



Censo Demográfico 2022

Notas metodológicas 03/2026

**Controle Estatístico de Confidencialidade dos
Microdados da Amostra do Censo Demográfico 2022**

Presidente da República

Luiz Inácio Lula da Silva

Ministro do Planejamento e Orçamento

Bruno Moretti

**INSTITUTO BRASILEIRO
DE GEOGRAFIA E
ESTATÍSTICA - IBGE**

Presidente

Marcio Pochmann

Diretora-Executiva

Flávia Vinhaes Santos

ÓRGÃOS ESPECÍFICOS SINGULARES

Diretoria de Pesquisas

Gustavo Junger da Silva

Diretoria de Geociências

Maria do Carmo Dias Bueno

Diretoria de Tecnologia da Informação

Marcos Vinícius Ferreira Mazoni

Centro de Documentação e Disseminação de Informações

José Daniel Castro da Silva

Escola Nacional de Ciências Estatísticas

Jorge Abrahão de Castro

Ministério do Planejamento e Orçamento
Instituto Brasileiro de Geografia e Estatística - IBGE

Censo Demográfico 2022

Notas metodológicas 03/2026

**Controle Estatístico de Confidencialidade dos
Microdados da Amostra do Censo Demográfico 2022**



Rio de Janeiro
2026

Instituto Brasileiro de Geografia e Estatística - IBGE

Av. Franklin Roosevelt, 166 - Centro - 20021-120 - Rio de Janeiro, RJ - Brasil

© IBGE. 2026

Sumário

Introdução	6
Avaliação e tratamento do risco de revelação	7
Modalidades de acesso aos microdados	8
Microdados de acesso público	8
Microdados de acesso controlado	9
Microdados de acesso restrito - Sala de Acesso a Dados Restritos	9
Fluxo de solicitação e autorização de acesso aos microdados de acesso controlado.....	10
Relação com os resultados oficiais.....	11
Referências.....	12

Introdução

Os Princípios Fundamentais das Estatísticas Oficiais, adotados pela *United Nations Statistics Division* (UNSD) em 1994 e endossados pela Assembleia Geral das Nações Unidas em 2014, estabelecem que os Institutos Nacionais de Estatística (INE) devem “divulgar, de forma imparcial, estatísticas de utilidade prática comprovada, para honrar o direito do cidadão à informação pública” (Princípio 1) mas também que “os dados individuais coletados pelos órgãos de estatística para elaboração de estatísticas, sejam referentes a pessoas físicas ou jurídicas, devem ser estritamente confidenciais (...)” (Princípio 6)¹. À luz desses princípios, o Controle Estatístico de Confidencialidade (CEC) é amplamente adotado pela comunidade estatística.

No Brasil, esse compromisso é reforçado pela Lei Geral de Proteção de Dados Pessoais – LGPD (Lei nº 13.709/2018)², que estabelece princípios e regras para o tratamento de dados pessoais e amplia a responsabilidade das organizações quanto à proteção dessas informações. Paralelamente, a expansão das possibilidades tecnológicas de captação, armazenamento, processamento e cruzamento de dados tem aumentado as possibilidades de identificação e reidentificação de indivíduos e domicílios, tornando necessário o aprimoramento contínuo dos mecanismos de proteção da confidencialidade.

Nesse contexto, o Controle Estatístico de Confidencialidade constitui um conjunto de métodos destinados a minimizar o risco de revelação de informações confidenciais e, simultaneamente, a reduzir a perda de informação decorrente dos procedimentos de proteção. No caso dos microdados da Amostra do Censo Demográfico 2022, a metodologia de CEC foi desenvolvida considerando as experiências anteriores de disseminação, particularmente aquelas relacionadas ao Censo Demográfico 2010, bem como os avanços metodológicos internacionais e as mudanças no ambiente tecnológico e regulatório.

A metodologia adotada compreende duas etapas principais: a avaliação do risco de revelação e a aplicação de tratamentos destinados à sua minimização. Para os microdados da Amostra do Censo Demográfico 2022, o IBGE adotou uma metodologia de disseminação estruturada em três formas de acesso, com diferentes níveis de detalhamento e diferentes mecanismos de proteção da confidencialidade:

- Microdados de acesso público;
- Microdados de acesso controlado;
- Microdados de acesso restrito, disponibilizados na Sala de Acesso a Dados Restritos.

