodologia e resultados* de 2013). Estes valores são apresentados a título de referência essencialmente para enquadrar a ordem de grandeza dos resultados agora obtidos (para além do universo objeto de comparação ser diferente também alguns dos conceitos não são os mesmos, ainda que próximos, como é o caso da: residência habitual^a; grupo etário quinquenal^b; âmbito geográfico agregado^c).

Quadro 25 – Quadro Síntese dos resultados obtidos na comparação entre os Microdados dos Censos 2011 e outras fontes

Variável	Nº indivíduos Censos	Nº indivíduos outras fontes		Nº de registos comparados	% Igualdade Censos/ outra fonte	ICG (%)
Freguesia de residência	10 562 178	BDIC 2010	11 565 714	9 308 384	87,0	97,7^a
Município de residência	10 562 178	IBDIC 2010	11 565 714	9 308 384	94,6	97,7
Sexo	10 562 178	BDIC 2010	11 565 714	9 308 384	99,9	99,0
Data de nascimento	10 562 178	BDIC 2010	11 565 714	9 308 384	92,6	95,7^b
Estado civil	10 562 178	BDIC 2010	11 565 714	9 308 384	95,3	97,4
Município de naturalidade	10 562 178	BDIC 2010	11 565 714	9 308 384	80,0	84,0
País de naturalidade	10 562 178	BDIC 2010	11 565 714	9 308 384	94,7	84,0^c
		SEF 2011	434 708	107 136	91,3	84,0^c
País de nacionalidade	10 562 178	BDIC 2010	11 565 714	9 308 384	99,4	97,8
		SEF 2011	434 708	107 136	90,3	97,8
Condição Perante a atividade económica	8 989 849	ISS 2011	7 066 838	4 910 073	81,2	
		SEF 2011	379 965	107 136	27,1	
		CGA 2010	1 103 980	716 264	92,1	
		IEFP 2011	702 215	454 479	42,1	
		IE 2011	39 884	17 732	84,3	
Local de trabalho (Município)	4 361 187	ISS 2011	4 107 425	2 788 758	56,6	77,6
		QP 2011	2 736 659	2 045 476	81,6	77,6
Profissão	4 361 187	QP 2011	2 736 659	2 045 476	61,9	
		IEFP 2011 (anterior)	629 748	27 590	30,9	
		SEF 2011	124 721	171 370	52,9	
		IE 2011	39 884	6 995	67,8	
Ramo de atividade económica	4 361 187	QP 2011 (estabelecimento)	2 736 659	2 045 476	74,1	
		QP 2011 (empresa)	2 736 659	2 045 476	73,8	
		IEFP 2011	702 215	171 370	49,2	
		IE 2011	39 884	6 995	77,6	
Situação na profissão	4 361 187	QP 2011	2 736 659	2 045 476	93,0	82,2
		ISS 2011	4 107 425	2 788 758	85,5	82,2
		IE 2011	39 884	6 995	86,5	
Número de trabalhadores da empresa	4 361 187	QP 2011	2 736 659	2 045 476	54,4	51,6
		IE 2011	39 884	6 995	60,6	
Número de horas de trabalho	4 361 187	QP 2011	2 736 659	2 045 476	56,8	
		IE 2011	39 884	6 995	72,6	
Nível de ensino completo	10 445 093	QP 2011	2 736 659	2 210 930	59,5	
		IE 2011	39 884	15 494	80,2	
Frequência de ensino	10 445 093	DGEEC 2011	1 965 842	15 494	87,4	
		IE 2011	39 884	17 732	86,5	
Nível de ensino a frequentar	10 445 093	DGEEC 2011	1 965 842	1 359 916	82,2	69,8

Considerando os resultados apresentados no quadro 25, e no que se refere às variáveis demográficas, a comparação dos microdados dos Censos com a BDIC 2010 e com o SEF 2011 mostra um elevado nível de concordância para as variáveis data de nascimento, sexo, país de nacionalidade, país de naturalidade e estado civil, entre os 90 e os 99%. Também a comparação, através da BDIC 2010, das variáveis relativas à residência obtiveram taxas de igualdade ou correspondência bastante elevadas: cerca de 87% ao nível da freguesia de residência e 95% quando se agrega ao nível do município de residência. Estes resultados são ilustrados na figura 12.

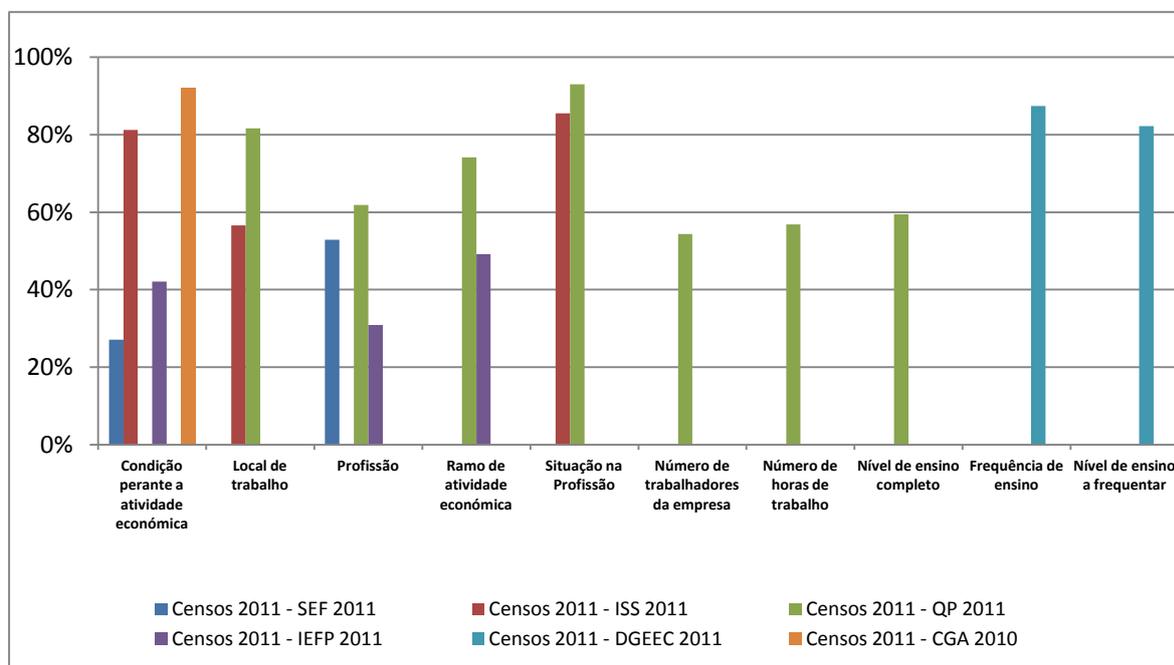
Figura 12 – Percentagem de igualdade dos registos ligados entre os Censos 2011 e os ficheiros administrativos, por variável demográfica e por fonte



Relativamente às variáveis de natureza socioeconómica, os resultados são mais heterogéneos. Porém, para a maior parte das variáveis / fontes, os resultados alcançados indicam que, também nesta área, se justifica prosseguir e aprofundar o trabalho descrito no presente relatório. Os resultados, para o conjunto de variáveis socioeconómicas selecionadas, são ilustrados, por variável e por fonte, na figura 13. A sua análise permite identificar três situações:

- Taxas de correspondência elevadas, por variável, obtidas na comparação com todas as fontes: caso da situação na profissão, com cerca de 86% de correspondência via ISS 2011 e 93% via QP 2011;
- Taxas de correspondência que variam por fonte: caso das variáveis profissão, ramo de atividade económica e condição perante a atividade económica (no caso desta última, obteve-se cerca de 92% de correspondência via CGA 2010 enquanto que considerando o IEFP 2011 esse valor diminuiu para os 42%);
- Taxas de correspondência com origem numa única fonte, medianas com valores que rondam os 50%: caso do número de trabalhadores da empresa e número de horas de trabalho, a elevadas (ultrapassam os 80%): caso da Frequência de ensino ou nível de ensino a frequentar.

Figura 14 – Percentagem de igualdade dos registos ligados entre os Censos 2011 e os ficheiros administrativos, por variável socioeconómica e por fonte



As diferenças conceptuais entre as fontes administrativas e os Censos 2011, assim como algum desfasamento temporal nos correspondentes períodos de referência nos vários ficheiros, podem estar na base das divergências encontradas na comparação das variáveis socioeconómicas. A este propósito refira-se ainda que, quando foi possível fazer uma análise mais pormenorizada (para níveis de informação mais desagregados), constatou-se que existe uma relação de vizinhança, entre categorias ou escalão de resposta nos Censos e a informação administrativa.

Por último refira-se o enquadramento dos resultados do comparativo Censos – ficheiros administrativos face ao comparativo Censos/IE e também Censos/ICG. No quadro 26, faz-se um paralelismo desses resultados considerando as percentagens de igualdade obtidas do comparativo Censos/ficheiros administrativos através da fonte mais representativa face ao universo dos Censos ou com a percentagem de igualdade mais elevada. Os valores obtidos são muito próximos, o que se considera aumentar a consistência global dos resultados.

Quadro 26 – Percentagem de igualdade Censos-IE e Censos-ficheiros administrativos

Variável	% Igualdade Censos-IE	% Igualdade Censos-ficheiros administrativos
Condição perante a atividade económica	84,3	81,2 (ISS 2011)
Profissão	67,8	61,9 (QP 2011)
Ramo de atividade económica	77,6	74,1 (QP 2011)
Situação na profissão	86,5	93,0 (QP 2011)
Número de trabalhadores da empresa	60,6	54,4 (QP 2011)
Número de horas de trabalho	72,6	56,8 (QP 2011)
Nível de ensino	80,2	59,5 (QP 2011)
Frequência de ensino	86,5	87,4 (DGEEC 2011)

No que se refere ao ICG, medida do Inquérito de Qualidade dos Censos 2011, vem acrescentar uma dimensão adicional à análise e enquadrar também os resultados.

No quadro 27 apresenta-se, para uma seleção de variáveis, o paralelismo entre o comparativo Censos/ficheiros administrativos e ICG. Os valores obtidos são muito próximos, pelo que também estes resultados contribuem para aumentar a consistência e robustez dos resultados do exercício de comparação Censos/ficheiros administrativos.

