



Censo Demográfico 2022

Notas metodológicas 04/2026

**Processo de expansão da amostra do Censo
Demográfico 2022 para os resultados definitivos**

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 04/2026

**Processo de expansão da amostra do Censo
Demográfico 2022 para os resultados definitivos**



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
Áreas de Ponderação	8
Áreas de Ponderação para os municípios com 100 mil habitantes ou mais	8
Áreas de Ponderação para os municípios com menos de 100 mil habitantes.	9
Método para obtenção dos pesos amostrais iniciais	12
Calibração dos pesos amostrais	13
Definição das variáveis auxiliares	14
Análise da qualidade da calibração	16
Subamostra para microdados	19
Estimação de totais	19
Avaliação da precisão das estimativas	21
Referências	24
Anexo I	25
Anexo II	26
Anexo III	27

Introdução

Assim como em censos demográficos anteriores, o IBGE adotou no Censo Demográfico 2022 dois modelos de questionário - um básico e um ampliado. O questionário ampliado – da Amostra – é aplicado em domicílios selecionados para a amostra, seguindo um esquema de seleção que considera a inexistência de um cadastro prévio atualizado para a maioria do território nacional. O questionário simplificado – Básico – é aplicado nos demais domicílios.

O plano amostral adotado para a aplicação do questionário ampliado também foi semelhante ao adotado no Censo Demográfico 2010, isto é, seleção de domicílios particulares e de unidades domiciliares em domicílios coletivos, independentemente em cada Setor Censitário, o que equivale a um plano de amostragem estratificada, onde os estratos são os Setores, e a seleção de domicílios é feita com equiprobabilidade em cada estrato, sendo que a investigação é feita para todas as pessoas moradoras nos domicílios selecionados para a amostra. Uma fração amostral foi aplicada para cada Setor, baseada no tamanho populacional do município (Tabela 1). Para este Censo Demográfico foram aplicadas também, em alguns Setores, frações diferentes daquelas utilizadas no município. Frações de 10%, 20%, 33%, 50% e até 100% foram aplicadas a diferentes Setores, com o objetivo de obter resultados da amostra para distritos e subdistritos de alguns municípios, além de obter resultados mais detalhados para Terras Indígenas, Territórios Quilombolas e Favelas ou Comunidades Urbanas. No Anexo, I encontra-se a tabela com os totais de questionários aplicados com frações diferenciadas.

Tabela 1 – Fração amostral, número total de municípios, população total, número de unidades domiciliares e número de pessoas por tipo de questionário, segundo as classes de população dos municípios

Classe de população dos municípios	Fração amostral	Nº de Municípios	População total	Questionário da Amostra		Questionário Básico	
				Unidades Domiciliares	Pessoas	Unidades Domiciliares	Pessoas
Até 2500	50%	289	606.617	112.996	300.389	114.473	303.721
Mais de 2.500 até 8.000	33%	1.808	8.814.725	1.058.445	2.909.443	2.143.265	5.879.050
Mais de 8.000 até 20.000	20%	1.673	21.076.063	1.474.121	4.214.089	5.878.193	16.757.921
Mais de 20.000 até 500.000	10%	1.751	109.956.835	3.952.428	11.179.394	35.089.889	98.503.958
Mais de 500.000	5%	49	62.626.516	1.091.973	2.936.264	22.281.075	59.684.054
Total	11%	5.570	203.080.756	7.689.963	21.539.579	65.506.895	181.128.704

Nota: A classificação dos municípios por população foi realizada baseada na estimativa populacional de 2019.