¹ https://www.ibge.gov.br/acesso-informacao/institucional/codigos-e-principios.html?option=com_content&view=article&id=16148

² **A Lei Geral de Proteção de Dados Pessoais (LGPD)**, Lei nº 13.709/2018, foi promulgada para proteger os direitos fundamentais de liberdade e de privacidade, e a livre formação da personalidade de cada indivíduo. A Lei fala sobre o tratamento de dados pessoais, dispostos em meio físico ou digital, feito por pessoa física ou jurídica de direito público ou privado, englobando um amplo conjunto de operações que podem ocorrer em meios manuais ou digitais.

Avaliação e tratamento do risco de revelação

A avaliação do risco de revelação parte da definição de um cenário de revelação, entendido como o modelo sobre quais informações um possível intruso poderia ter à sua disposição e como essas informações poderiam ser utilizadas para obter informações adicionais a partir do conjunto de dados. Willenborg e De Waal (2001) destacam que a definição desse cenário é um elemento fundamental para a avaliação do risco de revelação.

Para os microdados da Amostra do Censo Demográfico 2022, considerou-se um possível intruso, ou “vizinho intrometido” (ver Silva, 2020), que poderia ter acesso relativamente fácil a determinadas características de pessoas e domicílios e utilizá-las para localizar ou reidentificar registros. Nesse contexto, foram consideradas variáveis semi-identificadoras, isto é, características que não constituem identificadores diretos, mas que podem ser obtidas por terceiros e combinadas para aumentar a possibilidade de identificação.

A avaliação considerou, entre outras características, sexo, idade, número de moradores no domicílio, variáveis geográficas em diferentes níveis de agregação, ocupação, educação de nível superior e peso amostral. Foram avaliados os riscos de revelação de pessoas e de domicílios, sendo o risco dos domicílios obtido a partir dos riscos associados aos indivíduos.

A utilização dos pesos amostrais é particularmente importante nessa avaliação. Uma combinação rara ou única na amostra não implica necessariamente um risco elevado, uma vez que o registro pode representar um número elevado de unidades na população. Conforme discutido por Capobianchi, Poletti e Lucarelli (2001), o plano amostral e os pesos associados aos registros devem ser considerados na avaliação do risco individual. De modo geral, quanto maior o peso amostral associado a um registro raro ou único, menor tende a ser o risco de revelação efetiva desse registro.

Após avaliação dos riscos de revelação de informações confidenciais, foram aplicados tratamentos destinados a reduzir o risco a níveis compatíveis com cada modalidade de acesso. Esses tratamentos foram definidos levando em consideração o nível geográfico, as variáveis disponibilizadas e a finalidade de cada modalidade de microdado. Entre os principais procedimentos adotados estão:

- subamostragem;
- recodificação de variáveis;
- supressão local;
- supressão global;
- calibração dos pesos amostrais;
- restrição da disponibilização de determinadas variáveis;

- mecanismos de rastreabilidade e controle de acesso.

Assim, uma nova metodologia de divulgação dos microdados da amostra do Censo Demográfico foi proposta, alterando a forma como essas informações foram divulgadas no Censo Demográfico 2010.

Modalidades de acesso aos microdados

Microdados de acesso público

Os microdados de acesso público são disponibilizados para download no portal eletrônico do IBGE, sem necessidade de identificação individual do usuário para acesso ao arquivo. Essa modalidade apresenta menor nível de detalhamento geográfico e menor quantidade de variáveis, em comparação com as modalidades de acesso controlado e restrito. O processo de controle estatístico de confidencialidade aplicado aos microdados de acesso público compreendeu as seguintes medidas:

- a Unidade da Federação constituiu o nível geográfico mais desagregado disponibilizado nessa modalidade;
- foi adotada uma subamostra de 50% dos domicílios nos setores censitários que possuíam fração amostral de 100% (aproximadamente 116 mil domicílios, correspondentes a cerca de 1,5% dos domicílios, foram suprimidos nesse processo, sendo os registros restantes reponderados);
- a variável de idade foi recodificada em grupos quinquenais, com recodificação adicional para a categoria superior a 80 anos. Essa medida reduz a granularidade da informação e, conseqüentemente, a possibilidade de identificação a partir da combinação entre idade e outras características;
- após os controles anteriores, registros que permanecessem com risco de revelação superior ao limite estabelecido foram submetidos à supressão local de determinadas características. Entre as variáveis afetadas encontram-se sexo e idade. Nos casos correspondentes, também foram suprimidas variáveis que dependiam diretamente do sexo, como as relacionadas à fecundidade;
- após a aplicação dos procedimentos anteriores, foi realizada supressão global dos domicílios que permanecessem com risco elevado (aproximadamente 50 domicílios suprimidos nessa etapa final de controle);
- após a seleção das subamostras e as supressões, os pesos amostrais foram recalibrados. A calibração das Áreas de Ponderação (APONDs) foi realizada após a aplicação dos procedimentos de redução do risco, de modo a adequar os pesos à composição final da base disponibilizada;