Quadro 27 – Percentagem de igualdade Censos-IE e ICG

Variável	ICG (%)	% Igualdade Censos-ficheiros administrativos
Freguesia de residência	97,7	87,0 (BDIC 2010)
Município de residência	97,7	94,6 (BDIC 2010)
Sexo	99,0	99,9 (BDIC 2010)
Data de nascimento	95,7	92,6 (BDIC 2010)
Estado civil	97,4	95,3 (BDIC 2010)
Município de naturalidade	84,0	80,0 (BDIC 2010)
País de naturalidade	84,0	94,7 (BDIC 2010)
	84,0	91,3 (SEF 2011)
País de nacionalidade	97,8	99,4 (BDIC 2010)
	97,8	90,3 (SEF 2011)
Local de trabalho	77,6	56,6 (ISS 2011)
	77,6	81,6 (QP 2011)
Situação na profissão	82,2	93,0 (QP 2011)
	82,2	85,5 (ISS 2011)
Número de trabalhadores da empresa	51,6	54,4 (QP 2011)

PONTOS-CHAVE

- Os resultados agora obtidos apresentam uma elevada consistência entre a informação administrativa e os microdados dos Censos 2011.
- Foram comparadas, entre os Censos 2011 e os ficheiros administrativos, 8 variáveis demográficas. Estas variáveis, utilizadas nas chaves de ligação entre ficheiros, apresentam taxas de correspondência muito elevadas – superiores a 90% – ao nível dos microdados dos indivíduos.
- Das 10 variáveis socioeconómicas comparadas, entre os Censos 2011 e os ficheiros administrativos, obtiveram-se valores de taxas de correspondência superiores a 80% para algumas variáveis relativas ao mercado de trabalho.
- Em termos de consistência de fontes, e não obstante a questão de cobertura de universos, a informação provenientes dos Quadros de Pessoal, apresenta a maior taxa de correspondência para o conjunto de variáveis disponíveis.
- A análise conjunta dos resultados dos comparativos Censos/Ficheiros, Censos/IE e também os resultados do ICG do IQ dos Censos 2011, que revelou uma convergência a partir dessas três dimensões de análise, permitiu validar a qualidade da informação dos ficheiros e também a própria qualidade das ligações, não obstante a questão da cobertura dos universos.
- Em termos de trabalho futuro, de modo a reunir informação complementar aos resultados agora obtidos, que possibilite uma avaliação mais rigorosa do potencial de substituição das variáveis dos Censos por informação dos ficheiros, equaciona-se como fundamental o estudo dos universos e complementaridade entre fontes em termos de cobertura.

4. Desenho para os Censos 2021

Em cada década o modelo de inquérito censitário tem vindo a incorporar fatores de modernização que têm constituído desafios metodológicos e operacionais bem-sucedidos. Foi assim quando em 1970 os Censos introduziram o autopreenchimento; em 1991 aquando da introdução de um sistema de codificação automática de expressões alfabéticas; em 2001 os questionários foram tratados através de leitura ótica e reconhecimento e interpretação de caracteres (OCR/ICR) e foi também assim quando em 2011 foi utilizada pela primeira vez a internet para recolha da informação.

Neste documento, são apresentadas as linhas gerais do modelo para os Censos 2021, com particular ênfase para as áreas de inovação. São igualmente identificados os riscos das alterações propostas. O desenho apresentado reflete a experiência dos Censos 2011, integra a avaliação do que são as tendências e as melhores práticas a nível internacional e toma em consideração o desenvolvimento tecnológico e as mudanças na sociedade que irão ocorrer nos próximos anos. É importante sublinhar que o desenvolvimento da estratégia apresentada neste documento, será objeto de aprofundamento no quadro de um programa de testes, o primeiro dos quais terá lugar em 2016.

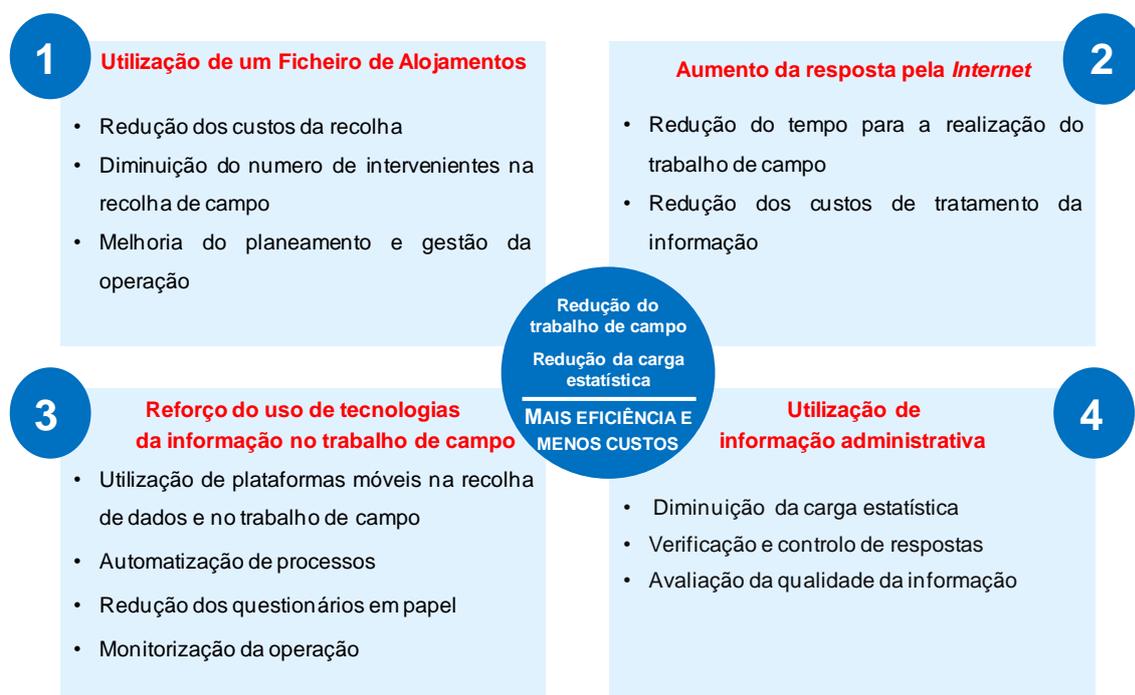
4.1. Desenho para os Censos 2021: a caminho de um censo digital

O desenho para os Censos 2021 deverá assentar em quatro áreas chave, de acordo com o diagrama abaixo (figura 15). Os objetivos visados são:

- Assegurar que os resultados tenham padrões de qualidade elevados;
- Reduzir custos e aumentar a eficiência da operação;
- Facilitar a resposta e reduzir a carga estatística sobre os cidadãos.

De seguida, descrevem-se resumidamente as características de cada uma dessas áreas de intervenção.

Figura 15 – Quatro áreas chave para os Censos 2021



4.1.1. Utilização de um ficheiro de alojamentos: alteração do modelo de distribuição de questionários

A utilização de um ficheiro de alojamentos vai permitir alterar o modelo de distribuição de questionários porta-a-porta pelo envio postal de uma carta com os códigos de acesso para resposta pela internet.

Esta mudança representa a alteração mais significativa no desenho do novo modelo para os Censos 2021.

Portugal nunca alterou a metodologia de distribuição dos questionários porta-a-porta. Este processo sempre foi assegurado por recenseadores que identificavam localmente todos os edifícios e alojamentos, constituindo uma lista – ficheiro de alojamentos –, no momento da distribuição. A qualidade da série censitária está muito associada a este processo, que garante elevada qualidade, em particular na contagem da população e dos alojamentos – cobertura do censo.

Os custos e a logística associados a este modelo (24 500 pessoas em campo, 75% das quais na função de recenseador nos Censos 2011 em Portugal) são muito elevados. O pagamento do trabalho de campo dos Censos 2011 representou cerca de 55% do orçamento global da operação.

Portugal, no contexto da UE, foi um dos poucos países que nos Censos 2011 ainda manteve o modelo de distribuição de questionários porta-porta-porta.

As condicionantes à utilização do ficheiro de alojamentos

Com base nos dados recolhidos nos Censos 2011, o INE constituiu o Ficheiro Nacional de Alojamentos (FNA). O ficheiro a utilizar nos Censos 2021 será baseado no FNA.

O conhecimento sobre a qualidade do FNA deverá permitir antecipadamente a construção de um mapa de risco para o todo o território tendo em vista a identificação de estratégias de recolha diferenciadas.

Este fator, em associação com o grau de adesão da população à resposta pela Internet, é um elemento importante para se determinar o número de recenseadores a envolver nos Censos 2021. Deste número dependerá também a dimensão da estrutura de enquadramento e gestão operacional da operação e respetivos custos.

A qualidade do FNA, para efeitos dos Censos, envolve três componentes:

- Exaustividade
- Qualidade dos endereços
- Qualidade do atributo “Forma de ocupação do alojamento”

O FNA tem como base de partida o momento censitário (março de 2011). A atualização do FNA (construções novas / demolições) através do Sistema de Indicadores das Operações Urbanísticas (SIOU) deverá, em grande parte,

garantir a exaustividade pretendida para a lista de alojamentos utilizada no arranque da futura operação censitária.

O FNA poderá ter qualidade em termos de exaustividade mas, ainda assim, subsistirem problemas que afetem os Censos. Dado que se pretende utilizar maioritariamente a via postal, a qualidade do endereço é fundamental, para se evitarem injustificadas devoluções de correspondência.

Um atributo no qual, naturalmente, é inevitável alguma desatualização num ficheiro deste tipo é o que se refere à situação / forma de ocupação do alojamento. As características do parque habitacional em Portugal têm forte influência na gestão desta questão. De acordo com os Censos 2011, eram de residência habitual apenas 68% dos 5,9 milhões de alojamentos do país. Existiam então cerca de 1,9 milhões de alojamentos vagos ou de residência secundária.

Esta estrutura do *stock* de alojamentos, muito diferente da existente noutros países que já utilizam um ficheiro de alojamentos no inquérito censitário, aconselha uma atenção muito especial ao atributo “Forma de ocupação”.

O FNA atualizado vai possibilitar ganhos de eficiência na realização do inquérito e em particular no trabalho de campo. O FNA deve apresentar a melhor qualidade possível, tendo em conta patamares mínimos que não ofereçam risco para a realização da operação e para a qualidade dos resultados.

4.1.2. Internet como principal canal de resposta aos Censos 2021

A utilização da internet como canal privilegiado de comunicação por parte dos cidadãos deverá continuar a crescer. Os serviços do Estado têm acelerado esse movimento promovendo a disponibilização de cada vez mais serviços com recurso à internet. Os casos do IRS, do e-facturas e mais recentemente dos contratos de arrendamento, são exemplos que impulsionaram o uso da internet na nossa sociedade.

Tendo em conta este contexto e os resultados alcançados nos Censos 2011 (50% de resposta pela internet), os Censos 2021 deverão encarar a internet como principal canal de resposta.

4.1.3. Reforço das tecnologias de informação e comunicação no trabalho de campo

Tendencialmente todo o trabalho de campo será assegurado através de dispositivos móveis com ligação à internet: *Smartphone*, *tablets*, computadores portáteis.