Numa pesquisa realizada por amostragem probabilística, cada unidade selecionada na amostra representa, também, outras unidades que fazem parte da população-alvo. Assim, para cada unidade domiciliar selecionada na amostra do Censo Demográfico 2022, foi associado um fator de expansão ou peso obtido através do ajuste de um peso amostral inicial. Este peso também foi atribuído a cada um de seus moradores. Por unidades domiciliares pesquisadas, entende-se os domicílios particulares ocupados e as famílias e pessoas sós, moradoras em domicílio coletivo (também denominadas unidades de habitação em domicílios coletivos). Este processo de

expansão foi realizado em cada Área de Ponderação (APOND), sendo esta, a unidade geográfica formada por um agrupamento de Setores Censitários, onde os pesos iniciais foram ajustados, usando as informações conhecidas para a população como um todo. Os resultados foram obtidos usando este peso calibrado. E, como em toda pesquisa probabilística por amostragem, os erros amostrais devem ser avaliados. Para os resultados definitivos da amostra, foram utilizados modelos de regressão do tipo Função de Variância Generalizada para os cálculos de coeficientes de variação aproximados.

Áreas de Ponderação

As Áreas de Ponderação (APONDS) são definidas por agrupamento de Setores Censitários, consistindo no menor recorte geográfico para a divulgação dos resultados da amostra do Censo Demográfico 2022. São unidades territoriais obtidas por meio das visões estatísticas e geográficas, permitindo análises temáticas e espaciais mais detalhadas e a identificação de processos e padrões espaciais. É o nível geográfico onde os pesos iniciais foram ajustados, usando as informações conhecidas para a população como um todo. O tamanho dessas áreas, em termos de número de domicílios e de população, não pode ser muito reduzido, sob pena de perda de precisão de suas estimativas. Para esta construção, assim como no Censo Demográfico 2010, foi considerado o número mínimo de 400 domicílios particulares ocupados na amostra (DPOAs). As APONDS do Censo Demográfico foram definidas considerando essa condição e, também, os níveis geográficos mais detalhados da base territorial.

Inicialmente, foi gerado um conjunto de APONDS Preliminares que pudesse ser empregada para a divulgação dos primeiros dados amostrais por municípios. Para a construção das APONDS Preliminares, foram necessárias duas etapas que envolveram dois grupos distintos de municípios – municípios com mais de 100 000 habitantes e municípios com até 100 000 habitantes.

Um segundo conjunto de APONDS, denominada de Áreas de Ponderação Definitivas (APONDS Definitivas), foi criado para a divulgação final dos resultados do questionário da Amostra do Censo Demográfico 2022, permitindo também, resultados em uma escala territorial intramunicipal. À exemplo das APONDS Preliminares, a construção das APONDS Definitivas também se utilizou de dois conjuntos distintos de municípios – municípios com mais de 100 000 habitantes e municípios com até 100 000 habitantes –, mas a sua concepção envolveu também uma etapa de consulta às Superintendências Estaduais e Agências do IBGE e às Prefeituras dos municípios com mais de 100 000 habitantes.

Áreas de Ponderação para os municípios com 100 mil habitantes ou mais

Para a elaboração das APONDS definitivas, de forma a tornar esse recorte territorial o mais aderente possível às necessidades dos municípios brasileiros com mais de 100 000 habitantes, criou-se um sistema de consulta às Agências e Superintendências Estaduais (SES) do IBGE e às prefeituras desses municípios. Com isso, estabeleceu-se um trabalho cooperativo entre o IBGE e as Prefeituras que optaram por participar do processo de delimitação das APONDS.

Para a construção das APONDS a serem propostas pelo IBGE às prefeituras dos municípios com população superior a 100 mil habitantes foram utilizadas como referências

iniciais duas classificações de APONDS e a base de setores censitários definitiva de 2022. Para os municípios que fizeram parte da publicação “Tipologia intraurbana: espaços de diferenciação socioeconômica do Brasil” (IBGE, 2017), foram utilizadas como referência as APONDS 2010 desse estudo, uma vez que os critérios para as suas definições consideravam uma maior homogeneidade socioeconômica interna. Já para os demais municípios acima de 100 mil habitantes, foram utilizadas as APONDS de 2010 empregadas na divulgação dos dados daquele Censo Demográfico.