- variáveis semi-identificadoras com número elevado de categorias não foram disponibilizadas na modalidade de acesso público, contribuindo para a redução do risco de revelação.

Microdados de acesso controlado

Os microdados de acesso controlado são destinados a usuários que necessitam de um nível de detalhamento superior ao disponível nos microdados de acesso público, especialmente no que se refere à desagregação geográfica e à disponibilidade de variáveis. Nessa modalidade, o nível geográfico mais desagregado disponibilizado é a Área de Ponderação.

O acesso é concedido exclusivamente para fins estatísticos, técnicos ou científicos, mediante cadastramento e identificação do usuário, bem como assinatura dos instrumentos de compromisso e responsabilidade aplicáveis. Essa modalidade preserva uma quantidade maior de informações coletadas pelo questionário e disponibiliza variáveis categóricas que, em razão do elevado número de categorias, não são disponibilizadas nos microdados de acesso público.

Os principais elementos dessa modalidade são:

- Áreas de Ponderação como nível geográfico mais desagregado;
- subamostra de 50% para os domicílios pertencentes a setores censitários com fração amostral de 100%;
- calibração das Áreas de Ponderação após a subamostragem;
- disponibilização de maior quantidade de variáveis do questionário;
- disponibilização de variáveis originais que foram restringidas no microdado de acesso público.

Assim como no microdado de acesso público, os procedimentos de subamostragem e calibração devem ser considerados na interpretação dos resultados produzidos pelos usuários. Além dos procedimentos estatísticos de proteção, os microdados de acesso controlado incorporam mecanismos destinados a tornar os arquivos disponibilizados individualmente rastreáveis.

Microdados de acesso restrito - Sala de Acesso a Dados Restritos

A terceira modalidade é constituída pelos microdados disponibilizados por meio da Sala de Acesso a Dados Restritos (SAR) do IBGE. Essa modalidade é destinada a situações em que o pesquisador necessita de informações que não podem ser disponibilizadas por meio dos arquivos de acesso público ou controlado. Os microdados disponibilizados nesse ambiente possuem maior nível de detalhamento espacial e permitem a realização de agregações com recortes territoriais distintos daqueles disponíveis diretamente nos microdados de acesso controlado. Além disso, essa modalidade disponibiliza os mesmos pesos amostrais empregados nas tabulações

oficiais dos resultados definitivos da amostra do Censo Demográfico 2022, sem os ajustes aplicados aos microdados de acesso público e controlado para fins de preservação da confidencialidade.

O acesso é solicitado mediante apresentação de projeto próprio, submetido à análise de viabilidade e pertinência por comitê específico do IBGE. O comitê avalia os projetos de pesquisa e pode analisar solicitações que envolvam dados com diferentes níveis de tratamento, inclusive solicitações relacionadas a dados sem remoção de identificadores. O acesso ocorre em ambiente seguro, destinado à proteção do sigilo estatístico.

Fluxo de solicitação e autorização de acesso aos microdados de acesso controlado

O processo de solicitação e autorização de acesso aos microdados de acesso controlado envolve etapas destinadas a assegurar a identificação do usuário, a finalidade declarada do uso e a aceitação formal das responsabilidades relacionadas à utilização dos dados. O fluxo compreende:

- autenticação do usuário por meio do Gov.br;
- geração e assinatura digital do Termo de Compromisso de Confidencialidade e Responsabilidade;
- envio da declaração de vínculo institucional;
- aceitação do Termo de Uso e Finalidade no momento da solicitação;
- disponibilização dos arquivos individualizados para download.
- A finalidade declarada pelo pesquisador deve estar relacionada a estudo técnico, científico ou estatístico.

