
PARTE II

DADOS E MÉTODOS BÁSICOS DA ANÁLISE DEMOGRÁFICA

O CENSO (RECENSEAMENTO) DEMOGRÁFICO

4.1 INTRODUÇÃO

Como a demografia tende a tratar mais da descrição e sistematização dos fenômenos estudados do que da sua explicação teórica, é natural que a coleta (recolha), a crítica e a correção dos dados cumpram um papel fundamental na sua prática. Diferentemente de outros cientistas sociais, o demógrafo tradicionalmente pode se basear em observações referentes a todo o universo estudado, mas ao mesmo tempo o seu controle sobre a natureza da informação é limitado. Normalmente, esta é produzida por órgãos administrativos, com finalidades que, no mínimo, vão além do seu interesse puramente científico (Hakkert, 1996). Os dados tradicionais podem ser complementados por levantamentos especiais, para investigar questões de interesse específico, mas os tamanhos amostrais necessários para investigar muitos fenômenos do interesse da demografia tendem a ser grandes, de modo que geralmente precisam de algum respaldo institucional. Sendo assim, a posição do demógrafo, da mesma forma que o epidemiólogo, é intermediária entre aquela do historiador, que precisa se basear inteiramente na informação colocada à sua disposição, e o psicólogo ou antropólogo, que frequentemente podem estudar a realidade social a partir de pesquisas com perguntas muito direcionadas e amostras muito reduzidas.

A qualidade da informação demográfica pode ser avaliada em termos de diferentes critérios:

- Relevância: Em que medida a informação produzida corresponde às necessidades de informação dos usuários potenciais ?

- Precisão: Em que medida a informação produzida descreve corretamente os fenômenos que pretende medir ?
- Confiabilidade: Qual é a variação dos resultados quando são medidos repetidas vezes ?
- Oportunidade: Quanto tempo passou entre o momento da medição da informação e a sua disponibilidade efetiva ?
- Coerência e comparabilidade: Até que ponto a informação é internamente consistente e satisfaz os mesmos critérios de produção de outros resultados obtidos ao longo do tempo ou em outros lugares ?
- Acessibilidade: A facilidade com que os usuários potenciais podem ter acesso à informação.
- Clareza: Os resultados são apresentados de forma clara, legível e compreensível ? São acompanhados de metadados e outras explicações que possibilitam avaliar o seu significado e forma de produção ?
- Independência: O grau de liberdade de interferências políticas ou comerciais que podem afetar os padrões, procedimentos, conteúdo e tempo de divulgação da informação.
- Disponibilidade de recursos: Os recursos humanos, financeiros e técnicos investidos na produção da informação foram suficientes para cumprir com as necessidades ?

Embora os estudiosos da população hoje em dia usem uma variedade de informações, inclusive de tipo qualitativo, o dado propriamente demográfico normalmente pode ser classificado como estatística de *estoque* ou de *fluxo* (ver Quadro 4.1). A primeira refere-se às características estáticas de uma população num determinado momento – como se fosse uma fotografia. Essas características incluem o tamanho da população, a sua distribuição territorial, por sexo, por idade ou por características socioeconômicas. A segunda categoria trata da dinâmica populacional, ou seja, dos eventos que transformam essas características – como se fosse um filme. As características principais de fluxo são os nascimentos, óbitos, migração e mobilidade entre categorias socioeconômicas. Tradicionalmente, os dados nos quais estes dois tipos de estatísticas se baseiam provêm de fontes distintas. Enquanto a informação para as estatísticas de estoque é tipicamente obtida por meio de censos ou inquéritos amostrais (“surveys”), a medição de fluxos pode ser realizada por meio de um sistema de registro de eventos ou, mais raramente, um levantamento de passagens repetidas (“multi-round survey”).

Quadro 4.1: Visão esquemática das variáveis de estoque e de fluxo

Estoque de População	Fluxo (Dinâmica) de População
Tamanho da População	Natalidade, Mortalidade
Composição da População	Mobilidade Social
Distribuição Espacial	Migração
Estudo Transversal	Estudo Longitudinal
Censo ou Inquérito	Estatística Contínua (Vital)

A identificação de cada tipo de estatística com uma ou outra fonte, porém, não é rígida ou absoluta. Muito pode ser aprendido sobre a natalidade (estatística de fluxo) de uma população no passado pela inspeção cuidadosa da pirâmide etária encontrada num censo (estatística de estoque). Também existem técnicas (chamadas *indiretas*, que serão discutidas no Capítulo 23) que permitem estimar a mortalidade, fecundidade ou migração a partir de dados censitários retrospectivos. Isso acontece particularmente em países como os PALOP, onde a qualidade dos censos costuma ser muito melhor do que a qualidade do registro dos eventos vitais. Mesmo em países onde o registro civil é mais completo, a única forma de obter dados confiáveis sobre determinados fenômenos demográficos, como a migração interna ou a fecundidade diferenciada por religião (onde esta consta no censo mas não no registro civil), pode ser por meio de dados censitários retrospectivos. Este é um dos argumentos levantados por Queiroz e Sacco (2018), para defender a inclusão de perguntas sobre mortalidade nos censos da América Latina, apesar da existência de fontes alternativas de boa qualidade, na maioria dos países. Por outro lado, o registro de nascimentos, óbitos ou casamentos às vezes é aproveitado de forma *sintomática*, ou seja, para obter alguma indicação do tamanho ou da composição da população que produziu estes eventos, quando não existem censos ou inquéritos suficientemente recentes para fornecer a mesma informação diretamente (ver seção 21.10 do Capítulo 21). A seguir serão discutidos os aspectos mais gerais das diferentes fontes de dados mais frequentemente usadas na análise da população. Maiores detalhes sobre o seu uso na análise de processos demográficos específicos podem ser encontrados nos respectivos capítulos deste livro.

4.2 CONCEITO E TIPOLOGIA

Apesar das desvantagens de alto custo, divulgação demorada e frequência reduzida, e a despeito de existirem hoje várias alternativas de obtenção de informação, o censo (em alguns países se usa mais o termo “recenseamento”) ainda é o principal instrumento para obter dados de estoque sobre a população, principalmente nos países em desenvolvimento, onde existem relativamente poucas fontes alternativas. Segundo a definição das Nações Unidas (United Nations, 2017 a: 1.4), um censo é “o processo total de coleta (recolha), processamento, avaliação, análise e publicação ou divulgação de dados demográficos, econômicos e sociais referentes a todas as pessoas dentro de um país ou de uma parte bem definida de um país num momento específico até o menor nível geográfico considerado apropriado”. Aprofundando esta definição, as Nações Unidas especificam que um levantamento de dados populacionais deve satisfazer, no mínimo, os critérios especificados no Quadro 4.2 para ser considerado um censo.

Quadro 4.2: Critérios mínimos para um censo de população

1.	Respaldo legal com especificação dos fins, orçamento, administração, sigilo e similares
2.	Periodicidade definida, de preferência de cinco ou dez anos
3.	Simultaneidade de todo o levantamento, com um tempo de referência predefinido
4.	Referência territorial pré-fixada. Normalmente, o território nacional
5.	Universalidade da enumeração dentro deste território
6.	Enumeração individual de todas as pessoas
7.	Disponibilidade dos resultados dentro de prazos compatíveis
8.	Capacidade de produzir estatísticas de pequenas áreas

Fonte: United Nations (2017 a).

A seguir, cada um destes pontos é discutido em algum detalhe.

Critério 1: Respaldo legal com especificação dos fins, orçamento, administração, sigilo e similares

O respaldo legal geralmente se baseia numa lei específica. A lei deve especificar o escopo e data de execução, a obrigatoriedade de cooperação do público e a obrigatoriedade do registro correto das respostas pela entidade executora, os fins, o orçamento, a administração, as garantias legais quanto ao sigilo da informação e as demais obrigações da entidade executora. Ao contrário do que acontece em outros tipos de levantamentos, nos quais não se pode obrigar o entrevistado a cooperar, o não fornecimento de informação ou o fornecimento de informação falsa à autoridade censitária normalmente está sujeito a sanções legais. O principal direito estabelecido para o entrevistado é o direito à confidencialidade da informação. No caso de censos baseados em registros administrativos (ver abaixo), a lei sobre o uso de estatísticas deve conter providências para poder usar a informação contida nas bases de dados existentes para os fins do censo. Se esse for o caso, não há necessidade de uma lei censitária específica.

A lei censitária normalmente não estabelece um direito individual a ser recenseado, embora em muitos casos se dê uma oportunidade ao público para ajudar a corrigir eventuais omissões. O critério de universalidade (Critério 3) exige que se envidem todos os esforços necessários para recensear todos os habitantes do país, mas a implementação prática deste critério geralmente é formulada em termos de uma obrigação da autoridade censitária, não como um direito individual da população para ser recenseada. As eventuais consequências negativas de omissões no censo geralmente se manifestam ao nível de entidades locais de administração pública (particularmente os municípios) e não individualmente. Portanto, os eventuais recursos legais para questionar tais omissões normalmente ocorrem nesse nível.

Critério 2: Periodicidade definida, de preferência de 5 ou 10 anos

Nas Américas, o único país que sistematicamente realiza os seus censos de 5 em 5 anos é o Canadá, desde 1951. Outros países que, em princípio, mantêm censos quinquenais, embora com algumas interrupções, incluem Austrália, Irlanda, Japão, México (desde 1990) e Nova Zelândia

(embora o Censo de 2011 tenha sido adiado para 2013) e algumas repúblicas insulares do Oceano Pacífico. Timor-Leste teve um censo em 2010 e outro em 2015, mas o anterior foi em 2004. Outros países têm conseguido manter uma rotina de censos decenais nas últimas décadas, sendo que a tradição mais longa neste sentido é a da Índia e Bangladesh, onde existe uma série histórica de censos decenais desde 1881, interrompida apenas pela guerra de independência de Bangladesh nos anos 70. Na América Latina, a série histórica mais longa é a do Panamá, desde 1920.

Desde o Congresso Internacional de Estatística de Petersburgo em 1897, houve esforços por parte de várias entidades internacionais para sincronizar os censos em anos terminando em “0” ou “5”, mas na prática tal objetivo se mostrou pouco viável. Em vez disso, foram definidas as “rodadas” (“rondas”) de censos, centradas em anos terminando em “0”. A rodada (ronda) dos censos de 2020 começou em 2015 e vai até 2024.

O Brasil teve censos decenais desde 1940, mas em 1990 a série foi interrompida com o adiamento do censo para 1991; o mesmo aconteceu em 2020, quando o censo teve que ser adiado para 2021 por conta da epidemia da COVID-19. Moçambique teve censos gerais desde 1930 até 1980, quando a série histórica teve que ser interrompida por causa da guerra, sendo retomada em 1997¹. Cabo Verde tem conseguido realizar censos decenais desde 1940. A série histórica dos censos de São Tomé & Príncipe também é razoavelmente regular. Os censos da Guiné-Bissau têm sido mais irregulares e em Angola o último censo colonial foi em 1970, enquanto o primeiro censo depois da independência foi realizado só em 2014 (Nzatuzola, 2015). Houve uma tentativa de realização de um censo em 1983, mas como isso foi no meio da guerra civil só foi possível enumerar as áreas controladas pelo governo e os resultados nunca foram completamente analisados. Em Moçambique algo parecido aconteceu em 1991 quando se realizou um *Inquérito Demográfico Nacional* nas áreas controladas pelo governo, em vez de um censo completo em todo o território nacional. Um outro país com uma grande lacuna na sua história censitária é o Uruguai, que foi um dos primeiros países latino-americanos a realizar um censo, em 1852, mas que suspendeu a rotina em 1908, retomando a série só em 1963. Em menor grau, o mesmo aconteceu na Argentina entre 1914 e 1947.

A Tabela 4.1 retrata a história censitária dos principais países latino-americanos e de língua portuguesa. África do Sul é incluído para fins comparativos, sendo que a série daquele país é bastante completa, mas de qualidade discutível antes de 1996, devido aos vieses decorrentes da política de segregação racial.

¹ Também houve censos parciais (da capital ou de grupos populacionais específicos) em 1894, 1899, 1904, 1909, 1912 e 1928.