A estratégia de reforço do uso das tecnologias de informação baseia-se no conceito BYOD (*Bring Your Own Device*) e implica que toda a estrutura deverá ter os seus próprios equipamentos. Esta estratégia coloca em evidência as aplicações de suporte à operação e as respetivas condições de acesso.

O desenvolvimento de aplicações para várias plataformas, que garantam funcionalidade, segurança e rapidez, representam novos e complexos desafios. Por sua vez, reduzir ao mínimo o volume de papel em circulação e os procedimentos manuais concentra responsabilidades acrescidas nos sistemas informáticos de suporte à recolha.

Em síntese, os riscos associados a um Censo digital devem ser bem avaliados e os sistemas informáticos devidamente testados. O período de recolha dos Censos é relativamente curto, cerca de 1,5 meses. Qualquer falha compromete a realização da operação.

4.1.4. Utilização de informação administrativa

A continuidade dos estudos em curso e designadamente o aperfeiçoamento metodológico do exercício de construção de protótipos da BPR irá permitir avaliar o potencial de utilização desta base no médio e longo prazo.

Paralelamente, o teste de 2016 irá possibilitar uma melhor compreensão do potencial de utilização de uma base desta natureza no apoio à realização dos Censos 2021. Um dos benefícios que a BPR poderá representar é o de permitir identificar os alojamentos ocupados como residência principal e contribuir de forma muito significativa para direcionar o trabalho de campo e diminuir o custo da operação.

O conhecimento aproximado das contagens da população e das suas características possibilita o desenho de indicadores de alerta para monitorização e acompanhamento da recolha.

A utilização de informação administrativa para tratamento de não respostas ou no desenho do controlo de qualidade são igualmente áreas que podem vir a ser exploradas.

4.2. Principais alterações e riscos face aos Censos 2011

O quadro seguinte, apresenta as principais alterações dos Censos 2021 face aos Censos 2011.

Quadro 28 – Principais diferenças entre os Censos 2011 e modelo Censos 2021

Fator	Censos 2011	Modelo para Censos 2021
Utilização de ficheiro de alojamentos	x	v
Distribuição e recolha de questionários porta a porta	v	x
Resposta pela Internet	v	v
Utilização de ficheiros administrativos	x	v
Carga estatística para o respondente	+	-
Complexidade e volume associado ao tratamento dos questionários	+	-
Custos com a operação	+	-

De modo a completar a informação do quadro 28, sintetizam-se a seguir alguns dos riscos associados a três dos pilares em que assentam as mudanças do modelo para 2021, face a 2011 (a dimensão de utilização dos ficheiros administrativos não se considera aqui por ter sido amplamente explorada nos capítulos 2 e 3).

FNA

- Gestão e tratamento de um volume excessivo de cartas devolvidas, por falta de qualidade dos endereços dos alojamentos;
- Taxas de resposta baixas, por deficiente classificação dos alojamentos quanto à forma de ocupação;
- Aumento de contactos através da Linha de Apoio;
- Aumento dos custos;
- Dificuldades e complexidade na operacionalização do trabalho de campo;
- Qualidade dos resultados.

Equipamentos móveis e sistemas aplicacionais

- Segurança dos sistemas;
- Funcionamento dos sistemas (dificuldade de acesso).

Estrutura de apoio ao respondente e às equipas de terreno

- Dimensionamento da estrutura de modo a assegurar um serviço eficiente e de qualidade.

O Teste de 2016 tem o objetivo de avaliar a viabilidade do modelo agora descrito em linhas gerais para 2021, nomeadamente identificar os pontos críticos associados às mudanças face a 2011. O documento metodológico do Teste 2016 será oportunamente apresentado para certificação técnica.

5. Conclusões e trabalho futuro

A transformação do modelo censitário para Portugal assenta numa estratégia de médio e longo prazo e numa estratégia de curto prazo.

A utilização da informação administrativa para substituição do atual modelo censitário, insere-se na visão de médio e longo prazo, tendo em conta a natureza estrutural dos obstáculos identificados:

- A BDIC não é uma base de população residente;
- Portugal não dispõe de um número único para cada cidadão;
- A informação administrativa não permite responder integralmente às necessidades censitárias. Das 27 variáveis obrigatórias em termos comunitários, os ficheiros administrativos permitem responder a 12 variáveis com uma cobertura total, e a 6 com uma cobertura parcial, devido a limitações dos universos de observação. Para algumas áreas, como a família, não existe qualquer informação disponível. Das variáveis de âmbito nacional os ficheiros administrativos não permitem obter informação para 33 das 50 variáveis disponibilizadas nos Censos 2011.

O estudo levado a cabo possibilitou a construção de um primeiro protótipo de Base de População Residente (BPR) a partir de ficheiros administrativos.

O desenvolvimento da metodologia de construção da BPR teve como objetivo ultrapassar a limitação da BDIC não ser um ficheiro de população residente, e como tal, incluir cerca de 1 milhão de indivíduos não residentes em Portugal, embora com uma morada em território nacional.

A BPR é uma base de dados de indivíduos considerados residentes em Portugal (mediante a aplicação de regras que indiciam residência) e das suas características, que resultou da interligação de diferentes ficheiros administrativos, e que reúne a informação disponibilizada pelas várias fontes relevantes para o apuramento de variáveis demográficas e socioeconómicas de cariz censitário.

Esta base de dados é constituída por cerca de 8,6 milhões de registos, que no seu conjunto permitirão o apuramento parcial ou total de 15 variáveis censitárias. O primeiro protótipo da BPR permite caracterizar cerca de 82% da população residente em Portugal (Censos 2011), prevendo-se o seu aperfeiçoamento, nomeadamente com o objetivo de ser possível ultrapassar algumas das limitações encontradas e permitir a sua atualização anual.

O modelo tem limitações no que se refere aos métodos de interligação das bases de dados administrativas, sendo que é necessário ainda avaliar, se o modelo de estimação da população pode ser baseado na metodologia de indícios de residência.

O estudo envolveu uma avaliação sobre a qualidade dos resultados, através do grau de coincidência entre a informação recolhida pelos Censos 2011 e as fontes administrativas. A comparação dos microdados entre os Censos 2011 e os ficheiros administrativos revelou taxas de coincidência elevadas, em particular para as variáveis de natureza demográfica.

A BPR pode vir a constituir um importante repositório de informação demográfica e social, atualizável por via administrativa, instrumento estatístico de longo alcance, que se associa mas ultrapassa a preparação dos Censos 2021. O desenvolvimento deste exercício e o caminho já percorrido são determinantes para encarar uma mudança progressiva do modelo censitário e poderá permitir alimentar séries censitárias de periodicidade anual, perspetiva já em debate no Eurostat.

Em face dos resultados até agora obtidos, o desenho proposto para os Censos 2021 assenta na realização de um inquérito, que inclui quatro áreas de inovação:

- Utilização de um ficheiro de alojamentos;
- Aumento da resposta pela Internet;
- Reforço do uso de tecnologias de informação no trabalho de campo;
- Utilização de informação administrativa, como suporte à realização da operação.

Em termos de futuro, continuarão a ser desenvolvidas duas áreas de trabalho.

Por um lado, na perspetiva de curto prazo, a realização de um Teste em 2016 para avaliar o desenho metodológico dos Censos 2021.

Por outro lado, a estratégia de médio e longo prazo deve continuar a ser objeto de investigação e desenvolvimento, nomeadamente:

- O aprofundamento do diálogo com a CNPD, no sentido de criar condições para acesso aos ficheiros com um enquadramento legal menos restritivo;
- A articulação com as entidades administrativas, dando conhecimento dos avanços até agora conseguidos neste projeto, e também das limitações associadas a cada fonte. O objetivo é o de formalizar as relações de cooperação que permitam o envio regular e periódico da informação necessária;
- A continuação do estudo prevê o aprofundamento dos métodos de interligação de ficheiros e da metodologia de estimação da população residente.

6. Bibliografia de referência

- Census Bureau. 2015. *Changing Times, Changing Methods, Changing Technologies: Innovations and Testing for the US 2020 Population and Housing Census*. Conference of European Statistics, Group of Experts on Population and Housing Censuses. Geneva.
- Central Statistical Bureau 2015. *Innovations preparing for the Population and Housing Census 2021 in Latvia* Conference of European Statistics, Group of Experts on Population and Housing Censuses. Geneva.
- Citro, C. 2014. “From multiple modes for surveys to multiple data sources for estimates”. Statistics Canada Catalogue no. 12-001-X. *Survey Methodology*. Vol. 40, Nº2, pp. 137-161.
- Instituto Nacional de Estatística (INE). 2014. *Estudo de viabilidade da utilização de dados de fontes administrativas no novo modelo censitário para 2021*. Relatório QUAR, Gabinete dos Censos 2021.
- Instituto Nacional de Estatística (INE). 2014. *Avaliação dos modelos censitários utilizados noutros países e a sua adequabilidade a Portugal*. Documento interno, Gabinete dos Censos 2021.
- Instituto Nacional de Estatística (INE). 2015. *Interligação das diferentes bases de dados provenientes de fontes administrativas, no âmbito do novo modelo censitário para 2021*. Relatório QUAR, Gabinete dos Censos 2021.
- Office for National Statistics (ONS). 2013. *Beyond 2011 Producing Population Estimates Using Administrative Data: Methods and policies (M6)*. ONS website.
- Office for National Statistics (ONS). 2013. *Beyond 2011 Producing Population Estimates Using Administrative Data: In Practice (M7)*. ONS website.
- Office for National Statistics (ONS). 2014. *The Census and Future Provision of Population Statistics in England and Wales: Recommendation from the National Statistician and Chief Executive of the UK Statistics Authority*. ONS website.

- Office for National Statistics (ONS). 2015. *Census Design Document*. Office for National Statistics. ONS website.
- Office for National Statistics (ONS). 2015. *Research for 2021 Census England and Wales: possible innovations under consideration*. Conference of European Statistics, Group of Experts on Population and Housing Censuses. Geneva.
- Royce, D. 2011. *Preliminary Report on Methodology Options for the 2016 Census*. Technical Report, 60p.
- Statistics Austria. 2012. *Challenges in the transition from traditional to register-based Census in Austria*. Conference of European Statistics, Sixtieth Plenary Session. Paris.
- Statistics Canada. 2015. *Plans for the 2016 Canadian Census, related to tests*. Conference of European Statistics, Group of Experts on Population and Housing Censuses. Geneva.
- Statistics New Zealand. 2013. *Evaluation of administrative data sources for subnational population estimates*. Statistics New Zealand.
- Statistics New Zealand. 2014. *An overview of progress on the potencial use of administrative data for census information in New Zealand – Census Transformation programme 2013*. Statistics New Zealand.
- Statistical Office of the Republic of Serbia. 2015. *2021 Population and housing census in the Republic of Serbia – Moving forward*. Conference of European Statistics, Group of Experts on Population and Housing Censuses. Geneva.
- United Nations Economic Commission for Europe (UNECE). 2011. *Using administrative and secondary sources for Official Statistics*. United Nations Publications.
- United Nations Economic Commission for Europe (UNECE). 2013. *Measuring population and housing – Practices of UNECE countries in the 2010 round of censuses*. United Nations Publications.