Uma vez definidas as APONDS de 2010 que serviriam como referência para a criação das APONDS Definitivas de 2022, foi realizada uma compatibilização de limites territoriais. Para tal, os setores censitários definitivos de 2022 foram interpostos às APONDS de 2010, de modo que os setores que estavam integralmente inseridos em uma APOND de 2010 receberam automaticamente o código dessa APOND. Para os setores censitários que estavam entre duas ou mais APONDS de 2010, foi calculada a área de interseção de cada parte e o setor censitário definitivo de 2022 foi integralmente associado à APOND 2010 cuja área de sobreposição era maior.

A partir da base de setores censitários definitivos de 2022 com os códigos das APONDS de 2010 associados, foi realizada uma avaliação manual (a partir da espacialização em aplicativos de geoprocessamento) pelos técnicos do IBGE, corrigindo quantitativos mínimos e buscando uma homogeneidade espacial.

As APONDS resultantes desse primeiro trabalho foram alimentadas no sistema criado pelo IBGE para consulta das Superintendências Estaduais do IBGE (SES) e Prefeituras. Tal sistema foi desenvolvido para ser utilizado em navegadores web (como Google Chrome e Microsoft Edge) e possuía todas as ferramentas para que as prefeituras e SES visualizassem em forma de mapa as APONDS propostas, trocassem os setores censitários de APONDS, criassem novas APONDS e removessem as que não interessavam, além de gerar automaticamente o relatório de conclusão do trabalho e dar alertas de falta de contiguidade espacial e de quantitativos menores que 400 e maiores que 900 DPOAs.

Às prefeituras foi facultada a possibilidade de alteração na composição das APONDS de forma a compatibilizá-las às suas necessidades de informação estatística, guardadas as restrições de números mínimo e máximo de DPOAs e de contiguidade espacial. Não houve qualquer impedimento, salvo os mencionados anteriormente, para a adequação das APONDS aos anseios das Prefeituras.

Áreas de Ponderação para os municípios com menos de 100 mil habitantes.

Para os municípios com população menor ou igual a 100 mil habitantes, não houve consulta pública nos moldes do que foi realizado para os municípios com mais de 100 mil

habitantes. Todavia, para garantir uma maior qualidade na divulgação estatística para esse grupo de municípios, foi facultado às SES a decisão de atuar na elaboração das APONDS para esse conjunto de municípios, na medida de suas possibilidades e desde que informassem à equipe responsável pelo trabalho na Sede. Não foi facultada a participação das Prefeituras nesse trabalho.

Nos casos em que houve manifestação das SES, a equipe da Sede do IBGE disponibilizou a malha de setores censitários definitivos de 2022, com o quantitativo de DPOAs e o conjunto de variáveis usadas para a confecção das APONDS propostas às prefeituras, exceto a variável de renda. Assim, as SES que optaram por realizar esse trabalho puderam criar, a partir do zero, ou seja, sem a existência prévia de um conjunto de APONDS elaborada pela equipe da Sede, as APONDS para esses municípios. Para tal, utilizaram softwares comuns de geoprocessamento.

Para as SES que não manifestaram a disponibilidade de trabalhar nas APONDS dos municípios com população igual ou inferior a 100 000 habitantes, as APONDS Definitivas foram geradas de forma automática. Para tal foi desenvolvido um algoritmo – programado em linguagem R e utilizando recursos de paralelismo – baseado em conceitos de otimização e teoria dos grafos. Tal algoritmo foi aplicado aos municípios, sendo considerados três possíveis níveis de agregação, a saber: 1º nível subdistrito; 2º nível distrito; e 3º nível geral (o próprio município). Inicialmente, o algoritmo é aplicado em cada um dos subdistritos que compõem um município. Havendo solução viável nesse nível, ou seja, APONDS que atendem às restrições supracitadas, utiliza-se tal solução. Caso contrário, aplica-se o algoritmo por distrito e, por fim, no 3º nível – geral. Além disso, quando possível, tentou-se manter as estruturas de Terras Indígenas, Territórios Quilombolas e Favelas e Comunidades Urbanas.