Tabela 4.1: Censos demográficos de países selecionados desde 1900

África do Sul	1911	1921	1936	1951	1960	1970	1980	1985	1991	1996	2001	2011	2021
Angola	1915	1940	1950	1970	(1983)	2014							
Argentina	1914	1947	1960	1970	1980	1991	2001	2010	2020				
Bolívia	1900	1950	1976	1992	2001	2012	2022						
Brasil	1900	1920	1940	1950	1960	1970	1980	1991	2000	2010	2021*)		
Cabo Verde	1900	1940	1950	1960	1970	1980	1990	2000	2010	2020			
Chile	1907	1920	1930	1940	1952	1960	1970	1982	1992	2002	(2012)	2017	
Guiné-Bissau	1940	1950	1979	1991	2009								
Guiné Equatorial	1965	1994	2001	2015									
Macau	1900	1911	1950	1960	1981	1991	2001	2011	2021				
México	1910	1921	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020	
Moçambique	1940	1950	1960	1970	1980	1997	2007	2017					
Paraguai	1936	1950	1962	1972	1982	1992	2002	(2012)	2015				
Portugal	1900	1911	1920	1930	1940	1950	1960	1970	1981	1991	2001	2011	2021
S Tomé & Príncipe	1900	1915	1921	1931	1940	1950	1960	1970	1981	1991	2001	2012	2022
Timor-Leste	1936	1970	1980	1990	2000	2004	2010	2015	2020				
Uruguai	1908	1963	1975	1985	1996	2004 ²	2011	2021					
Venezuela	1920	1926	1936	1941	1950	1961	1971	1981	1990	2001	2011	2021	

*) Os anos em cursivas se referem a censos que ainda não foram realizados no momento da edição deste livro.

Critério 3: Simultaneidade de todo o levantamento, com um momento de referência preestabelecido

Normalmente, o momento de referência é à meia-noite de uma determinada data para outra. Geralmente esse momento é chamado de “data censitária”, embora o significado seja mais preciso. Certas perguntas, como aquelas referentes à atividade econômica, exigem um período de referência mais extenso, normalmente de uma semana ou mês antes da data censitária. Embora a contagem efetiva possa demorar várias semanas ou até meses (um ano no Censo da Colômbia de 2005, quase cinco meses no Censo do Jamaica em 2011 e três meses no Censo brasileiro de 2010), toda a informação recolhida deve referir-se à situação existente na data censitária. Pessoas já falecidas antes da chegada do recenseador, mas ainda vivas no momento de referência, devem ser incluídas na contagem (desde que haja informantes para relatar o evento), enquanto crianças nascidas posteriormente não devem ser contadas. Na prática, a implementação destes procedimentos pode ser difícil quando o processo de contagem é demorado e o nível de instrução da população é baixo. Nestas condições a ideia de que todo o censo se refere à mesma data de referência pode implicar certo viés.

Na medida do possível, é recomendável que a mesma data de referência seja mantida em censos sucessivos. Nos EUA e em Porto Rico, por exemplo, o censo se realiza sempre em 1º de abril; em 1940 e 1950, a data censitária no Brasil foi o dia 1º de julho, entre 1960 e 1991, o 1º de

² Embora oficialmente seja considerado um censo, as características do Censo de 2004 foram mais parecidas com uma contagem. Os objetivos, segundo o INE de Uruguai, foram: “atualizar a cartografia, obter informações úteis para a estimativa de saldos migratórios e a revisão de projeções demográficas e obter informações básicas para renovar o desenho amostral de pesquisas domiciliares”.

setembro, e desde 2000 tem sido o dia 1º de agosto. Moçambique executa o seu censo na primeira metade de agosto.

Critério 4: Referência territorial prefixada. Normalmente, o território nacional

Normalmente este território abrange todo o país, mas existem exceções. Por exemplo, os censos coloniais dos países africanos geralmente tinham finalidades mais administrativas do que propriamente estatísticas, razão pela qual se limitavam a algumas regiões-chaves destes países. Em 1983, no meio da guerra civil, Angola realizou um censo limitado às áreas do país então controladas pelo governo. O problema nesses casos é estabelecer com precisão quais são as áreas cobertas pela definição e aplicar uma metodologia uniforme de enumeração dentro delas.

Em alguns países latino-americanos com áreas florestais de difícil acesso, habitadas por populações indígenas, como Brasil e Venezuela, não toda a população indígena é enumerada. Apesar dos avanços havidos nas últimas décadas para assegurar que a população indígena seja corretamente representada pelos censos (Pereira, 2012), existem grupos isolados (com talvez uns 10 mil indivíduos) no Brasil que até agora evitam o contato com o mundo externo e que, portanto, não são considerados no universo do censo. Em outros países da América Latina, onde não existem populações indígenas completamente isoladas, esta situação não se apresenta. Alguns países, como Peru realizam censos especiais para as comunidades indígenas em áreas como a Amazônia. Os diplomatas estrangeiros residentes no país também costumam ser excluídos. Alguns países (Austrália, Costa Rica, Índia, Indonésia, Irlanda, México, Reino Unido, Uruguai, Vietnã, entre outros) fazem um esforço para contar as populações de rua, eventualmente usando módulos e metodologias especiais. No Brasil até 2021 esse grupo tem sido excluído, embora existam estimativas realizadas por entidades governamentais fora da operação censitária (Natalino, 2016).

Critério 5: Universalidade da enumeração dentro do território preestabelecido

Universalidade não implica a ausência de erros de cobertura, pois nenhum censo está livre de omissões, mas indica o propósito explícito de uma enumeração completa. Existe a possibilidade de amostragem, desde que as informações básicas sejam recolhidas para o conjunto da população. Entretanto, o critério de universalidade não é satisfeito pelos censos inteiramente conduzidos com base em amostras, tais como o censo de Jamaica de 1953, que enumerou apenas 5% da população da capital e 10% nas áreas rurais. Estritamente falando, isso também exclui os chamados “by-censuses” de Macau de 2006 e 2016, que foram conduzidos com amostras de 1 em 7.

Critério 6: A enumeração individual de todas as pessoas, seja pelo critério da sua residência habitual (de iure) ou pelo critério de local de presença na data do censo (de facto)

Este critério não é satisfeito por muitos censos da antiguidade, que geralmente tinham finalidades militares ou fiscais e se limitavam à enumeração dos lares ou dos indivíduos relevantes do ponto de vista militar ou fiscal. Nos censos demográficos de hoje, evita-se ao máximo qualquer referência a finalidades tributárias ou de alistamento militar, para não comprometer a qualidade dos dados populacionais.

O critério de enumeração (de iure ou de facto) tem implicações importantes para a forma de execução do censo. Por exemplo, o censo chileno de 2012 fracassou em parte porque foi mudado o critério de enumeração de *de facto* para *de iure*, sem a devida preparação. O critério *de iure* é geralmente mais relevante do ponto de vista administrativo, mas o critério *de facto* é mais exequível, principalmente em países com populações pouco alfabetizadas ou com muitos migrantes. Classificar as pessoas pela sua residência habitual significa que uma pessoa que na data do censo estava temporariamente hospedada num domicílio (agregado familiar) diferente do seu deveria ser declarada pelos demais residentes da sua casa e ser excluída da enumeração da casa onde se hospedou. Isso pode ser difícil se não ficou ninguém para ser enumerado na sua residência habitual. O conceito de residência habitual pode ser problemático em certos casos, como o de estudantes, filhos de pais divorciados e trabalhadores itinerantes que moram em endereços diferentes em épocas distintas. Em países com muita migração, estabelecer com precisão quem é um migrante temporário, que em algum momento deve voltar para a sua residência habitual, e quem migrou definitivamente pode ser altamente arbitrário. Por isso os países africanos geralmente fazem censos baseados no critério *de facto*. Nos EUA, Canadá, Costa Rica, México, Israel, Filipinas (apesar do grande número de migrantes internacionais) e na grande maioria dos países europeus, incluindo Portugal, geralmente usa-se o critério *de iure*. Desde 1980 os censos brasileiros também seguem esse critério. Entre 1940 e 1970 os censos brasileiros foram realizados segundo ambos os critérios, como até hoje acontece em alguns outros países (Burquina Faso, Camarões, Irlanda, Mali, Moçambique, Zâmbia). O censo português de 2011 também usou ambos os critérios, mas em 2021 será usado apenas o critério *de iure* (INE Portugal, 2018).

A responsabilidade direta pelo preenchimento dos dados de cada pessoa pode caber ao recenseador, por meio de entrevistas individuais ou com informantes qualificados (*canvasser method*), ou ao responsável pelo domicílio (agregado familiar), que devolve o questionário preenchido ao recenseador, o envia por correio ou via internet³ (*householder method*). O primeiro método ainda é o mais comum nos países em desenvolvimento, onde muitos informantes têm baixo nível de escolaridade ou não têm um acesso fácil ao correio ou à internet. No caso de censos baseados inteiramente em registros existentes os indivíduos não fornecem a informação pessoalmente, mas esta é gerada a partir de dados que já foram fornecidos para outros propósitos.

Critério 7: Disponibilidade dos resultados dentro de prazos compatíveis

Outro critério é a exigência de que os resultados sejam publicados dentro de um prazo razoável, suficientemente curto para manter a sua relevância para o planejamento, tanto governamental como do setor privado. No passado houve vários episódios de censos, cujos resultados foram publicados com tanto atraso que perderam grande parte da sua relevância, como o censo brasileiro de 1960 cuja publicação só se terminou 15 anos mais tarde. Portugal teve um problema

³ Tanto o Censo brasileiro de 2010 como o português de 2011 ofereceram a opção de preenchimento e transmissão do questionário via internet pelos moradores, sem uma visita do enumerador ao domicílio. Embora esta modalidade de preenchimento seja mais barata do que uma visita tradicional de um recenseador, ela também tem os seus riscos. A revista *The Economist* relata que no Censo do Reino Unido de 2011 houve 177 mil respondentes on-line que identificaram a sua ocupação como “cavaleiro Jedi”. Em 2016, o sítio web da Agência Australiana de Estatística (ABS) ficou sobrecarregado e falhou, deixando milhares de respondentes sem possibilidade de entregar os seus questionários.

semelhante com o censo de 1970. Um exemplo mais recente é o censo de Camarões de 2005, cujos resultados – devido a problemas políticos – nunca foram completamente analisados e divulgados. Entretanto, de uma forma geral a tendência em décadas recentes tem sido que os prazos de publicação dos resultados censitários vêm diminuindo, com a ajuda dos avanços na informática e a tecnologia de coleta (recolha) e processamento de dados, especialmente o uso de Dispositivos Móveis de Coleta (DMC, Personal Digital Assistants ou PDAs, em inglês) ou tablets ou ainda a resposta “on-line”.

Critério 8: Disponibilidade da informação nos níveis mais desagregados considerados relevantes

Como o censo levanta informação sobre todos os indivíduos num determinado território e não só sobre uma amostra, é o instrumento ideal para obter dados sobre áreas geográficas locais ou sobre grupos sociais muito detalhados (por exemplo, mulheres não alfabetizadas de 15-19 anos que vivem sozinhas, em domicílios unipessoais). Mas de pouco serve esta característica se a informação só é divulgada para o país como um todo e algumas das divisões administrativas principais.

É de notar que, embora as Nações Unidas façam recomendações detalhadas sobre o conteúdo dos censos (United Nations, 2017 a), a lista de critérios acima não inclui um critério referente ao conteúdo mínimo de perguntas que precisam ser feitas num levantamento para que seja considerado um censo. Isso significa que as contagens rápidas, do tipo que o Brasil executou em 1996 e 2007⁴, em que só se fazem as perguntas básicas para poder identificar os domicílios (agregados familiares) e os indivíduos que moram neles (idade, sexo, relação com o chefe, estado civil), também podem ser consideradas censos. Indonésia é um dos países que sistematicamente fazem contagens das suas populações nos anos terminando em “5”, entre os dois censos.

O censo convencional, por meio de entrevistas a cada domicílio (agregado familiar), é uma operação custosa que exige uma logística de dimensões quase militares. Para o censo brasileiro de 2010 foram necessários mais de 190 mil recenseadores e um orçamento de mais de 5 bilhões de dólares para percorrer os mais de 8,5 milhões de quilômetros quadrados do território nacional, com um custo aproximado de quase 90 dólares por domicílio. O censo de Moçambique de 2017 teve um orçamento de 73,5 milhões de dólares, ou seja, 2,65 dólares por pessoa enumerada. O censo de 2010 dos EUA custou 92 dólares por domicílio enumerado. Como já se mencionou acima, desde os anos 70 do século passado há um número crescente de países que realizam os seus censos administrativamente. Com base em informações preliminares sobre o planejamento para a rodada de 2020, a tendência é que, dos 48 países da Comissão Econômica das Nações Unidas para Europa (UNECE), 14 realizem o Censo com base exclusivamente em registros administrativos, 12 países adotem o modelo de transição (25%) e apenas 22 ainda façam o Censo tradicional (46%). Portugal considera a possibilidade de um censo administrativo para 2021. O procedimento consiste em consolidar a informação já contida em diversas bases de dados mantidas para outros propósitos, tais como

⁴ Devido a restrições orçamentárias, a contagem de 2007 ficou limitada a 5.435 dos 5.564 municípios, excluindo a maioria dos municípios com mais de 170.000 habitantes.