Anexo 1: Fichas técnicas da informação administrativa, por fonte

ADSE – Direção Geral de Proteção Social aos Trabalhadores em F P

Metainformação	
Anos disponíveis e datas de receção	Universo
<p>13.02.2015 13.02.2015</p>	<p>Beneficiários da ADSE (Titulares - trabalhadores no ativo ou aposentados - e Familiares ou equiparados) nomeadamente: os trabalhadores com relação jurídica de emprego público da administração central, regional e local, inscritos na Caixa Geral de Aposentações ou na Segurança Social, e não beneficiem de outro subsistema de saúde integrado na Administração Pública; o pessoal docente do ensino particular e cooperativo, desde celebrado acordo com a ADSE; os aposentados que não sejam abrangidos por qualquer outro subsistema de saúde integrado na Administração Pública; os dependentes dos titulares: cônjuges, descendentes e ascendentes ou equiparados.</p>
Âmbito geográfico	Variáveis disponíveis
Portugal	Nº ADSE, NIC, NIF, NCGA, SEXO, NOME, DATA DE NASCIMENTO, ESTADO CIVIL, PARENTESCO, PAIS DISTRITO, MUNICÍPIO, FREGUESIA DE RESIDÊNCIA, LOCALIDADE, CÓDIGO POSTAL, DATA DE INSCRIÇÃO, SERVIÇO, MOVIMENTO DESPESAS REEMBOLSO: REGIME LIVRE E CONVENCIONADO.
Unidade de observação	
Indivíduo	

Total de registos		Registos únicos	
2012	2013	2012	2013
1 377 353	1 290 784	1 377 353	1 290 784

Taxas de matching	
ADSE2012	<i>Sem matching</i> (problemas no NIC encriptado em tratamento)
ADSE2013	<i>Sem matching</i> (problemas no NIC encriptado em tratamento)

Indicadores de qualidade			
Nome da variável	Designação da variável	Taxa de preenchimento (%)	
		2012	2013
NUB_ADSE	Número de identificação da ADSE	100,0	100,0
NIC	Número de identificação civil	67,5	66,5
NIF	Número de identificação fiscal	64,8	65,3
NCGA	Número de CGA	-	39,5
SEXO	Sexo	100,0	100,0
DTNASC	Data de nascimento	100,0	100,0
NOME	Nome	100,0	100,0
PAIS_RES_COD	País de residência	100,0	100,0
DT_RESID_COD	Distrito de residência	99,9	99,9
MN_RESID_COD	Município de residência	99,5	99,5
FR_RESID_COD	Freguesia de residência	37,6	34,6
MORADA_LOCALIDADE	Localidade	100,0	100,0
CP4	Código postal	100,0	100,0
CP3	Código postal auxiliar	85,5	86,7
ESTCIVIL	Estado civil	100,0	100,0
PARENTESCO	Situação parentesco na ADSE (Titular, filho, cônjuge, ascendente,...)	100,0	100,0
DT_INS_ADSE	Data de inscrição na ADSE	100,0	100,0
NOME_SERVIÇO	Nome serviço/entidade	100,0	100,0
MOVDESP_RL2012	Movimento de despesas Reembolso 2012 (Regime livre)	100,0	100,0
MOVDESP_RC2012	Movimento de despesas Reembolso 2012 (Regime convencionado)	100,0	100,0

CGA – Caixa Geral de Aposentações

Metainformação	
Anos disponíveis e datas de receção 	Universo Indivíduos subscritores (funcionários públicos e agentes administrativos civis e militares da Administração Pública Central, Local e Regional, professores do ensino particular e cooperativo e trabalhadores de algumas empresas públicas e sociedades anónimas de capitais públicos), os reformados da função pública e os pensionistas de sobrevivência de funcionários públicos ou equiparados.
Âmbito geográfico Portugal	Variáveis disponíveis ANO, NUMERO DE SUBSCRITOR, NIC, NIF, NOME, SEXO, DATA DE NASCIMENTO, CÓDIGO POSTAL, DISTRITO E MUNICÍPIO DE RESIDÊNCIA, SITUAÇÃO NA CGA, ORIGEM, MINISTÉRIO, SERVIÇO, MORADA DO SERVIÇO
Unidade de observação Indivíduo	

Total de registos		Registos únicos	
2010	2013	2010	2013
1 130 790	1 085 872	1 103 980	1 080 116

Taxas de matching	
	BDIC NIC
CGA 2010	94,8%
CGA 2013	96,1%

Indicadores de qualidade			
Nome da variável	Designação da variável	Taxa de preenchimento (%)	
		2010	2013
ANO	Ano	100,0%	100,0%
NUMERO_SUBSCR	Número de subscritor	100,0%	100,0%
NOME	Nome	100,0%	100,0%
BI_00	Bilhete de identidade	71,4%	77,0%
NIF	Número de identificação fiscal	84,6%	81,9%
CP4	Código postal	50,8%	57,6%
CP3	Código postal auxiliar	50,8%	57,6%
DT_RESID_COD	Código do Distrito de Residência	49,8%	56,6%
MN_RESID_COD	Código do Município de Residência	49,8%	56,6%
SEXO	Sexo	100,0%	100,0%
DTNASC	Data de Nascimento	100,0%	100,0%
SITUACAO	Situação	100,0%	100,0%
ORIGEM	Origem	81,2%	79,2%
MINISTERIO	Ministério	81,2%	79,2%
SERVICO	Serviço	81,9%	82,3%
RUA_SERV	Rua do Serviço	81,4%	81,8%
LOCALIDADE_SERV	Localidade do Serviço	39,0%	37,3%
CP4_SERV	Código de postal do Serviço	81,9%	82,3%
CP3_SERV	Código postal auxiliar do Serviço	81,9%	82,3%

Educação – Direção Geral de Estatísticas da Educação e Ciência/ Direção de Educação da Madeira

Metainformação	
Anos disponíveis e datas de receção	Universo
<p>25.03.2014</p> <p style="text-align: center;">↓</p> <p>2011/12</p> <hr style="border: 0.5px solid #4F81BD;"/> <p style="text-align: center;">2013/14</p> <p style="text-align: center;">→</p> <p style="text-align: center;">17.12.2014 23.01.2015</p> <p style="text-align: center;">↓</p>	<p>Alunos matriculados nos diferentes graus e modalidades de ensino (ensino básico, secundário e superior). Os níveis de ensino creche e pré-escolar apenas estão disponíveis para a Madeira. Foram enviados em 2011, 7 ficheiros distintos e em 2013, 5 tabelas distintas (4 do Continente e 1 da Madeira)</p>
Unidade de observação	Variáveis disponíveis
<p>Portugal continental e R.A.Madeira</p> <p>Indivíduo</p>	<p>ANO E ANO LETIVO, NIC, NISS, TIPO E NÚMERO DO DOCUMENTO DE IDENTIFICAÇÃO, NOME, SEXO, DATA DE NASCIMENTO, NACIONALIDADE, CÓDIGO POSTAL, DISTRITO, MUNICÍPIO E FREGUESIA DE RESIDÊNCIA, CICLO E MODALIDADE DE ENSINO, CURSO, ANO E SITUAÇÃO, MUNICÍPIO E FREGUESIA DA ESCOLA</p>
Total de registos	
<p>2011 *</p> <p>1 965 842</p>	<p>2013</p> <p>1 706 779</p>
Taxas de matching	
<p>DGEEC2011</p> <p>DGEEC/DEM2013</p>	<p>BDIC NIC</p> <p>88,0%</p> <p>98,3%</p>

**Educação – Direção Geral de Estatísticas da Educação e Ciência/
Direção de Educação da Madeira**

Indicadores de qualidade			
Nome da variável	Designação da variável	Taxa de preenchimento (%)	
		2011	2013
ANO	Ano	100,0%	100,0%
A0_ANOLETIVO	Ano letivo	100,0%	100,0%
NOME	Nome	100,0%	100,0%
A2_TIPODOCID	Tipo de documento de identificação	98,4%	100,0%
A2_NUMDOCID	Número de identificação	98,6%	100,0%
A2_NUMDOCID_CHKDIGIT	Número de controlo do documento de identificação	-	64,7%
NISS	Número de identificação da Segurança Social	-	62,5%
PAIS_NAC_COD	País de Nacionalidade	99,9%	100,0%
DT_NASC	Data de nascimento	100,0%	100,0%
SEXO	Sexo	100,0%	100,0%
CP4	Código postal	73,4%	66,1%
CP3	Código postal auxiliar	58,7%	64,7%
FR_RESID_COD	Freguesia de residência	65,1%	71,6%
MN_RESID_COD	Município de residência	84,6%	92,0%
DT_RESID_COD	Distrito de residência	84,9%	92,0%
A10_PLANOESTUDOS	Plano de Estudos	-	75,9%
A10_ENSINO	Nível de Ensino	100,0%	100,0%
A10_CICLO	Ciclo de Ensino	71,7%	98,8%
A12_MODALIDADE	Modalidade de Ensino	78,3%	78,4%
A12_CURSO	Curso	47,7%	97,6%
A13_ANOTIPO	Ano de Ensino	90,4%	78,6%
A16_SITUACAO	Situação de aprovação	93,8%	78,9%
A20_CODFREGUESIAESC OLA	Freguesia da Escola	2,8%	78,9%
A20_CODCONCELHOESC OLA	Município da Escola	2,8%	78,9%

Quadros de Pessoal – Gabinete de Estratégia e Planeamento (MSESS)