A definição das APONDS preconiza o agrupamento de Setores que satisfaçam, simultaneamente, duas restrições: (1) Conexidade: Setores alocados a um mesmo grupo devem ser contíguos entre si; (2) Capacidade: a soma do quantitativo de domicílios associados a cada um dos Setores – dentro de cada um dos grupos – deve ser maior ou igual a 400 Domicílios Particulares Ocupados na Amostra. Os grupos de Setores que satisfazem essas duas restrições correspondem às APONDS.

Além disso, de forma a nortear a construção de soluções de boa qualidade, traduzidas em APONDS mais homogêneas sob o ponto de vista socioeconômico, utiliza-se uma função de avaliação (função objetivo) que agrega a distância pelo método de Mahalanobis relacionada a um conjunto de oito variáveis quantitativas, cujos valores estão disponíveis para cada um dos Setores Censitários que serão agrupados: V1 - Proporção de domicílios particulares permanentes ligados à rede geral de água; V2 - Proporção de domicílios particulares permanentes ligados à rede geral de esgoto ou pluvial; V3 - Número médio de pessoas por domicílio particular permanente; V4 - Número médio de moradores por banheiro em domicílios particulares permanentes onde exista ao menos 1 banheiro; V5 - Rendimento médio total das pessoas responsáveis por domicílios particulares; V6 - Proporção de pessoas com 65 anos ou mais de idade em domicílios particulares; V7 - Razão de dependência dos domicílios particulares;

V8 - Proporção de pessoas com 15 anos ou mais de idade e que sabem ler e escrever. Basicamente, quanto menor o valor dessa função – soma das distâncias entre os Setores que compõem cada uma das APONDS construídas – melhor é a sua qualidade.

O algoritmo é aplicado em duas etapas. Na 1ª etapa, com o objetivo de contemplar de forma imediata a restrição de contiguidade, são construídas 200 árvores geradoras (AGs) distintas. Cada AG abarca possíveis ligações (relações de vizinhança) entre os Setores que irão compor as APONDS – ideia de agrupamento hierárquico – sendo cada AG, em um primeiro estágio, correspondente a uma única APOND.

Por sua vez, a construção de um conjunto de k APONDS é realizada aplicando-se, em cada uma das AGs, um procedimento de particionamento executado em k-1 iterações, onde, em cada iteração, são produzidas duas novas subárvores relacionadas com duas APONDS que cumpram a restrição de capacidade. Destaca-se que em todas as k-1 iterações, independente do particionamento realizado, levando em conta a característica de uma árvore geradora, são produzidas sempre duas novas subárvores conexas, que correspondem a grupos de Setores contíguos.

Por fim, na 2ª etapa, considerando a melhor solução produzida na 1ª etapa, ou seja, a AG particionada com menor valor de função objetivo, aplica-se um procedimento denominado “busca local” que, em linhas gerais, efetua diversas trocas entre Setores de APONDS distintas – sempre preservando as restrições de contiguidade e capacidade – de forma a aumentar a homogeneidade dentro de cada uma das APONDS. A solução produzida após a aplicação desse procedimento corresponde às APONDS definidas para o município em questão.

No total, foram criadas 14.270 APONDS considerando os dois métodos descritos anteriormente.

Para maiores detalhes sobre a metodologia de construção das APONDS, consultar o documento “Notas metodológicas 02/2026 - Metodologia de construção das Áreas de Ponderação”.

Método para obtenção dos pesos amostrais iniciais

Para o cálculo dos pesos amostrais iniciais, foram definidos estratos de ponderação dentro das APONDS. Os estratos foram criados em função das frações amostrais teóricas dos Setores Censitários que os compõem e os tipos desses Setores: Terras Indígenas, Territórios Quilombolas, Favelas e Comunidades Urbanas e demais Setores. Desta maneira, Setores com a mesma fração amostral teórica e mesmo tipo foram alocados no mesmo estrato de ponderação.