- Registro de população;
- Registro civil;
- Registro de endereços, prédios, construções e moradias;
- Registro de utilidades públicas;
- Registro da receita de impostos;
- Registro de empresas e negócios;
- Registro eleitoral;
- Registro de emprego e de pessoas buscando emprego;
- Registro de aposentadorias e seguro social;
- Registro de benefícios sociais (por exemplo, o Cadastro Único no Brasil, que se aplica ao Bolsa Família e outros benefícios); e
- Registro de estudantes.

De todos estes registros, o mais essencial é o registro de população que deve conter os dados básicos atualizados de todos os indivíduos residentes no país no momento do censo. São relativamente poucos os países que mantêm um registro de população (ver a seção 4.8 deste capítulo), mas sem ele um censo administrativo é praticamente impossível. O funcionamento do registro de população será discutido em mais detalhe abaixo. Outra condição, que talvez não seja absolutamente necessária, mas cuja ausência certamente dificulta o processo, é a existência de um número único de identificação de cada pessoa no país como o Cadastro da Pessoa Física (CPF) no Brasil, preferencialmente emitido automaticamente no momento do nascimento. Até 2018 o CPF no Brasil tinha que ser solicitado, mas desde então a emissão é feita automaticamente, como parte do registro do nascimento. Além disso, os registros administrativos em que o censo se baseia devem ser de alta qualidade, atualizados e passíveis de serem cruzados com outras fontes de dados.

Extrapolando a tendência de vários países europeus a adotar procedimentos alternativos na realização do censo de população, Coleman (2012) chega a falar no “ocaso do censo” no sentido convencional. As principais vantagens do censo baseado em cruzamento de dados administrativos são o baixo custo – consideravelmente menor do que o custo de um censo convencional – e o fato de não exigir a cooperação da população. Este último ponto é importante na medida que em vários países se verifica uma disposição cada vez menor por parte da população para fornecer informação em censos e inquéritos amostrais, apesar das penalidades legais que geralmente são previstas para a recusa ou para o fornecimento de informação falsa. A obtenção deste tipo de informação também se vê prejudicada pela redução dos tamanhos dos domicílios (agregados familiares) e a crescente

participação dos seus membros em atividades fora de casa, o que diminui a probabilidade de que o recenseador encontre alguém em casa para responder ao questionário⁵.

A principal desvantagem do censo administrativo é que só cobre informação que já está contida nos registros administrativos existentes. Como estes registros são mantidos para finalidades específicas, a sua definição da população nem sempre coincide com a do censo. Também há dados que costumam ser levantados nos censos, mas que não constam nos registros administrativos. Para acrescentar dados sobre temas novos é preciso usar outras estratégias, tais como o cruzamento com inquéritos por amostragem ou a aplicação de um formulário censitário curto para levantar as informações que não puderam ser obtidas por meio dos registros existentes. O uso ou não uso de fontes de dados adicionais, além dos contidos nos registros administrativos existentes, atualmente constitui a maior diferença metodológica dentro da categoria de censos realizados com base em registros administrativos.

Ainda existe um terceiro tipo de censo, conhecido como o “censo rolante”, no qual o universo é dividido em 5 partes, cada uma observada exaustivamente durante um ano, durante um ciclo de 5 anos. Até agora o único país que utiliza este modelo é a França que o aplicou pela primeira vez em 2004 (Sosa, 2004; Durr, 2012). Também houve tentativas para a aplicação desta metodologia na Colômbia e no Peru, mas como se trata de um procedimento complexo que exige muita preparação, os resultados não foram satisfatórios. Em teoria, o modelo tem como vantagem a diluição dos custos ao longo dos anos, mas além da sua complexidade ele não permite a caracterização simultânea do universo num momento específico de referência (o quarto da lista de critérios acima). A metodologia usa fontes administrativas para recalibrar o tamanho da população e o número de moradias no ano de referência, que é o ano central do último quinquênio.

4.3 UM POUCO DE HISTÓRIA DOS CENSOS

A história dos censos tem sido contada em vários livros (por exemplo, Dupâquier e Dupâquier, 1985: Cap. 3; Thorvalsen, 2018), com muito mais detalhe do que é possível fazer nestas breves notas. Devido à multiplicidade de critérios, não é fácil afirmar com certeza quando o primeiro censo foi realizado. Certamente, muito antes da era cristã os impérios do Egito, Babilônia, Palestina e Roma já realizavam contagens periódicas das suas populações para estabelecer as suas bases fiscais, de trabalhadores ou de soldados. Confúcio menciona um censo executado durante o reinado do rei Yao, em 2238 a.C. Entretanto, apenas os indivíduos relevantes destes pontos de vista eram enumerados: proprietários, chefes de família ou homens sujeitos ao alistamento militar. O censo romano, por exemplo, foi realizado de cinco em cinco anos durante quase oito séculos, mas se limitava à enumeração de cidadãos, para fins fiscais e militares. Da mesma forma, os raros levantamentos medievais, tais como o *Domesday Book*, contendo o levantamento de propriedades rurais, decretado em 1085 pelo Rei William I da Inglaterra, e a contagem domiciliar do *État des Subsidies* na França, de 1328, eram limitados no seu escopo. Parece, entretanto, que um levantamento exaustivo da população de Nürnberg (Alemanha) foi realizado em 1449.

⁵ Isso explica, por exemplo, por que os idosos frequentemente são sobrerrepresentados nos inquéritos (ver também a seção 4.9 deste capítulo).

Os levantamentos populacionais da Sicília, de 1501, e diversas outras regiões da Itália em anos próximos são considerados, às vezes, como os primeiros censos no pleno sentido da palavra. Outras experiências pioneiras são os levantamentos populacionais de Québec, no Canadá, a partir de 1666, e de diversas outras colônias francesas e inglesas na América do Norte. Islândia realizou um levantamento completo da sua população em 1703. No Brasil houve vários levantamentos regionais de população a partir de 1765 (ver Capítulo 15). Entretanto, a maioria dos estudiosos aponta o Censo sueco de 1749 como o primeiro que satisfaz quase todos os critérios considerados essenciais para um recenseamento moderno. Outros censos da época foram os da Noruega, em 1760, e da Dinamarca, em 1769. Os EUA foram o primeiro país a estabelecer uma rotina legal para a organização decenal de censos de população. Como foi mencionado no Capítulo 3, tal rotina baseava-se na própria Constituição, para garantir a proporcionalidade das vagas estaduais e da Câmara Federal com a população residente em cada distrito eleitoral. O primeiro Censo dos EUA foi realizado em 1790. Os Países Baixos fizeram um primeiro Censo em 1795 e Inglaterra e França seguiram em 1801⁶. Na América Latina, os primeiros censos após a Independência foram os da Colômbia (1825), Chile (1843) e Uruguai (1852). Considerando o período anterior a 1900 (já mostrado na Tabela 4.1), na América Latina os primeiros censos após a Independência foram os da Colômbia (1825), Chile (1843) e Uruguai (1852). Embora houvesse censos parciais anteriores, Portugal fez o seu primeiro censo geral em 1864, tendo depois, os censos de 1878 e 1890; Macau em 1866. Brasil teve censos em 1872 e 1890. Além dos censos gerais que ocorreram a partir de 1872, o Brasil também teve um censo do então Distrito Federal em 1906 e do Estado de São Paulo em 1934. A qualidade dos censos varia consideravelmente. Particularmente, os censos coloniais de diversos países africanos e asiáticos são de utilidade muito limitada.

Curto (1994) descreve a história dos primeiros censos em Angola, começando com a de Luanda em 1773. Um exemplo dos problemas desses censos pré-modernos na África, baseado nos censos de Angola no início do século XX, é mostrado no Capítulo 15. A história dos censos populacionais em Moçambique é descrita em Lemos (1987), Wagner (2007, 2009) e em INE Moçambique (1999). O Censo de 1930 foi o primeiro a enumerar toda a população do país; houve um Censo anterior em 1928, mas apenas da população não indígena. De 1930 a 1970, os censos foram realizados de dez em dez anos para toda a população e de cinco em cinco anos para a população não indígena e os resultados eram publicados em volumes separados para cada raça, indígena e não indígena, com a exceção de 1960 e 1970, onde, por causa da instabilidade política causada pelo início da guerra de libertação levada a cabo pela FRELIMO, houve apenas um volume para todas as raças (Arnaldo, 2007). O Censo de 1960 foi notoriamente subenumerado, particularmente nas Províncias (então Distritos) de Inhambane e Manica (Rita-Ferreira, 1972).

No Brasil, a qualidade dos censos também melhorou significativamente a partir de 1940, devido, em grande medida, ao desempenho pessoal de Giorgio Mortara, que chefiou a execução dos censos pelo IBGE, criado em 1936 como o sucessor da extinta Diretoria Geral de Estatística (Berquó e Bercovich, 1985). Algo parecido passou em Portugal, com a criação do INE em 1935. Mesmo assim, o frequente questionamento dos números populacionais levantados pelo IBGE por prefeitos e outros administradores locais, no intuito de reivindicar maiores parcelas

⁶ A Universidade de Essex, na Inglaterra, mantém um sítio web com diversos materiais históricos sobre os censos da Grã-Bretanha no endereço <http://chcc.arts.gla.ac.uk/>. Os Países Baixos também mantêm um sítio com informação histórica dos seus censos desde 1795 em <http://www.volkstelling.nl/en/> (em holandês).

de transferências da União, chamado de Fundo de Participação dos Municípios, mostra que a última observação do parágrafo em cima continua pertinente até hoje. Mesmo em países que já possuem uma longa tradição na organização de censos, os resultados são às vezes contestados, inclusive pela via judicial, como tem acontecido em várias cidades americanas, chegando em alguns casos a ordens judiciais para realizar uma recontagem da população local. No censo de 2010, o US Bureau of the Census estabeleceu um processo de rotina para atender às reclamações de autoridades locais (veja <http://www.census.gov/2010census/about/cqr.php> para mais detalhes). Foram recebidas 239 queixas deste tipo, a maioria das quais levou a revisões dos números oficialmente publicados.

4.4 O PROCESSO DO CENSO

O censo moderno consiste de três etapas principais. O *pré-recenseamento* consiste da criação das condições para que se possa realizar o trabalho de campo. Num censo convencional, não baseado no cruzamento de bases de dados existentes, uma das atividades mais importantes nesta fase é o levantamento cartográfico que resulta numa caracterização geográfica exata do setor que cada enumerador precisa visitar, de preferência com a localização de todos os domicílios (agregados familiares) a visitar. Nas palavras de um demógrafo francês (Clairin, 1981), “sem uma cartografia adequada, a coleta (recolha) de dados demográficos se faz no nevoeiro”. No Brasil os setores censitários (áreas de enumeração) de cada recenseador costumam ter um tamanho de 300 domicílios em áreas urbanas e 150 em áreas rurais, com uma variação de 50 para cima ou para baixo. Esta é uma atividade em que houve grandes mudanças tecnológicas nas últimas décadas, com a conversão de mapas convencionais de papel para sistemas de mapas digitais apoiados em Sistemas de Posicionamento Global (Global Positioning Systems ou GPS, em inglês) por satélite. Estas tecnologias não afetam só a preparação dos mapas em si, mas inclusive a forma como estes são usados no trabalho de campo (Charles, 2012; Charles e Marques de Sá, 2012; IBGE, 2013 a: Cap. 8; United Nations, 2008). O Brasil foi pioneiro no uso de DMCs como instrumento para a coleta (recolha) de dados censitários no campo, substituindo os questionários convencionais em papel. Na medida em que esses aparelhos tenham uma capacidade de posicionamento geográfico, eles podem guiar o recenseador e averiguar automaticamente na base cartográfica se este se encontra no lugar correto para realizar uma determinada entrevista, minimizando assim a necessidade por parte do recenseador para interpretar os mapas. Esta facilidade pode ajudar muito em reduzir a incidência de erros nos endereços dos domicílios (agregados familiares) entrevistados e de enumerações fora do setor assignado a cada enumerador.

Além da cartografia, é necessário desenvolver as seguintes atividades, em ordem cronológica:

- Desenho do questionário;
- Preparação das instruções para os entrevistadores;
- Organização do trabalho de campo: enumeradores, supervisores, recrutamento, capacitação;
- Campanha de comunicação;
- Coleta (recolha) de dados;

- Processamento de dados;
- Análise e divulgação de dados; e
- Avaliação.

A definição do questionário é delicada, pois a inclusão de uma única pergunta adicional pode causar um grande aumento nos custos, principalmente se as categorias de respostas possíveis são complexas, enquanto a omissão de uma pergunta essencial pode inviabilizar todo um trabalho de análise dos resultados. Além das muitas recomendações internacionais existentes para o desenho correto do questionário também costuma-se envolver os diferentes órgãos do governo interessados nos resultados, bem como organizações relevantes da sociedade civil. Outro aspecto que normalmente se decide já nesta fase preparatória é o formato do futuro *plano de tabulação* dos resultados e a sua forma de divulgação.