Metainformação							
Anos disponíveis e datas de receção				Universo			
				<p>Pessoas empregadas com contrato individual de trabalho, em empresas com pelo menos uma pessoa ao serviço e abrangidas pelo código do trabalho (não estão incluídos os trabalhadores por conta própria, os trabalhadores do serviço doméstico, os trabalhadores da administração pública, dos institutos públicos e os que possuem regime de contrato de trabalho em funções públicas).</p>			
Âmbito geográfico				Variáveis disponíveis			
Portugal				<p>ANO, NISS, NOME, SEXO, DATA DE NASCIMENTO, NACIONALIDADE, REGIME DE REFORMA APLICADO, DATA DE ADMISSÃO, ANTIGUIDADE NA EMPRESA, CATEGORIA PROFISSIONAL, SITUAÇÃO NA PROFISSÃO, TIPO DE CONTRATO, REGIME DE DURAÇÃO DO TRABALHO, HABILITAÇÕES LITERÁRIAS, NÍVEL DE QUALIFICAÇÃO, PROFISSÃO, PERÍODO DE TRABALHO SEMANAL, NIF EMPRESA, MORADA DA EMPRESA (DISTRITO, MUNICÍPIO, FREGUESIA, LOCALIDADE E CÓDIGO POSTAL) NATUREZA JURÍDICA, ATIVIDADE ECONÓMICA DA EMPRESA, NÚMERO DE PESSOAS A O SERVIÇO EMPRESA, MORADA DO ESTABELECIMENTO (DISTRITO, MUNICÍPIO, FREGUESIA, LOCALIDADE E CÓDIGO POSTAL) ATIVIDADE ECONÓMICA DO ESTABELECIMENTO, NÚMERO DE PESSOAS A O SERVIÇO ESTABELECIMENTO.</p>			
Unidade de observação							
Indivíduo							
Total de registos				Registos únicos			
2010	2011	2012	2013	2010	2011	2012	2013
2 843 391	2 798 081	2 617 252	2 610 344	2 780 487	2 736 659	2 562 816	2 554 641
Taxas de matching							
				SS NISS			
QP2010				94,8%			
QP2011				96,9%			
QP2012				96,1%			
QP2013				98,1%			

Quadros de Pessoal – Gabinete de Estratégia e Planeamento (MSESS)

Indicadores de qualidade					
Nome da variável	Designação da variável	Taxa de preenchimento (%)			
		2010	2011	2012	2013
NISS	Número de identificação da Segurança Social	98,8%	98,9%	98,9%	99,0%
NOME	Nome	100,0%	100,0%	100,0%	100,0%
REG_REFORMA	Regime de reforma aplicado	100,0%	100,0%	100,0%	100,0%
PAIS_NAC_COD	País de Nacionalidade	100,0%	100,0%	100,0%	100,0%
SEXO	Sexo	100,0%	100,0%	100,0%	100,0%
DTNASC	Data de nascimento	99,8%	99,9%	99,9%	99,9%
DT_ADM	Data de Admissão	100,0%	100,0%	100,0%	100,0%
ANTIG	Antiguidade na empresa	100,0%	100,0%	100,0%	100,0%
CTPRO	Categoria Profissional	100,0%	100,0%	100,0%	100,0%
SITPRO	Situação na profissão	100,0%	100,0%	100,0%	100,0%
TIPO_CONTR	Tipo de contrato	93,8%	93,6%	93,5%	93,5%
REG_DUR	Regime de duração do trabalho	93,8%	93,6%	93,5%	93,5%
HABIL	Habilitações literárias	100,0%	100,0%	100,0%	100,0%
NQUAL1	Nível de qualificações	100,0%	100,0%	100,0%	100,0%
PROF	Profissão - Classificação portuguesa de profissões	100,0%	100,0%	100,0%	100,0%
PNT	Período normal de trabalho semanal	100,0%	93,6%	93,5%	93,5%
NIF	Número de identificação fiscal da empresa	100,0%	100,0%	100,0%	100,0%
MORADA_EMPRESA	Morada da empresa	100,0%	100,0%	100,0%	100,0%
LOCAL_EMPRESA	Localidade da empresa	100,0%	100,0%	100,0%	100,0%
CP4	Código postal	100,0%	100,0%	100,0%	100,0%
CP3	Código postal auxiliar	100,0%	100,0%	100,0%	100,0%
DESCRICAO_POSTAL	Localidade postal da empresa	100,0%	99,8%	100,0%	100,0%
NATJU	Natureza jurídica	100,0%	100,0%	100,0%	100,0%
DTEMP	Distrito da empresa	100,0%	100,0%	100,0%	100,0%
IDEMP	Município da empresa	100,0%	100,0%	100,0%	100,0%
FREGEMP	Freguesia da empresa	100,0%	100,0%	100,0%	100,0%
CAEMP	Atividade económica da empresa	100,0%	100,0%	100,0%	100,0%
PEMP	Número de pessoas ao serviço na empresa (31 outubro)	100,0%	100,0%	100,0%	100,0%
MORADA_ESTAB	Morada do estabelecimento	100,0%	100,0%	100,0%	100,0%
LOCAL_ESTAB	Localidade do estabelecimento	100,0%	100,0%	100,0%	100,0%
CP4_ESTAB	Código postal do estabelecimento	100,0%	100,0%	100,0%	100,0%
CP3_ESTAB	Código postal auxiliar do estabelecimento	100,0%	100,0%	100,0%	100,0%
DESC_POST_ESTAB	Localidade postal do estabelecimento	99,9%	99,8%	100,0%	100,0%
DTEST	Distrito do estabelecimento	100,0%	100,0%	100,0%	100,0%
IDEST	Município do estabelecimento	100,0%	100,0%	100,0%	100,0%
FREGEST	Freguesia do estabelecimento	100,0%	100,0%	100,0%	100,0%
CAEST	Atividade económica do estabelecimento	100,0%	100,0%	100,0%	100,0%
PEST	Número de pessoas ao serviço do estabelecimento (31 outubro)	100,0%	100,0%	100,0%	100,0%

IEFP/IEM – Instituto de Emprego e Formação Profissional / Instituto de Emprego da Madeira

Metainformação					
Anos disponíveis e datas de receção			Universo		
23.12.2014 (IEM) 05.01.2015 (IEFP)			Indivíduos desempregados inscritos no centro de emprego		
Âmbito geográfico			Variáveis disponíveis		
Portugal continental e Região Autónoma da Madeira			ANO, ANO_MES, BI, NIF, NISS, NOME, LOCALIDADE, NÚMERO DE UTENTE, CODIGO POSTAL, TIPO DOCUMENTO IDENTIFICAÇÃO, DISTRITO, MUNICIPIO E FREGUESIA DE RESIDÊNCIA, SEXO, DATA DE NASCIMENTO E ESTADO CIVIL, PAIS NACIONALIDADE, HABILITACAO E AREA CURSO, CATEGORIA, DATA DESEMPREGO, DATA E MOTIVO DE INSCRICAO, PROFISSÃO E ATIVIDADE ECONÓMICA, TEMPO INSCRICAO E DE NÃO INSCRIÇÃO		
Unidade de observação					
Indivíduo					
Total de registos			Registos únicos		
2011	2012	2013	2011	2012	2013
702 215	853 817	897 258	702 215	853 817	897 258
Taxas de matching					
			SS NIF	SS NISS	BDIC NIC
		IEFP e IEM 2011	71,6%	70,0%	72,8%
		IEFP e IEM 2012	72,8%	72,1%	77,2%
		IEFP e IEM 2013	67,1%	67,0%	83,5%

IEFP/IEM – Instituto de Emprego e Formação Profissional / Instituto de Emprego da Madeira

Indicadores de qualidade				
Nome da variável	Designação da variável	Taxa de preenchimento (%)		
		2011	2012	2013
ANO	Ano	100,0%	100,0%	100,0%
ANO_MES	Ano e mês	96,5%	96,6%	96,8%
BI	Número de identificação civil	100,0%	100,0%	100,0%
NIF	Número de identificação fiscal	97,9%	98,8%	99,3%
NISS	Número de identificação Segurança Social	97,2%	98,1%	98,6%
NOME	Nome	100,0%	100,0%	100,0%
LOCALIDADE	Localidade	88,7%	87,6%	86,9%
UTENTE_ID	Número de utente	3,5%	3,4%	3,2%
CP4	Código postal	100,0%	100,0%	100,0%
CP3	Código postal auxiliar	96,8%	97,6%	97,9%
DCPOSTAL	Designação do Código Postal	100,0%	100,0%	100,0%
TIPO_DOC_ID	Tipo de documento de identificação	100,0%	100,0%	100,0%
DT_RESID_COD	Distrito de Residência	100,0%	100,0%	100,0%
MN_RESID_COD	Município de Residência	100,0%	100,0%	100,0%
FR_RESID_COD	Freguesia de Residência	99,5%	99,5%	99,4%
SEXO	Sexo	100,0%	100,0%	100,0%
DTNASC	Data de nascimento	100,0%	100,0%	100,0%
PAIS_NAC_COD	Código do país de nacionalidade	100,0%	100,0%	100,0%
ESTCIVIL	Estado civil	100,0%	100,0%	100,0%
HABILITACAO	Habilitação	100,0%	100,0%	100,0%
AREA_CURSO	Curso (apenas para habilitações superiores)	11,2%	13,4%	14,9%
DATA_DESEMPREGO	Data de início de contagem da situação de desemprego	96,5%	96,6%	96,8%
CATEGORIA	Categoria pedido de emprego	100,0%	100,0%	100,0%
DATA_INSCRICAO	Data da última inscrição para emprego	100,0%	100,0%	100,0%
MOTIVO_INSCRICAO	Motivo de inscrição do pedido de emprego	100,0%	100,0%	100,0%
INT_TEMPO_INSCRICAO	Contagem do tempo de inscrição	100,0%	100,0%	100,0%
INT_TEMPO_NAO_INSC	Contagem do tempo em que não esteve inscrito (só para reinscrições)	0,1%	0,2%	0,1%
CAE_ANTERIOR	Código de Atividade Económica anterior	92,7%	91,4%	88,8%
CNP_ANTERIOR	Código Nacional de Profissões anterior (até 2013)	89,7%	91,8%	90,0%
PROFISSAO_ANT_CPP_APARTIR_2014	Código Profissão anterior (Classificação Portuguesa de Profissões: a partir de 2014)	3,5%	3,4%	3,2%

Segurança Social (Qualificações ativas) – Instituto de Informática

Metainformação			
Anos disponíveis e datas de receção		Universo	
<p>19.03.2012 ↓ 2011</p> <p>17.02.2015 ↓ 2013</p>		<p>Indivíduos que têm uma qualificação ativa, no sentido de um enquadramento no sistema de segurança social. Abrange todas as pessoas singulares que mantêm uma inscrição ativa na segurança social, independentemente do tipo de qualificação: trabalhador por conta de outrem, trabalhador independente, pensionista de velhice, pensionista de invalidez, rendimento social de inserção,.....</p>	
Âmbito geográfico		Variáveis disponíveis	
Portugal		ANO, SEXO, DATA DE NASCIMENTO, ESTADO CIVIL, NÚMERO E TIPO DOCUMENTO IDENTIFICAÇÃO, NIF, NISS, PAIS, MUNICÍPIO E FREGUESIA DE NATURALIDADE, DISTRITO, MUNICÍPIO, FREGUESIA, ARTÉRIA E CÓDIGO POSTAL DE RESIDÊNCIA, NACIONALIDADE, DATA DE INÍCIO, DE FIM E TIPO DE QUALIFICAÇÃO, MOTIVO DE FIM, REGIME, NOME E NISS DA ENTIDADE EMPREGADORA, MORADA, DISTRITO, CONCELHO, FREGUESIA E MORADA DA ENTIDADE EMPREGADORA	
Unidade de observação			
Indivíduo			
Total de registos		Registos únicos	
2011	2013	2011	2013
9 062 616	9 058 682	7 209 027	7 124 233
Taxas de matching			
		BDIC NIC	
SS2011		95,7%	
SS2013		98,4%	