Os arquivos ficam disponíveis para download durante o período indicado no e-mail enviado ao usuário, que contém o respectivo link para acesso. Após o encerramento desse período, uma nova solicitação poderá ser realizada. O acesso controlado pressupõe o compromisso do usuário com o uso responsável e adequado dos microdados, em conformidade com as condições estabelecidas para essa modalidade de acesso. Entre as principais obrigações assumidas estão:

- utilizar os microdados exclusivamente para as finalidades declaradas na solicitação e no âmbito do projeto, pesquisa ou estudo apresentado;

- não alterar ou modificar o conteúdo dos arquivos disponibilizados, nem compartilhar, redistribuir, transferir ou ceder os arquivos a terceiros, sob qualquer forma;
- não publicar ou disponibilizar os arquivos em repositórios, plataformas de armazenamento em nuvem ou outros ambientes acessíveis a terceiros;
- não utilizar os microdados para finalidades comerciais, publicitárias ou de marketing, nem para finalidade alheia à pesquisa estatística declarada;
- abster-se de qualquer uso incompatível com o Termo de Uso, o Termo de Compromisso ou a legislação aplicável;
- reconhecer a existência de mecanismos internos de rastreabilidade utilizados pelo IBGE para fins de auditoria e, quando necessário, para subsidiar processos decorrentes de uso indevido;
- não tentar identificar, remover, modificar ou neutralizar os mecanismos de rastreabilidade;
- estar ciente de que o IBGE poderá solicitar informações relacionadas aos arquivos utilizados em trabalhos, análises, estudos ou atividades envolvendo os dados, com a finalidade de verificar o cumprimento das obrigações estabelecidas;
- estar ciente de que o uso indevido dos microdados poderá sujeitar o usuário às penalidades previstas na Lei nº 5.534/1968, na Lei nº 13.709/2018 e demais normas aplicáveis.

O Termo de Compromisso é gerado automaticamente pelo sistema a partir das informações fornecidas pelo usuário e pode ser armazenado pelo próprio solicitante.

Relação com os resultados oficiais

Os usuários devem considerar que os resultados produzidos a partir dos microdados de acesso público e de acesso controlado podem não coincidir exatamente com os resultados divulgados pelo IBGE nas tabulações do SIDRA e do BME. Para os microdados de acesso controlado, o processo de calibração desenvolvido assegura a coincidência, até o nível de área de ponderação, para a população total e por sexo. No caso dos microdados de acesso público, o processo de calibração garante, no nível de Unidade da Federação, a coincidência apenas dos resultados referentes à população total. Para os demais indicadores, podem ocorrer divergências em relação aos resultados divulgados pelo IBGE nas tabulações do SIDRA e do BME, em decorrência dos tratamentos aplicados aos microdados para a preservação da confidencialidade das informações, que incluem:

- subamostragem;
- supressão de registros;
- recodificação de variáveis;

- alterações decorrentes da calibração dos pesos;
- demais mecanismos de proteção aplicados a cada modalidade de acesso.

Assim, os microdados disponibilizados para acesso público e controlado devem ser compreendidos como produtos estatísticos derivados dos microdados da amostra, submetidos a procedimentos específicos de proteção da confidencialidade. A existência dessas diferenças não implica perda da finalidade estatística dos arquivos, mas deve ser considerada na interpretação e na comparação das estimativas.

Referências

CAPOBIANCHI, A.; POLETTINI, S.; LUCarelli, M. (2001). *Strategy for the implementation of individual risk methodology into μ -ARGUS*. Technical Report, CASC Project, Report No. 1.2-D1.

IBGE. *Código de Ética e Princípios Fundamentais das Estatísticas Oficiais*. Instituto Brasileiro de Geografia e Estatística.

SILVA, K. M. L. (2020). *Avaliação do Risco de Revelação em Microdados: Análise dos Microdados da Amostra do Censo Demográfico 2010*. Dissertação de Mestrado, ENCE, Rio de Janeiro.

WILLENBORG, L.; DE WAAL, T. (2001). *Elements of Statistical Disclosure Control*. Springer Science & Business Media.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. *Lei Geral de Proteção de Dados Pessoais – LGPD*.

BRASIL. Lei nº 5.534, de 14 de novembro de 1968. Dispõe sobre a obrigatoriedade de prestação de informações estatísticas e dá outras providências.