A fase da própria *enumeração* consiste no trabalho de campo e na sua supervisão. Como se mencionou anteriormente, uma das inovações que o Brasil introduziu no Censo de 2010 foi o uso de DMCs para o trabalho de campo. Esta alternativa tem muitas vantagens, como a estruturação automática da entrevista (por exemplo, saltando automaticamente as perguntas não aplicáveis), a verificação automática da consistência da informação, a supervisão “on-line” e a rapidez da transmissão dos resultados, uma vez que forem liberados pelos supervisores. Entretanto, a tecnologia pode não funcionar bem em áreas que não têm acesso fácil a energia elétrica ou onde não há cobertura de internet. A solução pode ser o uso de baterias adicionais e a recolha de dados off-line e a sua transmissão no momento em que houver acesso a internet. Também existe o risco de avarias ou extravio dos aparelhos. A maioria dos PALOP ainda usa questionários em papel, mas a nova tecnologia de DMCs já foi introduzida em vários países africanos, começando com Cabo Verde, Senegal, Côte d’Ivoire e São Tomé & Príncipe na rodada de censos de 2010. Atualmente há três centros de referência na África para a recolha digital que contam com o apoio do Brasil: Cabo Verde, Senegal e África do Sul. Uma boa supervisão é de importância primordial para garantir a consistência dos critérios e para evitar fraude. Muitas vezes a supervisão envolve visitar algumas casas para verificar a qualidade do trabalho dos enumeradores. Os supervisores também precisam certificar-se de que todos os domicílios (agregados familiares) tenham sido visitados e que aqueles onde não foram encontrados moradores na primeira tentativa sejam revisitados.

A fase da *pós-enumeração* envolve primeiro a captação eletrônica da informação e sua passagem para uma base de microdados que identifica a informação obtida de cada questionário. Se a informação foi submetida via internet pelos próprios moradores ou se foi recolhida com o uso de um DMC e transmitida ao centro de processamento, não é preciso montar uma operação especial para este fim, mas nos casos (ainda) mais comuns em que a informação chega na forma de questionários em papel, a sua passagem para o formato eletrônico pode ser trabalhosa. Entretanto, esta situação está mudando rapidamente. Ainda existem países (por exemplo, Moçambique em 1997 e 2007) onde este trabalho é feito por digitadores que passam cada questionário a um arquivo (ficheiro) de computador, mas atualmente a prática mais comum é a leitura automática por um aparelho de leitura ótica (Moçambique adotou esse sistema em 2017). As experiências internacionais com essa prática têm sido variadas. A leitura ótica certamente pode poupar tempo, mas ela também exige

muitos cuidados especiais e é sensível a certos tipos de erros. Por exemplo, o questionário deve ter um formato especial que possibilita a leitura ótica e não pode ser dobrado, rasgado, manchado ou excessivamente úmido. Os enumeradores também precisam ter o cuidado de escrever de uma forma padronizada e bem legível, por exemplo para evitar a confusão entre os números “5”, “6” e “8” ou entre as letras “l” e “i”. Na prática, muitos países têm encontrado que estas precauções não são tão fáceis de implementar e que a incidência de erros pode ser elevada caso a autoridade censitária não tenha uma infraestrutura informática própria ou trabalhe com um parceiro do setor privado com muita experiência no processo.

4.6 AVALIAÇÃO DA QUALIDADE DOS CENSOS

O próximo passo é que os dados precisam passar por uma revisão crítica, para detectar e corrigir inconsistências e omissões ao nível de indivíduos ou domicílios (agregados familiares). Por exemplo, se os dados indicam a existência de uma mulher de 13 anos com 4 filhos, é provável que haja um erro de digitação ou leitura ótica e que efetivamente ela tenha 43 anos ou que realmente tenha 13 anos, mas só 1 filho. A inspeção do questionário ou de outros dados da mesma mulher (por exemplo, a idade do cônjuge) podem esclarecer qual é o problema. Normalmente estes processos funcionam melhor quando executados de forma automatizada, com base em critérios explícitos, devido à tendência de operadores humanos a introduzir preferências e vieses pessoais. Posteriormente, os dados passam para o processamento e a publicação, normalmente na forma de tabulações. É possível fazer este trabalho usando softwares de uso geral como EXCEL, SAS ou “R”, mas também existe um software específico para o processamento de dados censitários, desenvolvido e distribuído gratuitamente pelo US Bureau of the Census, que se chama CSPro.

Além da publicação dos resultados, uma das últimas etapas no processo censitário é a avaliação da qualidade da informação. De um modo geral, existem três estratégias para esse fim que não são mutuamente excludentes, de modo que eventualmente podem ser aplicadas de forma conjunta:

1. A avaliação da *consistência interna* do censo ao nível agregado, por exemplo para verificar se o número de nascimentos declarados no(s) ano(s) anterior(es) ao censo é ou são consistente(s) com o número de crianças nas idades correspondentes. Em muitos censos existe uma tendência para a omissão de crianças menores de 1 ano que pode ser detectada desta forma.
2. A comparação com o censo anterior ou com inquéritos de diversos tipos no passado recente.
3. A realização de uma *Pesquisa de Avaliação*, também chamada *Inquérito de Pós-Enumeração* ou *Inquérito de Cobertura*.

O objetivo é avaliar tanto os erros de *cobertura* (pessoas ou casas que deixaram de ser contadas ou que foram contadas mais de uma vez) como os erros de *conteúdo* (pessoas ou domicílios que foram contados corretamente, mas com erros na informação recolhida). Algumas das técnicas envolvidas na aplicação da primeira e segunda estratégia serão discutidas no Capítulo 16. A

terceira estratégia é potencialmente a mais rigorosa, mas a sua aplicação é dispendiosa e exige um alto nível de competência técnica, além de precauções especiais na execução do próprio censo. Por este motivo, muitas Pesquisas de Avaliação fracassam total ou parcialmente. Segundo Durr (2010), dois terços de 134 países que responderam a um questionário da Divisão de Estatística das Nações Unidas sobre a rodada (ronda) dos censos de 2010 indicaram ter realizado uma pesquisa de avaliação. Na América do Sul e na África era 78% e 79%, respectivamente. Alguns desses inquéritos (por exemplo, África do Sul, Brasil, Uganda, Zâmbia e Zimbábue) foram publicados, mas outros não, por não terem chegado a uma conclusão satisfatória. No Quênia, por exemplo, tanto em 1999 como em 2009 a publicação foi suspensa porque não foi possível chegar a resultados consistentes. O mesmo aconteceu nos Censos de Timor-Leste em 2010 e em 2015. Cabo Verde em 2000 e 2010 também fez Pesquisas de Avaliação, mas os resultados não foram publicados. O Censo de Moçambique de 2017 contou com um Inquérito de Cobertura do Censo, que foi executado entre 16 e 30 de setembro, um mês e meio depois do próprio Censo, com uma amostra de 43.512 domicílios (agregados familiares).

A forma como se conduz uma Pesquisa de Avaliação é a seguinte: Pouco tempo depois do censo – quanto mais rápido melhor – se seleciona uma amostra de setores censitários (áreas de enumeração) (que geralmente são a unidade geográfica que corresponde ao conjunto de domicílios (agregados familiares) enumerados por um único enumerador) e nesta amostra se realiza a contagem outra vez, usando os mesmos procedimentos do censo, mas com um questionário reduzido e com enumeradores novos, que não tiveram nenhum envolvimento com a enumeração destas áreas no censo. Na avaliação do Censo brasileiro de 2010 se selecionaram 1,3% das áreas de enumeração e dentro delas uma amostra de 10% dos domicílios (agregados familiares). Outros países reenumeram todos os domicílios (agregados familiares) dentro das áreas de enumeração (setores censitários) amostradas. A análise posterior consiste em comparar os resultados da amostra reenumerada com os resultados originais do censo. Esta comparação deve ser feita não em termos de números agregados, mas para todas as pessoas e domicílios individuais, para poder estabelecer quem foi enumerado tanto no censo como na pesquisa (inquérito), quem foi enumerado só no censo em que foi enumerado só na pesquisa (inquérito). Isso exige um controle sofisticado do processo de enumeração no censo, para evitar que certos domicílios (agregados familiares) sejam enumerados fora da área à qual pertencem, e um bom sistema de identificação dos mesmos, de preferência com endereços físicos de rua, coordenadas geo-referenciadas por GPS ou outros identificadores de boa qualidade. Quando o país não dispõe de um sistema bem estabelecido de endereços postais ou outros bons identificadores e o censo não usou GPS, a execução de uma Pesquisa de Avaliação pode ser bastante difícil.

Tanto o Brasil como Portugal têm práticas de Pesquisas de Avaliação bem estabelecidas. Na rodada (ronda) de Censos de 2000, Argentina, Bolívia, Chile, Cuba, Equador, Paraguai e Venezuela também fizeram inquéritos desse tipo. As estimativas dos erros de omissão nos Censos brasileiros de 1970 (a primeira vez que se realizou uma Pesquisa de Avaliação), 1980, 1991, 2000 e 2010, segundo as respectivas pesquisas, foram as seguintes:

Tabela 4.2: Erros de cobertura dos censos brasileiros segundo as respectivas Pesquisas de Avaliação da Cobertura e correções aplicadas pela Divisão de População das Nações Unidas nas estimativas publicadas em 2019

	Pesquisas de Cobertura	Correções Aplicadas pela ONU
1970:	7,3%	2,6%
1980:	4,3%	1,9%
1991:	4,7%	3,6%
2000:	5,8%	3,4%
2010:	8,6%	1,7%

Fontes: IBGE, Pesquisas de Avaliação de Cobertura e Divisão de População das Nações Unidas, Revisão de 2019.

A metodologia de avaliação da cobertura não foi a mesma em todos os censos. Os números apresentados acima se referem às taxas de omissão de pessoas moradoras em domicílios (agregados familiares) particulares permanentes ocupados que foram omitidos nos censos. Mas em alguns casos é preciso acrescentar a isso a omissão de membros dos domicílios (agregados familiares) enumerados tanto no censo como na pesquisa de avaliação. Em função dessa subenumeração adicional, Carvalho e Campos (2006), por exemplo, aceitaram a estimativa dada acima para o Censo de 2000 para estimar a migração internacional entre 1991 e 2000, mas aumentaram a população de 1991 em 9,08%, em vez de 4,7%. Também há uma variação regional significativa. Segundo a pesquisa de 2013, o erro de cobertura variou de 4,1% na área urbana da Região Sul a 17,0% na área rural da Região Norte (IBGE, 2013 b). Os organismos internacionais têm aplicado correções bem mais modestas do que as indicadas pelas Pesquisas de Avaliação. A segunda coluna da Tabela 4.2 mostra as correções usadas pela Divisão de População da ONU, na sua Revisão de 2019. O CELADE estima a omissão do Censo de 2010 em 2,4% (Naciones Unidas, 2013).

A estimativa de omissão do Censo português de 2011 foi de 2,5%, segundo o relatório do Inquérito do Censo, mas esta porcentagem é compensada por 5,1% de inclusões errôneas, de modo que o resultado líquido foi uma sobre-enumeração de 2,6% (para mais detalhes sobre os censos portugueses, veja Pinheiro, 2003 e Leal, 2013). Esses números demonstram que os censos estão sujeitos a erros não desprezíveis, mesmo no Brasil e Portugal, onde a prática censitária é bem desenvolvida. Mesmo num censo com pouca subenumeração global, o erro pode ser muito maior em certas subpopulações específicas como a população favelada ou entre os homens de 20-29 anos. O Inquérito de Cobertura do Censo de Angola de 2014, realizado em 320 setores censitários, encontrou uma omissão de 5,6% ao nível nacional (5,9% em Luanda, 6,2% no resto urbano e 4,8% na área rural). Estes resultados foram usados para ajustar a população, que passou de 24.344.839 para 25.789.024⁷. Entretanto, não houve uma publicação da metodologia e dos resultados detalhados. Os resultados do Inquérito Pós-Censitário da Guiné-Bissau em 2009 tampouco foram publicados, exceto o erro total de cobertura (4,7%) e as suas variações regionais (de 3,8 a 6,1%). O erro de cobertura do Censo de Moçambique de 2017 foi estimado em 3,7% (comparado com 2,1% em 2007 e 5,1% em 1997), com um máximo de 5,7% na Província de Niassa. Entretanto, esses números representam apenas os números de pessoas que, no Inquérito de Cobertura, responderam que não foram enumeradas no censo; não houve um exercício sistemático de comparação da informação individual no censo e no inquérito.

⁷ Este nível de subenumeração parece baixo, visto que o Censo da África do Sul de 2011, considerado o melhor na África Subsaariana, oficialmente admitiu uma subenumeração de 14,6 % (Statistics South Africa, 2012).