Segurança Social (Qualificações ativas) – Instituto de Informática

Indicadores de qualidade			
Nome da variável	Designação da variável	Taxa de preenchimento (%)	
		2011	2013
ANO	Ano a que os dados dizem respeito	100,0%	100,0%
NISS	Número de identificação da Segurança Social	100,0%	100,0%
TIPO_DOC_CIVIL	Designação do documento de identificação	69,9%	79,4%
BI	Número de identificação Civil	73,9%	79,3%
NIF	Número de identificação Fiscal	96,4%	97,6%
SEXO	Sexo	100,0%	100,0%
DTNASC	Data de nascimento	99,9%	99,9%
NOME	Nome	-	100,0%
PAIS_NAC_COD	País de Nacionalidade	86,6%	90,5%
PAIS_NAT_COD	País de Naturalidade	100,0%	100,0%
MN_NAT_COD	Município de Naturalidade	87,7%	87,0%
FR_NAT_COD	Freguesia de Naturalidade	87,7%	87,0%
DT_RESID_COD	Distrito de residência	97,8%	97,4%
MN_RESID_COD	Município de residência	97,8%	97,4%
FR_RESID_COD	Freguesia de residência	97,8%	97,4%
ARTERIA_RES	Rua	98,6%	98,1%
LOCALIDADE_RES	Localidade	98,6%	98,1%
CP4	Código postal	86,4%	90,0%
CP3	Código postal auxiliar	87,1%	90,8%
ESTCIVIL	Estado Civil	100,0%	100,0%
TIPO_QLF	Tipo de qualificação	100,0%	100,0%
DATA_INICIO_QLF	Data de início da qualificação	100,0%	100,0%
DATA_FIM_QLF	Data de fim da qualificação	2,9%	3,4%
MOTIVO_FIM	Motivo do fim da qualificação	2,9%	3,4%
REGIME	Regime de enquadramento na Segurança Social	64,7%	63,6%
NISS_EE*	NISS da Entidade empregadora	100,0%	100,0%
NIPC*	Número de identificação de pessoa coletiva	-	95,0%
FIRMA_DENOMINACAO*	Nome da entidade empregadora	94,5%	95,1%
ARTERIA_EE*	Morada da entidade empregadora	94,5%	95,0%
CP4_EE*	Código postal da entidade empregadora	94,5%	89,7%
CP3_EE*	Código postal auxiliar da entidade empregadora	94,5%	89,7%
LOCALIDADE_EE*	Localidade da entidade empregadora	94,5%	95,0%
COD_DISTRITO_EE*	Distrito da entidade empregadora	94,5%	95,0%
COD_CONCELHO_EE*	Município da entidade empregadora	94,5%	95,0%
COD_FREGUESIA_EE_ORIG*	Freguesia de origem da entidade empregadora	94,5%	95,0%
COD_FREGUESIA_EE*	Freguesia da entidade empregadora	94,5%	95,0%

*Considerando qualificações DOMINANTES: MOE (MEMBRO DE ORGÃO ESTATUTÁRIO) e TCO (TRAB.POR CONTA DE OUTREM)

BDIC – Base de Dados de Identificação Civil (IRN)

Metainformação	
Anos disponíveis e datas de receção	Universo
<p>01.11.2011</p> <p>19.12.2014</p>	<p>2010: Cidadãos de nacionalidade portuguesa residentes em Portugal (11 565 714) ou no estrangeiro (1 323 280).</p> <p>2014: Indivíduos de nacionalidade Portuguesa e brasileiros com estatuto de porto seguro com morada de residência em Portugal ou morada desconhecida (inclui 11 884 913 indivíduos vivos e 2 489 817 falecidos).</p>
Âmbito geográfico	Variáveis disponíveis
Portugal	ANO, NIC, NOME, SEXO, DATA DE NASCIMENTO, CÓDIGO POSTAL, LOCALIDADE, DISTRITO, MUNICÍPIO e FREGUESIA DE RESIDÊNCIA, PAÍS DE NACIONALIDADE, PAÍS, MUNICÍPIO E FREGUESIA DE NATURALIDADE, ESTADO CIVIL, DATA DE EMISSÃO, DATA DE VALIDADE, ESTADO DO REGISTO
Unidade de observação	
Indivíduo	
Total de registos recebidos	Registos de residentes em Portugal
<p>2010</p> <p>12 888 994</p>	<p>2014</p> <p>14 374 730</p>
	<p>2010</p> <p>11 565 714</p>
	<p>2014</p> <p>11 884 913</p>

Indicadores de qualidade			
Nome da variável	Designação da variável	Taxa de preenchimento (%)	
		2010	2014
ANO	Ano a que os dados dizem respeito	100,0%	100,0%
DT_RESID_COD	Código de distrito	100,0%	98,8%
MN_RESID_COD	Código de município	100,0%	98,8%
FR_RESID_COD	Código de freguesia	100,0%	98,8%
NOME	Nome	100,0%	100,0%
BI	Número de identificação Civil	100,0%	100,0%
SEXO	Sexo	100,0%	100,0%
DTNASC	Data de nascimento	100,0%	100,0%
PAIS_NAC_COD	País de Nacionalidade	100,0%	100,0%
PAIS_NAT_COD	País de Naturalidade	100,0%	100,0%
MN_NAT_COD	Município de Naturalidade	100,0%	100,0%
FR_NAT_COD	Freguesia de Naturalidade	100,0%	100,0%
ESTCIVIL	Estado Civil	100,0%	100,0%
CP_LOCAL_POSTAL	Localidade postal	0,0%	72,3%
CP4	Código postal	0,0%	72,3%
CP3	Código postal auxiliar	0,0%	72,3%
DATA_EMISSAO	Data de emissão	100,0%	100,0%
DATA_VALIDADE	Data de validade	0,0%	87,5%
ESTADO	Vivo / óbito	0,0%	100,0%

SEF – Serviço de Estrangeiros e Fronteiras

Metainformação					
Anos disponíveis e datas de receção			Universo		
<p>27.01.2015</p> <p>2011 2012 2013</p>			Indivíduos de nacionalidade estrangeira com título de residência válido		
Âmbito geográfico			Variáveis disponíveis		
Portugal			ANO, RSN, NIF, NISS, NOME, SEXO, DATA E LOCAL DE NASCIMENTO, ESTADO CIVIL, PAÍS NACIONALIDADE E NATURALIDADE, PAÍS ÚLTIMA RESIDÊNCIA, MORADA, CÓDIGO POSTAL, ZONA POSTAL, LOCALIDADE, DISTRITO E MUNICÍPIO DE RESIDÊNCIA, Nº, TIPO E DATA DE VALIDADE DA AUTORIZAÇÃO DE RESIDÊNCIA, SITUAÇÃO, TIPO E GRUPO PROFISSIONAL, PROFISSÃO, HABILITAÇÕES		
Unidade de observação					
Indivíduo					
Total de registos			Registos únicos		
2011	2012	2013	2011	2012	2013
434 708	414 610	398 268	434 708	414 610	398 268
Taxas de matching					
			SS NISS		SS NIF
		SEF2011	69,3%		59,8%
		SEF2012	69,4%		58,5%
		SEF2013	67,1%		56,7%

SEF – Serviço de Estrangeiros e Fronteiras

Indicadores de qualidade				
Nome da variável	Designação da variável	Taxa de preenchimento (%)		
		2011	2012	2013
ANO	Ano a que os dados dizem respeito	100,0%	100,0%	100,0%
RSN	Número interno de processo	100,0%	100,0%	100,0%
NIF	Número de Identificação Fiscal	57,8%	59,9%	61,4%
NISS	Número de Identificação da Segurança Social	49,9%	50,6%	51,1%
NUMERO_AR	Número de autorização de residência	99,9%	99,9%	100,0%
TIPO_AR	Tipo de autorização de residência	100,0%	100,0%	100,0%
DATA_VALIDADE	Data de validade da autorização de residência	100,0%	100,0%	100,0%
NOME	Nome	100,0%	100,0%	100,0%
SEXO	Sexo	100,0%	100,0%	100,0%
ESTCIVIL	Estado civil	100,0%	100,0%	100,0%
DTNASC	Data de nascimento	100,0%	100,0%	100,0%
LOCAL_NASCIMENTO	Local de nascimento	92,2%	91,8%	91,2%
PAIS_NAC_COD	País de Nacionalidade	100,0%	100,0%	100,0%
PAIS_NAT_COD	País de Naturalidade	100,0%	100,0%	100,0%
DT_RESID_COD	Distrito de residência	100,0%	100,0%	100,0%
MN_RESID_COD	Município de residência	100,0%	100,0%	100,0%
CP4	Código postal	99,7%	99,8%	99,9%
CP3	Código postal auxiliar	99,7%	99,8%	99,9%
LOCALIDADE	Localidade	99,7%	99,8%	99,9%
MORADA	Morada	99,7%	99,8%	99,9%
PAIS_ULTIMA_RESID	País de residência anterior	27,6%	42,6%	53,1%
PROFISSAO_CNP	Profissão	13,3%	14,1%	14,2%
TIPO_OCUPACAO	Tipo de ocupação	36,8%	51,7%	61,9%
SITUACAO_PROFISSIONAL	Situação profissional	36,8%	51,7%	61,9%
GRUPO_PROFISSIONAL	Grupo profissional	28,7%	35,4%	40,2%
HABILITACOES	Localidade	53,8%	59,1%	63,6%

AT – Autoridade Tributária e Aduaneira (Cadastro)

Metainformação	
Ano disponível e data de receção	Universo
<p>11.06.2015</p>	Contribuintes singulares a 31 de dezembro de 2014
Âmbito geográfico	Variáveis disponíveis
Portugal	ANO, NIF, NOME, SEXO, DATA DE NASCIMENTO, DISTRITO, MUNICÍPIO E FREGUESIA DE NATURALIDADE, PAÍS NACIONALIDADE, LOCALIDADE DA RESIDÊNCIA, CÓDIGO POSTAL E RESIDENTE EM PT (V/F)
Unidade de observação	
Indivíduo	
Total de registos	Registos únicos
2014 15 034 581	2014 14 029 321 (Residentes PT) 1 005 260 (Residentes estrangeiro)
Taxas de matching	
-	