Em cada estrato de ponderação h da APOND a , o peso amostral inicial da unidade domiciliar i foi dado por:

$$w_{ahi} = \frac{N_{ah}}{n_{ah}}, \quad (1)$$

onde:

w_{ahi} é o peso inicial da unidade domiciliar i no estrato de ponderação h da APOND a ;

N_{ah} é o total de unidades domiciliares no conjunto universo no estrato de ponderação h da APOND a ;

n_{ah} é o total de unidades domiciliares na amostra no estrato de ponderação h da APOND a .

Nos estratos de ponderação compostos por Setores contendo frações amostrais teóricas iguais a 1, a não resposta foi tratada por imputação. Deste modo, foi atribuído a todas as unidades domiciliares, destes estratos, pesos iniciais iguais a 1.

Calibração dos pesos amostrais

Para assegurar a coerência entre as estimativas da amostra e os totais populacionais do universo, para um conjunto de variáveis conhecidas, foi realizado um processo de calibração para ajustar os pesos amostrais iniciais do Censo Demográfico. Os ajustes foram realizados de maneira independente em cada APOND, para os estratos de ponderação com frações teóricas diferentes de 1. Além de fornecer estimativas mais consistentes para as variáveis pesquisadas, o processo de calibração contribui para a melhoria da precisão de algumas estimativas.

Como apresentado em Silva et al (2021), a calibração consiste na obtenção de novos pesos (pesos calibrados), com menor distância para os pesos iniciais, tal que os totais estimados das variáveis auxiliares com os novos pesos w_{ai}^* satisfaçam as restrições de calibração definidas por:

$$\hat{T}_{ax}^* = \sum_{i \in s} w_{ai}^* x_{ai} = T_{ax}, \quad (2)$$

onde:

x_{ai} é o valor da variável auxiliar x da unidade domiciliar i da APOND a ;

T_{ax} são os totais populacionais das variáveis auxiliares da APOND a ;

$\hat{T}_{ax}^*$ são os totais estimados após a calibração dos pesos da APOND a .

Para evitar a obtenção de pesos considerados extremos ou não representativos (como $w_{ai}^* < 1$), foram impostos limites definidos por:

$$w_{ai}/5 \leq w_{ai}^* \leq 5 \times w_{ai}, \quad (3)$$

onde $w_{ai}/5 \geq 1$.

O cálculo dos pesos calibrados foi baseado no método dos Mínimos Quadrados Generalizados – MQG. A ideia é solucionar o problema de encontrar novos pesos (w^*) que tenham uma distância mínima D para os pesos amostrais iniciais (w):

$$D(\mathbf{w}^*, \mathbf{w}) = \sum_{i \in s} G_{ai}(w_{ai}^*, w_{ai}), \quad (4)$$

onde G é a função de distância. Se utilizarmos os estimadores de regressão como estimadores de calibração, estes seriam definidos como:

$$G_{ai}(w_{ai}^*, w_{ai}) = \frac{(w_{ai}^* - w_{ai})^2}{w_{ai}}, \quad \forall i \in s. \quad (5)$$

A solução deste problema é dada por:

$$w_{ai}^* = w_{ai} \times g_{ai}, \quad (6)$$

onde:

$$g_{ai} = 1 + (\mathbf{T}_x - \hat{\mathbf{T}}_x^*)' \left(\sum_{i \in s} w_{ai} x_{ai} x_{ai}' \right)^{-1} x_{ai}. \quad (7)$$

Maiores detalhes da metodologia também podem ser obtidos em Särndal et al. (1992) e Deville et al. (1993).

O produto da aplicação do processo de calibração é um peso ajustado para cada unidade domiciliar da amostra, que é repetido nos registros de cada pessoa moradora na unidade domiciliar.