Além dos erros de cobertura, também existem erros de conteúdo que podem ser revelados pelas inconsistências entre as respostas dadas no censo e na pesquisa de avaliação. Alguns erros são raros, por exemplo as inconsistências entre o sexo declarado do entrevistado, mas outros são bem mais comuns, principalmente no que se refere à informação sobre a atividade econômica. No Censo de Portugal de 2011, por exemplo, houve 17,8% de inconsistências nas respostas dadas à pergunta sobre o modo de exercício da profissão (condição na atividade) e 17,4% de inconsistências na informação sobre a procura de trabalho entre pessoas sem trabalho (relatório do Inquérito: Quadro 4.20).

Apesar de todos os esforços normalmente empenhados na organização dos censos, erros mais graves do que os indicados acima acontecem. Historicamente o Brasil enfrentou problemas sérios com o seu Censo de 1900 (que foi seriamente deficiente), com o de 1920 (cujos números superaram a realidade em 50% na Região Norte) e com o Censo de 1960. Inclusive, é preciso reconhecer que, apesar dos avanços tecnológicos, a qualidade dos censos nem sempre melhora com o tempo. Novas tecnologias, como a introdução da leitura ótica, podem poupar tempo e evitar determinados tipos de erros, mas também podem gerar novos problemas. Muito depende também da situação política do país que pode repercutir negativamente na preparação e execução do censo. Como foi mencionado acima, o censo brasileiro de 1980 teve uma das melhores coberturas da história do país, mas de uma forma mais geral a qualidade dos censos latino-americanos da década de 80 foi inferior à da década anterior (Guzmán, 1989). Alguns censos latino-americanos recentes até foram rejeitados por apresentarem problemas inadmissíveis. Em 2005, Peru fez uma tentativa mal-sucedida de adotar a metodologia francesa do “censo rolante” e teve que refazer a enumeração em 2007. Depois de uma auditoria do Censo chileno de 2012, executada em 2014, o governo teve que refazer o censo em 2017 (Carvajal e Henriques, 2013). Também houve problemas significativos com a cobertura do Censo paraguaio de 2012, com uma omissão que o CELADE estima em 26,0% (Naciones Unidas, 2013).

4.6 O CONTEÚDO E OS RESULTADOS DO CENSO

Como foi mencionado antes, parte dos resultados publicados de um censo pode ser obtida com base em amostras. Isso pode acontecer por diferentes motivos, mas a razão principal é para reduzir o trabalho de aplicação e de processamento do questionário completo que pode ser muito extenso. Nos censos brasileiros de 1960 em diante, por exemplo, aplicou-se um formulário resumido (chamado de “questionário básico” ou “curto” ou “questionário da não amostra”) a todos os domicílios, suplementado com perguntas adicionais que foram levantados apenas para uma amostra (o “questionário longo” ou “questionário da amostra”). As frações amostrais têm variado no tempo. Até 1980, a amostra cobria 25% dos domicílios (agregados familiares). Nos Censos de 1991 e 2000 aplicou-se uma fração de 10% de cada setor nos municípios com mais de 15.000 habitantes e de 20% naqueles menores. Nos recenseamentos posteriores, cálculos mais complexos aperfeiçoaram a amostra em função do tamanho populacional dos municípios. A circunstância de que parte da informação do censo provém de uma amostra coloca o problema da *expansão da amostra*, que precisa ser feita de uma forma consistente com os resultados gerais, inclusive a não amostra, naquelas variáveis que são investigadas no questionário curto. Para garantir esta consistência existem procedimentos específicos cuja descrição vai além do escopo deste livro. Para maiores detalhes, o leitor pode consultar a metodologia do Censo de 2010 (IBGE, 2005, 2013 a).

Há vários outros países que seguem a prática de questionários curtos e longos. Na rodada (ronda) dos Censos de 2010, Argentina, Bangladesh, Canadá (2006), China (inclusive Hong Kong), Etiópia, Itália, Jamaica, a República de Coreia, México, Nepal, Porto Rico, a Federação Russa, Sudão, Tanzânia e Vietnã se valeram deste recurso, com frações amostrais variadas. Os EUA usaram esta metodologia de 1970 até 2000, mas em 2010 se limitaram a um questionário curto, com apenas dez perguntas. A informação antes recolhida no questionário longo doravante será obtida por meio de um inquérito amostral, a American Community Survey (ACS). O processo resultante às vezes é chamado um “censo misto”. Embora pareça improvável que o Brasil possa abrir mão das informações do censo na mesma medida que os EUA, pelo menos no futuro previsível, a prática de redução do questionário do censo e a procura de outras fontes de informação também está acontecendo no Brasil e outros países em desenvolvimento.

Os censos variam consideravelmente em termos da quantidade de informação solicitada, sendo que o Brasil tem um questionário muito extenso comparado com outros países⁸. Julgando pelo número de perguntas numeradas (um critério algo discutível porque as perguntas variam em grau de complexidade e alguns são compostos), o questionário longo do Censo do Brasil de 2010 teve 32 perguntas sobre o domicílio e 72 individuais. Em comparação, o Censo português de 2011 só teve 36 perguntas individuais, igual ao número de Moçambique em 2007, o Censo de Angola de 2014 teve 39, Cabo Verde em 2010 teve 43, Guiné-Bissau em 2009 teve 38, Macau em 2006 teve 28 e Timor-Leste em 2015 teve 54. O questionário do censo brasileiro também é mais extenso do que no Censo mexicano de 2010, que teve 41 perguntas individuais, ou no Censo argentino, que teve 39. Como se pode ver, o padrão típico é de 30-50 perguntas individuais, além de aproximadamente 30 perguntas sobre o domicílio (agregado familiar).

Embora não exista uma norma rigorosa sobre o mínimo de informações que o censo deve levantar, o manual das Nações Unidas (2017: 175-176) considera os seguintes itens como centrais (“core”, em inglês):

1. Nome e sobrenome (só para identificação do questionário, nunca para publicação);
2. Idade e sexo⁹;
3. Relação de parentesco com o chefe ou pessoa responsável do domicílio ou da família ou com outras pessoas de referência no domicílio (agregado familiar);
4. Estado civil e/ou conjugal;
5. Ocupação e demais características econômicas;
6. Produção para consumo próprio;

⁸ Jannuzzi (2018 a) compara o questionário do Censo brasileiro de 2000 com os censos dos EUA, Itália e México. O número de perguntas nos censos mexicanos é parecido com os brasileiros, mas os outros dois questionários são consideravelmente mais curtos. Para um resumo de todas as perguntas feitas nos censos brasileiros até 2000, ver Oliveira e Simões (2005: Quadro 1).

⁹ A grande maioria dos censos distingue apenas entre o sexo masculino e feminino, mas alguns, como o da Índia, em 2011, oferecem uma alternativa “Outro”, para acomodar a categoria de pessoas transgênero. Não se trata aqui de classificar o conceito de “gênero” que é diferente de “sexo” na medida em que descreve os papéis sociais atribuídos a pessoas de cada sexo.

7. Alfabetização, assistência escolar e nível de escolaridade;
8. Lugar de nascimento e nacionalidade;
9. Residência anterior e tempo de residência no lugar atual;
10. Lugar de residência numa data fixa do passado;
11. Condição de deficiência;
12. Filhos tidos em toda a vida; filhos sobreviventes e data de nascimento do último nascido vivo (nado vivo);
13. Óbitos no domicílio (agregado familiar) nos últimos 12 meses;
14. Residência habitual (*de iure*) ou lugar de enumeração (*de facto*) (de preferência ambos).

Naqueles censos que têm um questionário curto e um longo, estas são as perguntas que tipicamente aparecem no questionário curto, embora haja certa variação entre países também. Por exemplo, muitos países não têm uma pergunta sobre raça e/ou etnia, mas aqueles que a têm, como os EUA e o Brasil, geralmente a consideram suficientemente importante para incluí-la no questionário curto. No Brasil ela foi incluída já em 1872. O Brasil inclusive pergunta o rendimento no questionário básico. Além destas perguntas sobre os moradores, normalmente se pergunta também sobre a moradia: condição de ocupação, tipo de construção, área construída ou número de cômodos (divisões), materiais usados na construção e acesso a serviços básicos. No demais, os conteúdos variam entre países. A Divisão de Estatística das Nações Unidas mantém uma página web com os questionários de quase todos os censos recentes no mundo em <http://unstats.un.org/unsd/demographic/sources/census/censusquest.htm>.

4.6.1 Dados básicos de identificação

Além de perguntar a idade (completa ou presumida) da pessoa, os censos brasileiros desde 1940 (salvo o de 1960) também perguntam a data de nascimento (pelo menos mês e ano). Como a data de nascimento é fixa enquanto a idade varia no tempo, esta informação pode ser mais confiável. Ela também está menos sujeita ao arredondamento e fornece um critério independente para avaliar a consistência da declaração das idades. Em 1991 perguntou-se, além disso, qual era o grupo etário da pessoa: 0-4, 5-9 ou 10+. Em 2010 não se perguntou o dia do nascimento, só o mês e ano, mas para 2021 a intenção é voltar ao formato completo. Existem outros países, como Angola (2014), Cabo Verde (2010) e Guiné-Bissau (2009), que adotam a prática de perguntar ambos, mas a maioria pergunta um ou outro. O Censo de Portugal de 2011 e o de Macau, de 2006, perguntaram só a data de nascimento, enquanto Moçambique, em 2007, e Timor-Leste, em 2015, só perguntaram a idade.

Cada questionário do censo corresponde a um domicílio. Um domicílio pode consistir de uma só pessoa ou então pode ser um grupo de pessoas, não necessariamente consanguíneas, que compartilham uma mesma moradia e um mesmo orçamento doméstico para providenciar comida

e outras necessidades básicas. Os censos de Angola e Moçambique usam o termo “agregado familiar” em vez de “domicílio”, razão pela qual este livro usa ambos os termos. Todos os censos perguntam qual é a relação de cada pessoa com o (a) chefe do domicílio (agregado familiar). O que varia é a forma como se define o conceito de chefe (responsável, pessoa de referência) e o grau de detalhe com que são distinguidas as categorias de relações. Alguns dos censos brasileiros também têm a particularidade de distinguir entre as diferentes famílias que podem compor um domicílio e conseqüentemente a relação de cada pessoa com o(a) chefe da família. Maiores detalhes sobre estas questões serão discutidos no Capítulo 13.

Alguns censos perguntam se o nascimento da pessoa foi registrado e/ou se ela tem certidão de nascimento. No Brasil esta pergunta foi incluída pela primeira vez no Censo de 2010 e mostrou que 97,25% das crianças menores de 1 ano tinham uma certidão de nascimento ou tinham sido registradas em hospitais; 93,94% tinham sido registradas em cartório. Castanheira e Kohler (2015) usaram estas percentagens para corrigir os dados do SINASC (ver no próximo capítulo) e do Registro Civil para 2010. A pergunta também foi feita no Censo de Angola, de 2014, e no Censo de Timor-Leste, de 2015 (só para crianças até 5 anos). Os Censos de Moçambique de 2007 e 2017, além de perguntar se o nascimento da pessoa foi registrado, perguntaram sobre as razões do eventual não registro. A Tabela 4.3 mostra os resultados para o Censo de 2017.

Tabela 4.3: Percentagens de pessoas menores de 5 anos sem registro de nascimento ou status desconhecido, por sexo, área de residência e motivo, segundo o Censo de Moçambique (2017)

		Urbano		Rural	
		Homens	Mulheres	Homens	Mulheres
Status desconhecido		6,04	6,17	6,63	6,80
NÃO REGISTRADO	Fica longe	2,22	2,24	14,81	14,73
	Não é importante	0,43	0,41	0,49	0,48
	Falta conhecimento	1,64	1,66	5,55	5,56
	Burocracia	1,67	1,66	1,93	1,90
	É caro	2,34	2,36	4,27	4,24
	Falta tempo	14,56	14,51	8,70	8,75
	Faltam documentos	5,20	5,20	3,39	3,37
	Ainda é muito novo	7,26	7,20	7,63	7,69
	Outros	0,96	0,94	0,75	0,72
	Total com problemas		42,32	42,35	54,15

Fonte: Resultados Definitivos do IV Recenseamento da População e Habitação 2017 de Moçambique.

Além da pergunta padrão sobre o estado civil ou marital/conjugal, existem outras perguntas que são feitas em alguns censos. Em países como o Brasil, onde muitas uniões não são formalizadas, convém esclarecer de que tipo de união se trata, algo que é feito nos censos brasileiros desde 1960. A pergunta também foi feita no Censo da Guiné-Bissau, de 2009, e em São Tomé & Príncipe, em 2012. Outra pergunta se refere à data ou à idade que a pessoa tinha quando se casou ou uniu pela primeira vez. Essa pergunta pode ser importante em países onde muitas mulheres são forçadas ou induzidas a casar-se muito jovens, muitas vezes com homens muito mais velhos. Mas

trata-se de uma pergunta difícil de administrar em países onde muitas das uniões são informais. No Brasil ela foi incluída nos Censos de 1890, 1960 e 1991, mas não nos demais. Ela também foi feita no Censo da Guiné-Bissau, de 2009. No Censo português de 1991 e no Censo brasileiro do mesmo ano perguntou-se não só a data do primeiro casamento (também união, no caso do Brasil), mas também do último, o que possibilita averiguar se a pessoa esteve casada/unida mais do que uma vez e com que idades. Para maiores detalhes, veja a seção 12.3 do Capítulo 12.