Indicadores de qualidade		
Nome da variável	Designação da variável	Taxa de preenchimento (%) 2014
ANO	Ano a que os dados dizem respeito	100,0%
NIF	Número de Identificação Fiscal	100,0%
NOME	Nome	100,0%
SEXO	Sexo	100,0%
DTNASC	Data de Nascimento	100,0%
PAIS_NAC_COD	Pais de Nacionalidade (PT/Outro)	100,0%
DT_NAT_COD	Distrito de Naturalidade	100,0%
MN_NAT_COD	Município de Naturalidade	100,0%
FR_NAT_COD	Freguesia de Naturalidade	100,0%
RESIDÊNCIA_LOCALIDADE	Localidade de Residência	100,0%
CP4	4 primeiros dígitos do Código Postal	100,0%
CP3	3 primeiros dígitos do Código Postal	100,0%
RESIDÊNCIA_PT	Localidade em Portugal (V/F)	100,0%

Anexo 2: População residente Censos 2011, população estimada no protótipo da BPR 2011, e diferenças absolutas e relativas, por município

Código Município	Designação Município	Censos 2011	BPR 2011	Diferença absoluta BPR 2011 – Censos 2011	Diferença relativa BPR 2011 – Censos 2011
	TOTAL	10562178	8649135	-1913043	-18,1
0101	Águeda	47729	39804	-7925	-16,6
0102	Albergaria-a-Velha	25252	20678	-4574	-18,1
0103	Anadia	29150	24212	-4938	-16,9
0104	Arouca	22359	18379	-3980	-17,8
0105	Aveiro	78450	64495	-13955	-17,8
0106	Castelo de Paiva	16733	13987	-2746	-16,4
0107	Espinho	31786	27664	-4122	-13,0
0108	Estarreja	26997	21992	-5005	-18,5
0109	Santa Maria da Feira	139309	114293	-25016	-18,0
0110	Ílhavo	38598	31298	-7300	-18,9
0111	Mealhada	20428	16794	-3634	-17,8
0112	Murtosa	10585	7687	-2898	-27,4
0113	Oliveira de Azeméis	68611	57423	-11188	-16,3
0114	Oliveira do Bairro	23028	18559	-4469	-19,4
0115	Ovar	55398	46102	-9296	-16,8
0116	São João da Madeira	21713	19062	-2651	-12,2
0117	Sever do Vouga	12356	10100	-2256	-18,3
0118	Vagos	22851	18257	-4594	-20,1
0119	Vale de Cambra	22864	19009	-3855	-16,9
0201	Aljustrel	9257	8169	-1088	-11,8
0202	Almodôvar	7449	6294	-1155	-15,5
0203	Alvito	2504	2039	-465	-18,6
0204	Barrancos	1834	1477	-357	-19,5
0205	Beja	35854	30180	-5674	-15,8
0206	Castro Verde	7276	6202	-1074	-14,8
0207	Cuba	4878	4019	-859	-17,6
0208	Ferreira do Alentejo	8255	7173	-1082	-13,1
0209	Mértola	7274	6209	-1065	-14,6
0210	Moura	15167	13232	-1935	-12,8
0211	Odemira	26066	19894	-6172	-23,7
0212	Ourique	5389	4517	-872	-16,2
0213	Serpa	15623	13060	-2563	-16,4
0214	Vidigueira	5932	4988	-944	-15,9
0301	Amares	18889	16302	-2587	-13,7
0302	Barcelos	120391	99085	-21306	-17,7
0303	Braga	181494	152953	-28541	-15,7

Código Município	Designação Município	Censos 2011	BPR 2011	Diferença absoluta BPR 2011 – Censos 2011	Diferença relativa BPR 2011 – Censos 2011
0304	Cabeceiras de Basto	16710	14024	-2686	-16,1
0305	Celorico de Basto	20098	15698	-4400	-21,9
0306	Esposende	34254	27830	-6424	-18,8
0307	Fafe	50633	43289	-7344	-14,5
0308	Guimarães	158088	136059	-22029	-13,9
0309	Póvoa de Lanhoso	21918	18124	-3794	-17,3
0310	Terras de Bouro	7253	6476	-777	-10,7
0311	Vieira do Minho	13001	10920	-2081	-16,0
0312	Vila Nova de Famalicão	133832	113254	-20578	-15,4
0313	Vila Verde	47888	39661	-8227	-17,2
0314	Vizela	23736	20507	-3229	-13,6
0401	Alfândega da Fé	5104	4430	-674	-13,2
0402	Bragança	35341	30134	-5207	-14,7
0403	Carraceda de Ansiães	6373	5170	-1203	-18,9
0404	Freixo de Espada à Cinta	3780	3054	-726	-19,2
0405	Macedo de Cavaleiros	15776	13500	-2276	-14,4
0406	Miranda do Douro	7482	5980	-1502	-20,1
0407	Mirandela	23850	20174	-3676	-15,4
0408	Mogadouro	9542	8041	-1501	-15,7
0409	Torre de Moncorvo	8572	7082	-1490	-17,4
0410	Vila Flor	6697	5562	-1135	-16,9
0411	Vimioso	4669	3861	-808	-17,3
0412	Vinhais	9066	7738	-1328	-14,6
0501	Belmonte	6859	5592	-1267	-18,5
0502	Castelo Branco	56109	47457	-8652	-15,4
0503	Covilhã	51797	43222	-8575	-16,6
0504	Fundão	29213	23835	-5378	-18,4
0505	Idanha-a-Nova	9716	7787	-1929	-19,9
0506	Oleiros	5721	4408	-1313	-23,0
0507	Penamacor	5682	4370	-1312	-23,1
0508	Proença-a-Nova	8314	6764	-1550	-18,6
0509	Sertã	15880	12887	-2993	-18,8
0510	Vila de Rei	3452	2729	-723	-20,9
0511	Vila Velha de Ródão	3521	2714	-807	-22,9
0601	Arganil	12145	10160	-1985	-16,3
0602	Cantanhede	36595	30120	-6475	-17,7
0603	Coimbra	143396	124213	-19183	-13,4
0604	Condeixa-a-Nova	17078	14184	-2894	-16,9
0605	Figueira da Foz	62125	53072	-9053	-14,6
0606	Góis	4260	3533	-727	-17,1

Código Município	Designação Município	Censos 2011	BPR 2011	Diferença absoluta BPR 2011 – Censos 2011	Diferença relativa BPR 2011 – Censos 2011
0607	Lousã	17604	14684	-2920	-16,6
0608	Mira	12465	10173	-2292	-18,4
0609	Miranda do Corvo	13098	11095	-2003	-15,3
0610	Montemor-o-Velho	26171	21609	-4562	-17,4
0611	Oliveira do Hospital	20855	17719	-3136	-15,0
0612	Pampilhosa da Serra	4481	3283	-1198	-26,7
0613	Penacova	15251	12867	-2384	-15,6
0614	Penela	5983	4911	-1072	-17,9
0615	Soure	19245	16384	-2861	-14,9
0616	Tábua	12071	10318	-1753	-14,5
0617	Vila Nova de Poiares	7281	6113	-1168	-16,0
0701	Alandroal	5843	5113	-730	-12,5
0702	Arraiolos	7363	6281	-1082	-14,7
0703	Borba	7333	6296	-1037	-14,1
0704	Estremoz	14318	12612	-1706	-11,9
0705	Évora	56596	47574	-9022	-15,9
0706	Montemor-o-Novo	17437	14855	-2582	-14,8
0707	Mora	4978	4491	-487	-9,8
0708	Mourão	2663	2375	-288	-10,8
0709	Portel	6428	5545	-883	-13,7
0710	Redondo	7031	6177	-854	-12,1
0711	Reguengos de Monsaraz	10828	9455	-1373	-12,7
0712	Vendas Novas	11846	10134	-1712	-14,5
0713	Viana do Alentejo	5743	4998	-745	-13,0
0714	Vila Viçosa	8319	7255	-1064	-12,8
0801	Albufeira	40828	29873	-10955	-26,8
0802	Alcoutim	2917	2502	-415	-14,2
0803	Aljezur	5884	3885	-1999	-34,0
0804	Castro Marim	6747	5336	-1411	-20,9
0805	Faro	65019	53500	-11519	-17,7
0806	Lagoa	22975	17557	-5418	-23,6
0807	Lagos	31049	22919	-8130	-26,2
0808	Loulé	70163	50675	-19488	-27,8
0809	Monchique	6045	4936	-1109	-18,3
0810	Olhão	45396	35186	-10210	-22,5
0811	Portimão	55614	44190	-11424	-20,5
0812	São Brás de Alportel	10662	8311	-2351	-22,1
0813	Silves	37126	28524	-8602	-23,2
0814	Tavira	26167	20278	-5889	-22,5
0815	Vila do Bispo	5258	3809	-1449	-27,6

Código Município	Designação Município	Censos 2011	BPR 2011	Diferença absoluta BPR 2011 – Censos 2011	Diferença relativa BPR 2011 – Censos 2011
0816	Vila Real de Santo António	19156	15946	-3210	-16,8
0901	Aguiar da Beira	5473	4629	-844	-15,4
0902	Almeida	7242	5766	-1476	-20,4
0903	Celorico da Beira	7693	6434	-1259	-16,4
0904	Figueira de Castelo Rodrigo	6260	5216	-1044	-16,7
0905	Fornos de Algodres	4989	4286	-703	-14,1
0906	Gouveia	14046	11942	-2104	-15,0
0907	Guarda	42541	37210	-5331	-12,5
0908	Manteigas	3430	3016	-414	-12,1
0909	Meda	5202	4401	-801	-15,4
0910	Pinhel	9627	8385	-1242	-12,9
0911	Sabugal	12544	9664	-2880	-23,0
0912	Seia	24702	21957	-2745	-11,1
0913	Trancoso	9878	7858	-2020	-20,4
0914	Vila Nova de Foz Côa	7312	5908	-1404	-19,2
1001	Alcobaça	56688	47167	-9521	-16,8
1002	Alvaiázere	7287	5801	-1486	-20,4
1003	Ansião	13092	10644	-2448	-18,7
1004	Batalha	15805	13192	-2613	-16,5
1005	Bombarral	13193	10587	-2606	-19,8
1006	Caldas da Rainha	51729	42055	-9674	-18,7
1007	Castanheira de Pêra	3191	2795	-396	-12,4
1008	Figueiró dos Vinhos	6169	5336	-833	-13,5
1009	Leiria	126884	104603	-22281	-17,6
1010	Marinha Grande	38699	32692	-6007	-15,5
1011	Nazaré	15158	12762	-2396	-15,8
1012	Óbidos	11772	9360	-2412	-20,5
1013	Pedrógão Grande	3915	3172	-743	-19,0
1014	Peniche	27753	22784	-4969	-17,9
1015	Pombal	55245	44296	-10949	-19,8
1016	Porto de Mós	24342	20721	-3621	-14,9
1101	Alenquer	43267	33199	-10068	-23,3
1102	Arruda dos Vinhos	13391	10157	-3234	-24,2
1103	Azambuja	21814	16433	-5381	-24,7
1104	Cadaval	14228	11089	-3139	-22,1
1105	Cascais	206479	155361	-51118	-24,8
1106	Lisboa	552700	449644	-103056	-18,6
1107	Loures	199494	157304	-42190	-21,1
1108	Lourinhã	25735	20134	-5601	-21,8