Este processo foi desenvolvido em um programa em linguagem R por técnicos do IBGE, usando a biblioteca *survey*.

Definição das variáveis auxiliares

As variáveis auxiliares constituem um subconjunto das variáveis comuns ao Questionário da Amostra e ao Questionário Básico referentes aos domicílios e às pessoas da amostra. A escolha das variáveis auxiliares utilizadas como restrições no processo de ajuste do qual derivam os pesos é um aspecto importante do método aplicado. A forma ou prioridade de tratamento dessas variáveis, sobretudo, quando não existe uma solução que atenda simultaneamente a todas as restrições, é outro ponto sensível do método.

A metodologia de ajuste deve satisfazer algumas condições. Uma delas é que as variáveis não sejam linearmente dependentes ou quase linearmente dependentes.

Outra condição imposta para a aplicação dessa metodologia, que pode influenciar na sua qualidade, é o tamanho da restrição, entendido como o número de domicílios aos quais a restrição se aplica em uma dada APOND. Quando uma restrição não atinge um número mínimo de unidades domiciliares na população de uma APOND, neste caso fixado em 50 unidades domiciliares, essa restrição foi considerada rara, sendo descartada da lista de variáveis auxiliares.

Dessa forma, o programa de ajuste do modelo incorpora procedimentos de eliminação de restrições que se enquadrem nas condições acima, observando a ordem que segue: restrições raras, restrições redundantes ou quase redundantes.

Convém ressaltar que a eliminação de restrições pode implicar diretamente no fato de não se ter a garantia da calibração desejada para as variáveis eliminadas na respectiva APOND.

As restrições inicialmente definidas para cada uma das áreas de ponderação, encontram-se na relação a seguir. Elas constituem o conjunto denominado conjunto 1 de restrições para calibração. Para sua definição realizou-se reuniões internas com técnicos da área de Projeções e Indicadores Sociais da Diretoria de Pesquisas do IBGE, partindo do conjunto inicial adotado no Censo Demográfico 2010.

Restrições em unidades domiciliares:

1. Número total de pessoas
2. Número total de unidades domiciliares
3. Número de pessoas do sexo masculino
4. Número de pessoas na faixa de idade de 0 a 4 anos
5. Número de pessoas na faixa de idade de 5 a 9 anos
6. Número de pessoas na faixa de idade de 10 a 14 anos
7. Número de pessoas na faixa de idade de 15 e 19 anos
8. Número de pessoas na faixa de idade de 20 a 24 anos
9. Número de pessoas na faixa de idade de 25 a 29 anos
10. Número de pessoas na faixa de idade de 30 a 34 anos
11. Número de pessoas na faixa de idade de 35 a 39 anos
12. Número de pessoas na faixa de idade de 40 a 44 anos
13. Número de pessoas na faixa de idade de 45 a 49 anos
14. Número de pessoas na faixa de idade de 50 a 54 anos
15. Número de pessoas na faixa de idade de 55 a 59 anos
16. Número de pessoas na faixa de idade de 60 a 64 anos
17. Número de pessoas na faixa de idade de 65 a 69 anos
18. Número de pessoas na faixa de idade de 70 a 74 anos
19. Número de pessoas na faixa de idade de 75 a 79 anos
20. Número de pessoas na faixa de idade de 80 anos ou mais