Conforme a situação de cada país, o interesse em determinadas informações evidentemente varia. Por exemplo, não todos os países fazem perguntas sobre a religião, mas em países onde as Igrejas recebem apoio financeiro do governo (por exemplo, alguns países do Caribe) esta informação é considerada altamente relevante. Mesmo no Brasil, ela evoca grande interesse por parte das comunidades religiosas (ver, por exemplo, Alves, Barros e Cavenaghi, 2012; Alves et al., 2017; CERIS, 2002; Lopes, 2013). Por outro lado, países com minorias religiosas muito pouco expressivas ou onde a questão religiosa não tem maior transcendência social e política, como Itália, Colômbia, Venezuela e Cuba, não levantam informação sobre este item. Os países de língua portuguesa, com a exceção de Macau, sim fazem a pergunta, embora em Portugal, onde a pergunta é feita desde 1900, o entrevistado não seja obrigado a respondê-la. No Brasil ela só não foi feita no Censo de 1920.

4.6.2 Dados sobre migração

Quase todos os censos contêm alguma informação sobre os movimentos migratórios, mas dependendo da situação específica do país o conteúdo específico e o grau de detalhe da informação podem variar. O Capítulo 11 discutirá mais detalhadamente quais são os méritos de diferentes tipos de perguntas e como elas podem ser analisadas para gerar conhecimentos sobre as diferentes formas de migração. Aqui basta listar as perguntas mais comuns e a sua inclusão em censos diferentes. À primeira vista pode parecer um pouco estranho que a migração se meça por meio do censo, já que se trata de uma quantidade de fluxo e os censos são feitos principalmente para investigar a situação da população num determinado momento do tempo. A justificativa é que existem poucas fontes alternativas para esta informação, já que o registro civil não cobre a migração ou apenas capta uma parte muito pequena da informação relevante. As perguntas feitas nos países de língua portuguesa, além da cidadania e do lugar atual de residência, são as seguintes:

- a) Naturalidade ou lugar de nascimento (especificado em diferentes graus de detalhe geográfico).
- b) Tempo de residência no lugar de residência atual, onde “lugar de residência”, no caso do Brasil, pode ser a Unidade da Federação (UF), o município ou ambos.
- c) Último lugar de residência anterior (também especificado em diferentes graus de detalhe geográfico).
- d) Lugar de residência (incluindo outro país) 1 e/ou 5 anos antes da data do censo ou em alguma data significativa, outra vez com as mesmas distinções geográficas.

- e) A maioria dos países, inclusive os de língua portuguesa, pergunta a nacionalidade/cidadania da pessoa.
- f) No caso de estrangeiros ou pessoas não nascidas no país, geralmente há uma pergunta de continuação sobre o ano de chegada no país.
- g) Uma pergunta muito particular para o censo brasileiro, introduzida em 1980, mas retirada em 2000, verifica se a pessoa alguma vez morou na outra zona (rural ou urbana) do mesmo município onde reside atualmente.
- h) Uma outra pergunta pouco comum, mas que foi feita no Censo de Timor-Leste, em 2015, é a razão da migração. Portugal pretende incluir esta pergunta no Censo de 2021.
- i) Vários países, especialmente na Europa, entre os quais Portugal, perguntam se a pessoa já residiu fora do país alguma vez durante um período de pelo menos um ano. Esta pergunta é particularmente relevante em países com muita imigração ou com muita emigração temporária.
- j) No caso de uma resposta afirmativa em i), em seguida se pergunta qual foi o país onde residiu?
- k) No caso de uma resposta afirmativa em i), quando imigrou ou retornou? Desde o Censo de 1991, o Brasil também faz esta pergunta, mas só para estrangeiros residentes. Cabo Verde (2010) perguntou o ano da entrada da pessoa no país.
- l) Outra forma de abordar o tema de emigração é com a pergunta se alguém que já fez parte do domicílio (agregado familiar) atualmente mora no exterior. Esta pergunta foi feita no Censo brasileiro de 2010 e também costuma ser parte dos censos em vários outros países latino-americanos, mas ela foi retirada do Censo brasileiro de 2021.
- m) Se a resposta a l) for afirmativa, geralmente se pedem maiores detalhes, como o nome, idade e sexo da pessoa, para onde foi e quando partiu.
- n) Os censos brasileiros desde 1970, com a exceção do Censo de 1991, têm feito a pergunta se a pessoa trabalha ou estuda em outro município, para poder estimar a mobilidade pendular.
- o) A mesma pergunta também é feita para deslocamentos relacionados com o trabalho. Portugal em 2011 juntou esta pergunta com a anterior.
- p) Alguns países também perguntam se o domicílio (agregado familiar) recebe remessas. Tecnicamente não é uma pergunta de migração, mas está estreitamente associada à migração.
- q) O Censo de Macau, de 2006, perguntou ainda se a pessoa fez alguma viagem de negócios fora do país nos últimos 7 dias.

As perguntas a-f são recomendadas nos *Princípios e Recomendações para Censos de População e Habitação* (United Nations, 2017 a). As demais perguntas são específicas para os países, embora algumas (l e m, por exemplo) sejam bastante comuns em países com uma emigração

significativa. Informação mais específica sobre as perguntas de mobilidade e migração feitas nos diferentes censos é apresentada no Capítulo 11.

4.6.3 Dados sobre fecundidade e mortalidade

Mais inesperado do que a medição da migração por meio do censo é que muitos censos também contêm perguntas sobre a fecundidade e mortalidade. Embora a fonte natural para esta informação seja o registro civil, este não funciona adequadamente em muitos países (ver o próximo capítulo). Por isso desenvolveram-se nos anos 60, 70 e 80 diversas técnicas demográficas que permitem estimar esta informação a partir do censo. Como funcionam estas técnicas será explicado em mais detalhe no Capítulo 23. As perguntas feitas sobre estes temas são basicamente as seguintes:

- a) Número de filhos nascidos vivos (independentemente da sua sobrevivência) tidos (de preferência de cada sexo) ao longo da vida (fecundidade retrospectiva). Esta pergunta é feita a mulheres de 15-49 anos, 10-50 anos (Moçambique em 2017), todas as mulheres maiores de 12 ou 15 anos ou mais raramente a pessoas maiores de 15 anos de ambos os sexos.
- b) Os censos brasileiros estão entre os poucos que também perguntam sobre filhos nascidos mortos. A pergunta foi incluída pela primeira vez em 1940, retirada em 1950 e 1960, mas desde 1970 fez parte de todos os censos, sendo novamente cortada em 2021.
- c) Número de filhos nascidos vivos (independentemente da sua sobrevivência) nos últimos 12 meses antes do censo ou data de nascimento do último filho. Muitos países também distinguem, tanto aqui como na pergunta a), entre nascimentos de meninos e meninas.
- d) Alguns países, como Timor-Leste em 2015, também perguntam onde o parto ocorreu e que tipo de assistência médica a mãe recebeu.
- e) Entre os filhos nascidos vivos (nados vivos) declarados em a), quantos estavam vivos na data do censo ?
- f) Alguns países, como Angola em 2014 e Moçambique em 2007, também perguntam quantos dos filhos nascidos nos últimos 12 meses declarados em c) ainda estavam vivos na data do censo.
- g) Vários países, entre os quais o Brasil desde o Censo de 1980, perguntam se a mãe da pessoa ainda está viva. Alguns também perguntam sobre a sobrevivência do pai, mas esta informação é considerada menos confiável.
- h) Uma outra forma para medir a mortalidade é perguntando quantas pessoas faleceram no domicílio (agregado familiar) nos últimos 12 meses, qual foi a data do seu falecimento e qual era o seu sexo e sua idade no momento do falecimento.

- i) Nos PALOP que fizeram as perguntas descritas em h) acrescentou-se uma pergunta mais na rodada (ronda) dos censos de 2010, no caso de tratar-se de uma mulher em idade fértil. Ao perguntar se a mulher estava grávida, estava dando à luz ou tinha dado à luz nas últimas 6 semanas, cria-se uma oportunidade para estimar a mortalidade materna.
- j) Outra pergunta que pode ser usada para estimar a mortalidade materna, embora não seja recomendada em censos, investiga a sobrevivência das irmãs da pessoa e sua situação de gravidez (mesmas categorias de h) no momento da sua morte.

4.6.4 Dados sobre educação

Quase todos os censos também contêm algumas perguntas sobre a educação das pessoas. As mais comuns são as seguintes:

- a) Alfabetização. Esta pergunta é a mais tradicional. Ela faz parte do Censo brasileiro desde 1890 e do Censo português desde 1878. Os censos do Brasil, Cabo Verde, Guiné-Bissau e Portugal formulam a pergunta simplesmente em termos de “saber ler e escrever”, mas outros (Angola, Moçambique, São Tomé & Príncipe) dividem a informação em componentes. Em países com mais de uma língua também convém especificar “em qualquer língua”. Timor-Leste é o único país de língua portuguesa que divide a informação tanto por componente de alfabetização como por língua. O Censo de Macau de 2006 perguntou se a pessoa sabia escrever uma nota simples.
- b) Assistência escolar. No Brasil esta pergunta foi feita nos Censos de 1872, 1940 e todos os censos a partir de 1960. A formulação exata da pergunta pode variar. Geralmente pergunta-se tanto se a pessoa atualmente frequenta o ensino ou se já frequentou no passado. No Brasil pergunta-se também se a instituição de ensino é pública ou privada e se a escola está localizada no mesmo município, em outro município ou no exterior.
- c) No caso de uma resposta afirmativa a c), pergunta-se também qual foi o nível e o ano mais alto de escolaridade alcançado. Também se pergunta se o último nível cursado foi concluído.
- d) Vários países africanos também perguntam se a pessoa concluiu algum curso superior e qual é a sua área de formação. O censo brasileiro identifica a educação superior, mas não a área.
- e) Hoje em dia muitos censos perguntam se a pessoa usa regularmente o computador e a internet. Os Censos de Cabo Verde, de 2010, Portugal, de 2011, e Guiné-Bissau, de 2009, ainda não fizeram a pergunta. Nos censos do Brasil e Timor-Leste até agora se pergunta apenas se existe um computador e uma conexão à internet no domicílio (agregado familiar), mas os últimos censos de Angola, Macau e Moçambique também perguntaram sobre o uso por cada membro. O Censo de Moçambique de 2017 também perguntou sobre a frequência de uso e o meio de acesso.

4.6.5 Dados sobre ocupação e demais características econômicas

Uma seção importante do censo se dedica à atividade econômica, apesar da notória dificuldade de medir esse tipo de conceitos mediante o conjunto limitado de perguntas que normalmente podem ser feitas num censo. As perguntas principais são as que estabelecem a condição de atividade da pessoa, o número de horas trabalhadas, a condição na ocupação, o ramo de atividade e as características da empresa onde trabalha. Alguns censos, como o do Brasil, também perguntam sobre o nível de renda e as contribuições para a Previdência Social. As principais perguntas que se fazem no censo sobre este tema são as seguintes:

- a) Na semana de X a Y, durante pelo menos 1 hora
 - Trabalhou ganhando em dinheiro, produtos, mercadorias ou benefícios ?
 - Tinha algum trabalho remunerado do qual estava temporariamente afastado(a) ?
 - Ajudou sem qualquer pagamento no trabalho remunerado de membro do domicílio (agregado familiar) ?
 - Trabalhou na plantação, criação de animais ou pesca, somente para alimentação dos membros do domicílio (agregado familiar) ?
- b) Alguns censos, como o de Portugal de 2011, fazem só uma pergunta genérica sobre se a pessoa trabalhou durante a semana de referência, mas entram em mais detalhe sobre as condições inativas (dona de casa, estudante, aposentado, incapacitado) ou de desemprego que poderiam explicar porque a pessoa não trabalha.
- c) Quantos trabalhos tinha e qual era o principal ?
- d) Setor/ramo de atividade, ou seja, a atividade principal do empreendimento onde trabalhava.
- e) Posição na ocupação: empregador, empregado, autônomo, trabalhador não remunerado etc.
- f) Número de empregados no estabelecimento onde trabalha. O censo brasileiro faz esta pergunta apenas para empregadores, mas outros, como o de Portugal, a fazem a empregados também.
- g) Uma pergunta específica dos censos brasileiros desde 1991 é se a pessoa contribui para a Previdência Social. A pergunta sobre o contrato de trabalho (carteira assinada), que foi introduzida em 1991.
- h) O censo brasileiro também é um dos poucos, junto com o de Macau, que, desde 1960, pergunta sobre o nível de renda, diferenciado por ocupação principal e ocupações secundárias.
- i) Número de horas trabalhadas por semana, informação que serve, entre outras coisas, para caracterizar o subemprego.