Código Município	Designação Município	Censos 2011	BPR 2011	Diferença absoluta BPR 2011 – Censos 2011	Diferença relativa BPR 2011 – Censos 2011
1109	Mafra	76685	57962	-18723	-24,4
1110	Oeiras	172120	135261	-36859	-21,4
1111	Sintra	377835	292939	-84896	-22,5
1112	Sobral de Monte Agraço	10156	7991	-2165	-21,3
1113	Torres Vedras	79465	62657	-16808	-21,2
1114	Vila Franca de Xira	136886	108152	-28734	-21,0
1115	Amadora	175136	133504	-41632	-23,8
1116	Odivelas	145142	113252	-31890	-22,0
1201	Alter do Chão	3562	3074	-488	-13,7
1202	Arronches	3165	2556	-609	-19,2
1203	Avis	4571	3709	-862	-18,9
1204	Campo Maior	8456	7149	-1307	-15,5
1205	Castelo de Vide	3407	2783	-624	-18,3
1206	Crato	3708	3173	-535	-14,4
1207	Elvas	23078	19297	-3781	-16,4
1208	Fronteira	3410	2913	-497	-14,6
1209	Gavião	4132	3480	-652	-15,8
1210	Marvão	3512	2909	-603	-17,2
1211	Monforte	3329	2829	-500	-15,0
1212	Nisa	7450	6043	-1407	-18,9
1213	Ponte de Sor	16722	14357	-2365	-14,1
1214	Portalegre	24930	22126	-2804	-11,2
1215	Sousel	5074	4334	-740	-14,6
1301	Amarante	56264	47634	-8630	-15,3
1302	Baião	20522	16943	-3579	-17,4
1303	Felgueiras	58065	48630	-9435	-16,2
1304	Gondomar	168027	141021	-27006	-16,1
1305	Lousada	47387	39426	-7961	-16,8
1306	Maia	135306	110823	-24483	-18,1
1307	Marco de Canaveses	53450	44296	-9154	-17,1
1308	Matosinhos	175478	144084	-31394	-17,9
1309	Paços de Ferreira	56340	44793	-11547	-20,5
1310	Paredes	86854	70294	-16560	-19,1
1311	Penafiel	72265	59282	-12983	-18,0
1312	Porto	237591	201223	-36368	-15,3
1313	Póvoa de Varzim	63408	51855	-11553	-18,2
1314	Santo Tirso	71530	61682	-9848	-13,8
1315	Valongo	93858	78795	-15063	-16,0
1316	Vila do Conde	79533	65277	-14256	-17,9
1317	Vila Nova de Gaia	302298	252452	-49846	-16,5
1318	Trofa	38999	31978	-7021	-18,0

Código Município	Designação Município	Censos 2011	BPR 2011	Diferença absoluta BPR 2011 – Censos 2011	Diferença relativa BPR 2011 – Censos 2011
1401	Abrantes	39325	34098	-5227	-13,3
1402	Alcanena	13868	11809	-2059	-14,8
1403	Almeirim	23376	19739	-3637	-15,6
1404	Alpiarça	7702	6514	-1188	-15,4
1405	Benavente	29019	23140	-5879	-20,3
1406	Cartaxo	24462	20391	-4071	-16,6
1407	Chamusca	10120	8425	-1695	-16,7
1408	Constância	4056	3403	-653	-16,1
1409	Coruche	19944	17415	-2529	-12,7
1410	Entroncamento	20206	16836	-3370	-16,7
1411	Ferreira do Zêzere	8619	7200	-1419	-16,5
1412	Golegã	5913	5071	-842	-14,2
1413	Mação	7338	6408	-930	-12,7
1414	Rio Maior	21192	17661	-3531	-16,7
1415	Salvaterra de Magos	22159	18598	-3561	-16,1
1416	Santarém	61752	52060	-9692	-15,7
1417	Sardoal	3939	3439	-500	-12,7
1418	Tomar	40677	33973	-6704	-16,5
1419	Torres Novas	36717	31165	-5552	-15,1
1420	Vila Nova da Barquinha	7322	6115	-1207	-16,5
1421	Ourém	45940	37446	-8494	-18,5
1501	Alcácer do Sal	13046	10989	-2057	-15,8
1502	Alcochete	17569	12922	-4647	-26,4
1503	Almada	174030	138356	-35674	-20,5
1504	Barreiro	78764	64206	-14558	-18,5
1505	Grândola	14826	11986	-2840	-19,2
1506	Moita	66029	55306	-10723	-16,2
1507	Montijo	51222	39046	-12176	-23,8
1508	Palmela	62831	49422	-13409	-21,3
1509	Santiago do Cacém	29749	24735	-5014	-16,9
1510	Seixal	158269	125267	-33002	-20,9
1511	Sesimbra	49500	38511	-10989	-22,2
1512	Setúbal	121185	99268	-21917	-18,1
1513	Sines	14238	11848	-2390	-16,8
1601	Arcos de Valdevez	22847	17341	-5506	-24,1
1602	Caminha	16684	13819	-2865	-17,2
1603	Melgaço	9213	6257	-2956	-32,1
1604	Monção	19230	13991	-5239	-27,2
1605	Paredes de Coura	9198	7388	-1810	-19,7
1606	Ponte da Barca	12061	10053	-2008	-16,6
1607	Ponte de Lima	43498	35298	-8200	-18,9

Código Município	Designação Município	Censos 2011	BPR 2011	Diferença absoluta BPR 2011 – Censos 2011	Diferença relativa BPR 2011 – Censos 2011
1608	Valença	14127	11166	-2961	-21,0
1609	Viana do Castelo	88725	74176	-14549	-16,4
1610	Vila Nova de Cerveira	9253	7287	-1966	-21,2
1701	Alijó	11942	10281	-1661	-13,9
1702	Boticas	5750	4925	-825	-14,3
1703	Chaves	41243	33021	-8222	-19,9
1704	Mesão Frio	4433	3878	-555	-12,5
1705	Mondim de Basto	7493	6551	-942	-12,6
1706	Montalegre	10537	8612	-1925	-18,3
1707	Murça	5952	5153	-799	-13,4
1708	Peso da Régua	17131	14864	-2267	-13,2
1709	Ribeira de Pena	6544	5739	-805	-12,3
1710	Sabrosa	6361	5236	-1125	-17,7
1711	Santa Marta de Penaguião	7356	6156	-1200	-16,3
1712	Valpaços	16882	14081	-2801	-16,6
1713	Vila Pouca de Aguiar	13187	11240	-1947	-14,8
1714	Vila Real	51850	43310	-8540	-16,5
1801	Armamar	6297	5207	-1090	-17,3
1802	Carregal do Sal	9835	8043	-1792	-18,2
1803	Castro Daire	15339	13106	-2233	-14,6
1804	Cinfães	20427	16683	-3744	-18,3
1805	Lamego	26691	23147	-3544	-13,3
1806	Mangualde	19880	16748	-3132	-15,8
1807	Moimenta da Beira	10212	8511	-1701	-16,7
1808	Mortágua	9607	8001	-1606	-16,7
1809	Nelas	14037	12058	-1979	-14,1
1810	Oliveira de Frades	10261	8812	-1449	-14,1
1811	Penalva do Castelo	7956	6834	-1122	-14,1
1812	Penedono	2952	2301	-651	-22,1
1813	Resende	11364	9935	-1429	-12,6
1814	Santa Comba Dão	11597	9700	-1897	-16,4
1815	São João da Pesqueira	7874	6238	-1636	-20,8
1816	São Pedro do Sul	16851	14284	-2567	-15,2
1817	Sátão	12444	10197	-2247	-18,1
1818	Sernancelhe	5671	4479	-1192	-21,0
1819	Tabuaço	6350	4861	-1489	-23,4
1820	Tarouca	8048	6445	-1603	-19,9
1821	Tondela	28946	23915	-5031	-17,4
1822	Vila Nova de Paiva	5176	4325	-851	-16,4
1823	Viseu	99274	83013	-16261	-16,4

Código Município	Designação Município	Censos 2011	BPR 2011	Diferença absoluta BPR 2011 – Censos 2011	Diferença relativa BPR 2011 – Censos 2011
1824	Vouzela	10564	8963	-1601	-15,2
3101	Calheta (R.A.M.)	11521	10194	-1327	-11,5
3102	Câmara de Lobos	35666	32428	-3238	-9,1
3103	Funchal	111892	101916	-9976	-8,9
3104	Machico	21828	19970	-1858	-8,5
3105	Ponta do Sol	8862	8073	-789	-8,9
3106	Porto Moniz	2711	2612	-99	-3,7
3107	Ribeira Brava	13375	12345	-1030	-7,7
3108	Santa Cruz	43005	37690	-5315	-12,4
3109	Santana	7719	7050	-669	-8,7
3110	São Vicente	5723	5105	-618	-10,8
3201	Porto Santo	5483	6432	949	17,3
4101	Vila do Porto	5552	4114	-1438	-25,9
4201	Lagoa	14442	10429	-4013	-27,8
4202	Nordeste	4937	3691	-1246	-25,2
4203	Ponta Delgada	68809	51380	-17429	-25,3
4204	Povoação	6327	4787	-1540	-24,3
4205	Ribeira Grande	32112	23788	-8324	-25,9
4206	Vila Franca do Campo	11229	7681	-3548	-31,6
4301	Angra do Heroísmo	35402	28504	-6898	-19,5
4302	Vila da Praia da Vitória	21035	15939	-5096	-24,2
4401	Santa Cruz da Graciosa (R.A.A.)	4391	3582	-809	-18,4
4501	Calheta (R.A.A.)	3773	3037	-736	-19,5
4502	Velas (R.A.A.)	5398	4093	-1305	-24,2
4601	Lajes do Pico	4711	3682	-1029	-21,8
4602	Madalena	6049	4722	-1327	-21,9
4603	São Roque do Pico	3388	2648	-740	-21,8
4701	Horta	14994	11638	-3356	-22,4
4801	Lajes das Flores	1504	1058	-446	-29,7
4802	Santa Cruz das Flores	2289	1771	-518	-22,6
4901	Corvo	430	313	-117	-27,2
0000	Desconhecido	-	6221	-	-