21. Número de pessoas do sexo masculino na faixa de idade de 0 a 4 anos
22. Número de pessoas do sexo masculino na faixa de idade de 5 a 9 anos
23. Número de pessoas do sexo masculino na faixa de idade de 10 a 14 anos
24. Número de pessoas do sexo masculino na faixa de idade de 15 a 19 anos
25. Número de pessoas do sexo masculino na faixa de idade de 20 a 24 anos
26. Número de pessoas do sexo masculino na faixa de idade de 25 a 29 anos
27. Número de pessoas do sexo masculino na faixa de idade de 30 a 34 anos
28. Número de pessoas do sexo masculino na faixa de idade de 35 a 39 anos
29. Número de pessoas do sexo masculino na faixa de idade de 40 a 44 anos
30. Número de pessoas do sexo masculino na faixa de idade de 45 a 49 anos
31. Número de pessoas do sexo masculino na faixa de idade de 50 a 54 anos
32. Número de pessoas do sexo masculino na faixa de idade de 55 a 59 anos
33. Número de pessoas do sexo masculino na faixa de idade de 60 a 64 anos
34. Número de pessoas do sexo masculino na faixa de idade de 65 a 69 anos
35. Número de pessoas do sexo masculino na faixa de idade de 70 a 74 anos
36. Número de pessoas do sexo masculino na faixa de idade de 75 a 79 anos
37. Número de pessoas do sexo masculino na faixa de idade de 80 anos ou mais
38. Número de pessoas moradoras na situação urbana
39. Número de pessoas em Terras Indígenas
40. Número de pessoas em Territórios Quilombolas
41. Número de pessoas em Favelas ou Comunidades Urbanas

Restrições em domicílios particulares ocupados:

42. Número total de pessoas
43. Número total de unidades domiciliares
44. Número de domicílios na situação urbana
45. Número de domicílios particulares permanentes sem banheiros
46. Número de domicílios particulares permanentes com 1 banheiro
47. Número de domicílios particulares permanentes com 2 ou mais banheiros

Análise da qualidade da calibração

As restrições, apresentadas na lista anterior, foram agrupadas em 6 conjuntos alternativos que foram utilizados em ordem de prioridade. Esse procedimento foi adotado para garantir que alguma calibração fosse feita mesmo utilizando um conjunto menor de restrições. O primeiro conjunto foi formado por todas as restrições, como listadas, e os demais formados pela

agregação de faixas etárias, agregação de faixas de número de banheiros por domicílio ou mesmo a retirada de grupos de restrições.

No cálculo dos pesos calibrados, para cada APOND, foi utilizado inicialmente o conjunto de restrições número 1. Quando não se obteve uma solução, a área foi processada novamente, utilizando o conjunto 2 e assim sucessivamente até o conjunto 6. A composição de cada um dos outros cinco conjuntos alternativos de restrições está apresentada no Anexo II.

A Tabela 2 a seguir, mostra o número de áreas de ponderação por conjunto de restrições utilizado no ajuste de calibração.

Tabela 2 - Número de áreas de ponderação, por conjunto de restrições utilizado na calibração segundo o tipo de área de ponderação

Tipo de Área de Ponderação	Conjunto de restrições utilizado na calibração						Total
	1	2	3	4	5	6	
Município inteiro	2.858	3	1		43	53	2.958
Consulta à prefeituras	4.330	60	22		702	181	5.295
SES IBGE	1.491	8	6		59	53	1.617
Equipes da Sede do IBGE	91	2	2	1	4	8	108
Automáticas	3.859	38	13		224	154	4.288
Total	12.629	111	44	1	1.032	449	14.266

Fonte: IBGE, Diretoria de Pesquisas 2026.

No final do processo, foi garantido que pelo menos as restrições referentes ao total de unidades domiciliares, total de pessoas e total de pessoas por sexo fossem respeitadas para todas as áreas de ponderação.

Como mostra a Tabela 2, cerca de 90% das APONDS foram calibradas com o primeiro conjunto de restrições, que considera a lista com 47 restrições, exceto aquelas que possam ter sido excluídas por raridade ou por serem combinações lineares de outras presentes nesse mesmo conjunto, para uma APOND específica. Do total de 14.270 APONDS, 4 delas foram formadas totalmente por domicílios com aplicação do questionário de amostra e por tanto não precisaram de calibração.

Os domínios geográficos formados por agregação de APONDS onde pelo menos uma destas não tenha sido calibrada com todas as restrições, não apresentará estimativas coincidentes com os parâmetros populacionais para essas restrições.