- j) No caso de pessoas que atualmente não trabalham, pergunta-se se a pessoa tomou alguma ação para encontrar trabalho, para poder estabelecer se a pessoa está desempregada ou economicamente inativa.
- k) O Brasil, Angola e Cabo Verde também perguntam sobre a disponibilidade da pessoa para trabalhar, supondo que houvesse trabalho disponível.
- l) Fontes de renda: trabalho, aposentadoria, bolsa de estudos, aluguéis, rendimentos financeiros etc.
- m) Na rodada (ronda) dos censos de 2010, muitos países, entre os quais o Brasil, Portugal e Cabo Verde, introduziram perguntas sobre a localização do trabalho, os meios de transporte para chegar nele e o tempo de viagem tipicamente gasto no deslocamento.

Maiores detalhes sobre as perguntas específicas feitas em diferentes censos serão dados no Capítulo 14.

4.6.6 Dados sobre línguas faladas e identificação étnica-racial

As perguntas sobre línguas faladas, cor, raça e etnia estão entre as mais específicas para o contexto de cada país. No Brasil os Censos de 1940 e 1950 fizeram perguntas sobre os idiomas falados, mas depois o assunto perdeu relevância na medida em que o número de pessoas que não dominassem o português diminuiu rapidamente. A pergunta voltou no Censo de 2010, mas exclusivamente para as línguas indígenas. Os censos mais recentes de Portugal não fazem a pergunta. Cabo Verde, que não tem muita variedade linguística, também não fez a pergunta em 2010. Mas nos outros PALOP, onde existe um grande número de línguas nacionais, a pergunta tem grande importância. O Censo de Angola, em 2014, distinguiu 12 idiomas e uma categoria de surdo/mudo. O Censo de São Tomé & Príncipe, de 2012, distinguiu 8. O Censo de Moçambique, em 2017, perguntou se a pessoa entende português, em que língua aprendeu a falar e qual língua usa em casa. O Censo da Guiné-Bissau, em 2009, perguntou sobre o domínio do português e também do crioulo, espanhol, inglês, francês e russo. O Censo de Timor-Leste, de 2015, distinguiu 38 línguas maternas, 38 línguas adicionais e além disso perguntou sobre a alfabetização da pessoa em português, tetum (principal língua nacional), inglês e indonésio. Macau, por outro lado, no seu Censo de 2006, usou 9 categorias, incluindo o português e uma categoria de “outros”.

Neste contexto pode ser interessante fazer uma avaliação rápida do contingente da população dos países de língua portuguesa que efetivamente domina o idioma. O domínio do português é praticamente universal no Brasil, Portugal e São Tomé & Príncipe. Mas em Angola e Moçambique, só 71,2% e 50,4% da população, respectivamente, falam português, segundo os Censos de 2014 e 2007. No Censo de Timor-Leste de 2015, 60,7% da população declarou falar, entender ou ler o português, mas em muitos casos a compreensão se limitava à leitura. Na Guiné-Bissau, o Censo de 2009 só registrou 27,1% de domínio de português (enquanto 90,4% fala crioulo). Guiné Equatorial e Macau são países de língua portuguesa para efeitos oficiais, mas o percentual da

população que efetivamente sabe usar o idioma é mínimo. Em Macau o Censo de 2011 só captou 0,7% de usuários – menos do que a percentagem de usuários de inglês ou tagalog –, embora recentemente pareça haver certo interesse das autoridades chinesas em estimular o estudo da língua portuguesa (Antero Reto, Machado e Esperança, 2016).

Hoje em dia o Brasil faz parte de uma minoria de países lusófonos cujo censo indaga sobre questões de etnia, cor ou raça. O censo de Macau pergunta qual é o grupo étnico ao qual a pessoa pertence (7 categorias), mas não a sua cor ou raça. Cabo Verde, Guiné-Bissau, São Tomé & Príncipe e Timor-Leste perguntaram sobre etnia, cor e/ou raça apenas nos censos coloniais. Nos Censos de 1940, 1950 e 1960, usou-se o termo “tipo somático” (preto, mestiço, branco) e em 1940 e 1950 também se fez uma distinção entre a população “civilizada” e a “não civilizada”, com diferenças significativas no grau de detalhe da enumeração. No Censo colonial de Angola de 1940, por exemplo, a população “civilizada” (44.083 brancos, 23.244 mestiços, 24.231 pretos e 63 diversos) foi enumerada em termos do seu lugar de residência, sexo, idade, tipo somático, nacionalidade, naturalidade, grupo étnico, estado civil, instrução, religião, período de permanência na colônia, fecundidade e famílias segundo o número de filhos, enquanto a população “não civilizada” (4.791 mestiços e 3.641.608 pretos) foi enumerada apenas em termos do seu lugar de residência, sexo e idade. Em Moçambique alguns Censos (1928, 1935, 1945, 1955) foram realizados só para a população não africana.

O próprio Censo português em 2011 não fez a pergunta sobre etnia, cor ou raça por razões que se justificam com base na própria Constituição. Entretanto, existem planos para incluir uma pergunta (talvez de resposta opcional) sobre a origem étnico-racial no Censo de 2021 para identificar pelo menos grupos importantes como os afrodescendentes, indo-asiáticos e ciganos. Os países africanos, de um modo geral, têm muita resistência aos três conceitos, devido às conotações negativas destes conceitos na sua herança colonial e ao medo de provocar conflitos étnicos. O fato de que o Censo de Ruanda de 1991 perguntou sobre etnia – e, portanto, estabeleceu o tamanho de cada grupo étnico – às vezes é visto como um fator que contribuiu para o massacre étnico de 1994. O estabelecimento da identidade étnica, mesmo quando necessária para efeitos de representação política, também levanta enormes sensibilidades em outros contextos de conflito interétnico como nos Balcãs (Kertzer e Arel, 2002). Entretanto, os Censos moçambicanos de 1997 e 2007 perguntaram sobre raça, especificamente se o indivíduo era negro, mestiço, branco, indiano, paquistanês ou outra categoria. O Censo de 2017 também incluiu “chinês” como categoria. Guiné-Bissau em 2009 fez a pergunta de forma aberta.

No Brasil, as práticas censitárias relativas à cor ou raça têm variado. O Censo de 1872 se referiu apenas à cor, com quatro categorias: branca, parda, preta e cabocla. Em 1890, a categoria “parda” foi trocada por “mestiça” e o termo “raça” apareceu pela primeira vez. Na época as perguntas eram abertas e codificadas posteriormente. Os Censos de 1900 e 1920 não fizeram a pergunta. Em 1940, houve só três categorias, agora fechadas: “branca”, “preta” e “amarela”. A categoria “parda” foi reintroduzida em 1950 e 1960. O Censo de 1970 retirou a pergunta sobre cor ou raça por razões políticas, justificadas como a suposta irrelevância da pergunta no contexto brasileiro. O Censo de 1980 reintroduziu a pergunta, com as mesmas categorias de 1960 que excluíam os indígenas. Estes foram classificados como “pardos”. Foi só a partir de 1991 que os indígenas começaram a ser identificados com a sua própria categoria (ver Anjos, 2013; Fusco e Ojima, 2014; Petrucelli, 2012; Piza e Rosemberg, 2012), de modo que em 2010 o Censo contemplou 5 categorias:

“branca”, “preta”, “amarela”, “parda” e “indígena”. Além disso, o Censo de 2010 perguntou se a pessoa pertencia a alguma etnia indígena, a especificar. No Censo de 2021 foi introduzida uma pergunta para identificar as populações quilombolas.

4.6.7 Dados sobre deficiências

Desde a década de 80, há um número crescente de censos que incluem dados sobre deficiências ou incapacidades, atendendo a recomendações neste sentido por parte das Nações Unidas. O Brasil teve perguntas sobre este tema até o Censo de 1950, mas depois as retirou, reintroduzindo-as no Censo de 1991¹⁰. Em Portugal o tema também tem uma história longa, desde o Censo de 1878, mas foi retirado dos Censos de 1981 e 1991, sendo reintroduzido em 2001. Um problema tradicional desta pergunta é que ela pode ser feita de muitas maneiras diferentes, com resultados bastante variados, o que impossibilita a comparação internacional. Por exemplo, o Censo brasileiro de 2010 encontrou que 23,9% da população sofria de algum tipo de deficiência enquanto o Censo angolano de 2014 estimou uma prevalência de apenas 2,5%. Em 2001 a Comissão de Estatísticas das Nações Unidas autorizou o chamado Grupo de Washington, formado por delegados dos INEs e organizações que promovem investigação e pesquisa sobre incapacidades¹¹ e que tem feito recomendações sobre maneiras padronizadas de formular as perguntas. Atualmente se recomenda que o censo pergunte a cada pessoa qual é o grau de dificuldade (nenhuma, pouca, severa ou incapacidade total) que enfrenta para:

- a) Enxergar;
- b) Ouvir;
- c) Caminhar ou subir degraus;
- d) Pegar ou levantar pequenos objetos ou abrir e fechar tampas de garrafas;
- e) Lembrar-se ou concentrar-se;
- f) Lavar-se ou vestir-se (autocuidado); e
- g) Comunicar-se.

Infelizmente a aplicação das recomendações do Grupo de Washington ainda não é uniforme entre os países. Por exemplo, o Censo de Portugal de 2011 fez todas as perguntas, exceto d), mas distinguiu apenas 3 categorias de grau de dificuldade. O Censo do Brasil de 2010 fez apenas as perguntas a), b) e c), com 4 categorias de resposta, e uma pergunta binária (sim/não) sobre a deficiência intelectual. Timor-Leste em 2015 usou um formato parecido com o do Brasil, mas com 4 categorias para a pergunta de deficiência intelectual e uma pergunta adicional sobre a causa da dificuldade principal. Os PALOP, na rodada (ronda) dos censos de 2010 e no Censo moçambicano de 2017, não seguiram as recomendações do Grupo de Washington e usaram outros formatos

¹⁰ Entretanto, o formato da pergunta em 1991 foi diferente do usado nos Censos de 2000 e 2010.

¹¹ <http://www.washingtongroup-disability.com/about/>.

de perguntas ou, no caso de São Tomé & Príncipe (2012), não pesquisaram o tema. No caso de Angola e Moçambique esta decisão se justifica parcialmente pela importância de obter informação sobre os mutilados de guerra (ver Eide e Kamaleri, 2009, para o caso de Moçambique). Macau (2006) também usou um formato distinto de pergunta. Em resumo, pode-se concluir que nenhum dos países de língua portuguesa seguiu rigorosamente as recomendações do Grupo de Washington, embora o Censo de Portugal (2011) se aproxime mais do padrão proposto. Para maiores detalhes, ver Borges e Mendes (2004), Gonçalves (2003) e Lenzi (2012).

No Censo brasileiro de 2021, foi acrescentada a categoria d) e a pergunta sobre deficiência intelectual foi convertida para o mesmo formato das demais. Um acréscimo de última hora foi a pergunta sobre pessoas com autismo no domicílio, a pedido da Primeira Dama. Embora haja um precedente para esta pergunta no Censo de 2011 da Escócia, é a primeira vez que ela é feita de forma isolada e não como parte de um conjunto de condições. Existem dúvidas sobre a qualidade da informação, já que as deficiências do espectro autista têm muitas gradações e nem sempre são óbvias.

4.6.8 Dados variados

Finalmente, sempre há algumas perguntas nos censos de um ou outro país que são muito particulares e fogem aos padrões internacionais. É o caso famoso do Censo de Butão, de 2005, que perguntou “Você é feliz ? sim/não”. Alguns países na Oceania perguntam se a pessoa fuma. O Censo de Camboja, de 2008, contém várias perguntas sobre causas de morte das pessoas falecidas no domicílio (agregado familiar). Outros censos têm um interesse específico em beneficiários de determinados subsídios governamentais ou na situação dos seus militares. O Censo moçambicano de 2017 perguntou se a pessoa tem uma conta bancária e acesso a crédito financeiro formal. Entretanto, as perguntas discutidas nesta seção cobrem mais de 95% dos temas investigados nos censos do mundo inteiro.

4.6.9 Dados sobre o domicílio (agregado familiar) e a moradia

A maioria dos censos não são apenas censos de população, mas de população e habitação, razão pela qual também recolhem informação sobre a moradia. As perguntas mais frequentemente feitas são as seguintes:

- Tipo de moradia (permanente, improvisada, coletiva);
- Localização (independente, em conjunto residencial popular, em aglomerado subnormal);
- Condição de ocupação (própria inclusive o terreno, própria sem o terreno, alugada, cedida);
- Qualidade da construção (em alguns países);
- No caso de ser alugada, valor do aluguel;
- Tipo de material das paredes, do teto e do chão;

- Número de cômodos, eventualmente divididos por função (dormitório, outras funções);
- A moradia possui cozinha ?
- A moradia possui banheiro ou qual é o número de banheiros ?
- Área total da moradia (em alguns países);
- Abastecimento de água;
- Acesso à rede de energia elétrica;
- Tipo de instalação sanitária;
- Tipo de combustível usado para cozinhar e energia usada para a iluminação;
- Forma de disposição do lixo; e
- Propriedade de bens duráveis do domicílio (agregado familiar).

A informação sobre a moradia e a propriedade de bens duráveis é muito menos usada para fins demográficos do que a informação sobre pessoas, mas ela tem outras aplicações importantes. Evidentemente ela é essencial para o planejamento habitacional e de infraestrutura. Mas ela também é importante como indicador do nível de prosperidade do domicílio (agregado familiar), principalmente em países onde não se fazem perguntas diretas sobre a renda. A posse de bens duráveis em particular é considerada um indicador bastante revelador sobre o nível de vida dos moradores. Esta informação também é usada frequentemente na chamada demografia de negócios, para estudos de segmentação de mercados. Maiores detalhes serão discutidos no Capítulo 13.

4.6.10 Dados sobre atividades agrícolas

Finalmente é preciso fazer menção de um conjunto de perguntas incluídas em muitos censos que levantam informação sobre atividades agrícolas exercidas pelo domicílio (agregado familiar). Em países como o Brasil, que possuem uma tradição de Censos Agropecuários bem desenvolvida¹², esta prática é menos necessária, mas ela pode ser de muita utilidade nos países que não possuem um censo agropecuário ou que limitam o seu censo agropecuário ao setor empresarial e que, portanto, dependem do censo demográfico para providenciar informações básicas sobre a agricultura familiar (FAO/UNFPA, 2012). Em seguida, apresenta-se uma amostra das perguntas feitas nos censos dos países de língua portuguesa:

- Algum membro deste domicílio (agregado familiar) pratica (ou praticou nos últimos 12 meses) atividade agrícola por conta própria ?

¹² Cabe mencionar que Moçambique realizou um Censo Agropecuário em 2010 e Angola, no momento da publicação deste livro, está iniciando o trabalho de campo do Recenseamento Agro-Pecuário e Pescas (RAPP), a partir de 25 de março de 2021..

- Que tipo de atividade praticou (cultura de sequeiro / irrigação / silvicultura / fruticultura / pesca artesanal) ?
- Que tipo de pesca (continental / marítima artesanal / marítima semi-industrial ou industrial / aquacultura) ?
- Este domicílio (agregado familiar) tem tanques de aquacultura ?
- Condição de ocupação das terras cultivadas (cedida por autoridade tradicional / cedida por governo / cedida por familiares / arrendada / ocupada de fato / comprada / herdada)
- Área das terras cultivadas durante os últimos 12 meses
- Regime de trabalho (só familiares / mutirão comunitário / trabalho contratado)
- Uso de diferentes tecnologias como fertilizantes naturais ou industrializados, irrigação, sementes melhoradas, herbicidas etc.
- Este domicílio (agregado familiar) possui cajueiros ?
- Este domicílio (agregado familiar) possui coqueiros ?
- Quantos dos seguintes animais o domicílio (agregado familiar) tem ?
 - Vacas/Bois
 - Cabras/Cabritos
 - Ovelhas/Carneiros
 - Porcos/Leitões
 - Coelhos
 - Galinhas
 - Patos

Como as práticas da agricultura podem variar significativamente de um país para outro, as perguntas também variam.

4.7 A RODADA DE CENSOS DE 2020

Como foi assinalado na Introdução, a rodada (ronda) de censos de 2020 começou em 2015 e se estende até 2024. Guiné Equatorial fez o seu Censo em 2015 e Moçambique em 2017. Guiné-Bissau deveria ter um censo em 2019, mas o adiou e ainda não há previsão de quando será a data efetiva. Outros países tiveram que adiar os seus censos devido à epidemia da COVID-19 que interferiu

não só com a própria operação censitária como com a sua preparação. O Censo brasileiro de 2020 teve que ser adiado para 2021. Timor-Leste também adiou o seu censo para 2021, talvez até 2022. Cabo Verde adiou a data do censo de junho para mais tarde em 2020, mas possivelmente terá que adiá-lo mais. Portugal e Macau pretendiam fazer os seus censos em 2021 e ainda não se sabe se serão afetados. O próximo censo de São Tomé & Príncipe está previsto para 2022 e de Angola para 2024, ambos segundo o calendário decenal regular.

Na maioria destes países as mudanças em relação à rodada de 2010 serão relativamente pequenas. Portugal pela primeira vez seguirá o exemplo do Censo brasileiro de 2010, realizando uma operação completamente digital. Nos demais países o formato do levantamento ainda seguirá o modelo tradicional, com questionários em papel. No que diz respeito ao conteúdo, na maioria dos países haverá apenas ajustes menores. Por exemplo, Portugal pretende eliminar a pergunta sobre a residência há 5 anos, mantendo apenas a pergunta sobre a residência há 1 ano. Em vez disso se pretende incluir uma pergunta sobre a razão da migração. Outros itens que se pretende eliminar são a presença na data censitária, o nível de ensino frequentado, a utilização de um segundo meio de transporte, o número de horas trabalhadas e o número de trabalhadores da empresa, bem como diversos quesitos referentes à moradia (INE Portugal, 2018).

O questionário do Censo brasileiro de 2020 experimentou uma redução no volume de perguntas com relação ao censo anterior: o questionário básico passou de 37 para 25 perguntas e o da amostra de 112 para 76. Em termos do conteúdo, os principais cortes nas perguntas referem-se a:

- Perguntas de confirmação, para verificar se todos os membros do domicílio (agregado familiar) foram enumerados;
- Algumas perguntas sobre existência de bens no domicílio (geladeira, acesso à internet e similares);
- Algumas perguntas sobre migração, nupcialidade, educação e trabalho e renda; e
- Alguns dos dados sobre a pessoa que forneceu a informação.

Sobre estas mudanças, cabe notar que, no geral, as novas versões de questionários, modalidades e técnicas de levantamentos censitários se implementando em muitos países – planejando eliminar o questionário de papel ou entrevistas presenciais – precisam de profundas análises que avaliem as perdas e ganhos que tais modificações ocasionam e seu impacto nos sistemas de informação nacionais.

4.8 DIVULGAÇÃO E ANÁLISE SECUNDÁRIA DO CENSO

A divulgação dos resultados do censo pode ser feita de diversas maneiras. A forma mais tradicional, ou seja, a de publicar tabelas padronizadas, vem gradualmente perdendo importância, uma vez que desde o final do século passado, um número cada vez maior de institutos oficiais de estatística nacional oferecem aos usuários as próprias bases de dados (ou microdados) resguardadas as exigências do sigilo e anonimato. Nos EUA, este procedimento é praticado desde o Censo de 1960, por meio de amostras de 1% e 5%. Como se trata de amostras, a informação pode ser disseminada

sem grandes riscos de prejudicar a confidencialidade. No Brasil o mesmo foi implementado na década de 70, com a liberação pela Fundação IBGE de uma fita contendo uma amostra de 1% do Censo de 1970. Na década de 80 foi disponibilizada uma amostra de 3% e posteriormente uma de 25% do Censo de 1980. Desde então a Fundação IBGE vem construindo uma reputação como uma das instituições estatísticas nacionais que mais têm contribuído para a democratização do uso de estatísticas no mundo. O INE de Portugal também disponibiliza os microdados dos Censos de 2001 e 2011 no formato de amostras de 5%, exigindo aos pesquisadores credenciamento e assinatura de um termo de compromisso.

Há, geralmente, um *plano de tabulações* que é uma atividade que precisa conciliar os interesses de profissionais e instituições das mais diversas áreas, para que os resultados contenham o máximo de informação relevante, sem se tornarem excessivamente volumosos. Dois critérios importantes neste processo de seleção são a *viabilidade prática* e a *necessidade de sigilo*. Em função do primeiro critério, normalmente não se publica cada tabela no âmbito das unidades geográficas mais desagregadas. Em vez disso, as unidades menores são caracterizadas apenas em termos das suas características demográficas principais, enquanto as tabelas mais específicas são compiladas ao nível de unidades mais abrangentes, como o estado ou a província. Isso vale particularmente para a migração, já que uma tabela de dupla entrada para todos os fluxos ao nível geográfico mais desagregado ficaria enorme e difícil de ler, mesmo sem detalhes adicionais sobre o tempo de residência, a idade ou outras informações sobre os migrantes.

O outro motivo para a não publicação de determinadas informações censitárias é a necessidade de sigilo ou confidencialidade (também chamado *segredo estatístico*) que costuma proibir a divulgação de informações individuais¹³. Embora o nome do indivíduo conste como uma das informações básicas levantadas em praticamente todos os censos, o propósito da pergunta é apenas a identificação de cada questionário para fins de processamento. Quando as precauções oficiais para garantir a separação posterior do nome e dos dados do informante são percebidas pela população como insuficientes, a qualidade do censo pode ser gravemente prejudicada, como aconteceu, por exemplo, com o Censo de 1983 da então República Federal da Alemanha que, por este motivo, foi adiado para 1987. O problema que ocasionou a consternação pública foi o anúncio por parte do governo da sua intenção de aproveitar os dados do Censo para atualizar os registros de pessoas físicas nas prefeituras. Na sua sentença, a Suprema Corte recomendou estudos sobre a viabilidade de levantamentos futuros baseados na participação voluntária da população. À medida que evoluem os mecanismos de divulgação de dados e se expandem as possibilidades de cruzamento de informações de fontes diferentes para uma variedade de propósitos administrativos, os choques potenciais entre aplicações estatísticas legítimas e a necessidade de sigilo sobre dados individuais tendem a tornar-se mais frequentes. Mesmo quando os dados são separados da identidade da pessoa durante o processamento, é preciso tomar certas precauções para evitar a identificação indireta de indivíduos por meio de tabulações excessivamente desagregadas. Uma desagregação por nacionalidade e religião ao nível municipal, por exemplo, pode facilmente revelar informações sigilosas, pois muitos municípios contam apenas com um indivíduo de determinada nacionalidade.

¹³ O tema de abusos potenciais de dados demográficos não será tratado aqui. Para uma introdução aos problemas que podem surgir neste sentido, ver Seltzer e Anderson (2001).

Em países menores, onde as autoridades estatísticas podem não ter as mesmas facilidades para preparar amostras para usuários externos, o IPUMS (Integrated Public Use Microdata Series), da Universidade de Minnesota, nos EUA, fornece um serviço de limpeza, padronização e arquivamento de amostras para um grande número de censos do mundo inteiro. O usuário pode pedir uma amostra para fazer análises no seu próprio computador, mas também pode pedir tabulações “on-line”, usando um programa padronizado de tabulação desenvolvido pelo IPUMS. Além da preparação de amostras para usuários, o IPUMS cumpre uma função importante de depositório de cópias dos censos o que pode ser importante se os originais nos países se perdem. O acervo de censos disponíveis evidentemente evolui no tempo, mas inclui os censos do Brasil, Chile, México e a República Dominicana desde 1960, Costa Rica desde 1963, Argentina desde 1970, Indonésia desde 1971, Portugal desde 1981 e Moçambique desde 1997. Para maiores detalhes sobre este programa, os censos disponíveis no IPUMS e os tamanhos das amostras, ver <https://international.ipums.org/international/>. Entre outras plataformas de bases de dados demográficos internacionais destaca-se a North Atlantic Population Project, com censos a partir de 1850, para os EUA, Grã-Bretanha, Canadá e outros países do Atlântico Norte (<https://www.nappdata.org/napp>), e o Mosaic (<http://www.censusmosaic.org>).

A preparação de amostras para o usuário resolve grande parte das necessidades de informação, mas não todas. Por um lado, sempre haverá usuários que não podem facilmente gerar suas próprias análises e que continuam dependendo de tabulações especiais. Por outro lado, existem tipos de uso, como a análise de dados ao nível local, que exigem a base inteira em vez de só uma amostra. Os dois instrumentos mais usados para este propósito são os programas REDATAM, desenvolvido pelo Centro Latino-americano e Caribenho de Demografia (CELADE), e DevInfo, desenvolvido pelo UNICEF. Existe uma versão especial de DevInfo, chamada CensusInfo, adaptada pela Divisão de Estatísticas das Nações Unidas para as necessidades específicas da divulgação de dados censitários. Os países europeus, principalmente os nórdicos e do Leste Europeu, geralmente usam um pacote alternativo chamado PC-Axis (agora reformulado como PX-Win), que foi desenvolvido pelo Instituto Nacional de Estatística da Suécia. Para mais informação sobre estes pacotes, veja a seção 17.2 do Capítulo 17.