Na Tabela 3 é apresentada a comparação entre os totais populacionais das restrições do conjunto inicial e suas estimativas obtidas a partir da amostra utilizando os pesos calibrados, para o Brasil.

Tabela 3 – Totais populacionais e estimados das restrições de calibração para Brasil

Indicador	Total Populacional	Estimativa pontual	Variação (%)
Total de pessoas	203.080.756	203.080.756	0,0000
Total de unidades domiciliares	73.313.618	73.313.618	0,0000
Total de homens	98.532.431	98.532.431	0,0000
Pessoas de 0 a 4 anos	12.704.860	12.720.625	0,1241
Pessoas de 5 a 9 anos	13.749.440	13.752.153	0,0197
Pessoas de 10 a 14 anos	13.674.961	13.665.051	-0,0725
Pessoas de 15 a 19 anos	14.375.942	14.364.148	-0,0820
Pessoas de 20 a 24 anos	15.466.463	15.443.172	-0,1506
Pessoas de 25 a 29 anos	15.469.723	15.450.407	-0,1249
Pessoas de 30 a 34 anos	15.473.117	15.459.318	-0,0892
Pessoas de 35 a 39 anos	16.172.791	16.148.812	-0,1483
Pessoas de 40 a 44 anos	16.072.170	16.065.839	-0,0394
Pessoas de 45 a 49 anos	13.640.112	13.626.931	-0,0966
Pessoas de 50 a 54 anos	12.598.581	12.598.794	0,0017
Pessoas de 55 a 59 anos	11.569.106	11.572.536	0,0296
Pessoas de 60 a 64 anos	9.944.389	9.960.980	0,1668
Pessoas de 65 a 69 anos	7.876.232	7.904.840	0,3632
Pessoas de 70 a 74 anos	5.858.536	5.881.889	0,3986
Pessoas de 75 a 79 anos	3.847.379	3.861.008	0,3542
Pessoas de 80 ou mais anos	4.586.954	4.604.254	0,3772
Homens de 0 a 4 anos	6.461.689	6.471.567	0,1529
Homens de 5 a 9 anos	7.011.282	7.011.247	-0,0005
Homens de 10 a 14 anos	6.992.746	6.985.562	-0,1027
Homens de 15 a 19 anos	7.317.515	7.310.456	-0,0965
Homens de 20 a 24 anos	7.767.306	7.759.609	-0,0991
Homens de 25 a 29 anos	7.627.458	7.618.047	-0,1234
Homens de 30 a 34 anos	7.537.285	7.530.418	-0,0911
Homens de 35 a 39 anos	7.827.333	7.816.705	-0,1358
Homens de 40 a 44 anos	7.781.059	7.779.495	-0,0201
Homens de 45 a 49 anos	6.549.109	6.545.589	-0,0538
Homens de 50 a 54 anos	6.014.391	6.015.868	0,0246
Homens de 55 a 59 anos	5.419.505	5.418.487	-0,0188
Homens de 60 a 64 anos	4.605.834	4.613.444	0,1652
Homens de 65 a 69 anos	3.588.052	3.599.286	0,3131
Homens de 70 a 74 anos	2.615.350	2.626.915	0,4422
Homens de 75 a 79 anos	1.657.786	1.662.427	0,2799
Homens de 80 ou mais anos	1.758.731	1.767.310	0,4878
Pessoas na situação urbana	177.508.417	177.514.945	0,0037
Pessoas em Domicílios particulares	202.243.505	202.262.452	0,0094
Total de domicílios particulares	72.522.372	72.532.083	0,0134
Total de domicílios particulares urbanos	64.057.890	64.064.953	0,0110
Total de domicílios particulares permanentes sem banheiro	1.414.140	1.445.788	2,2380
Total de domicílios particulares permanentes com um banheiro	47.163.906	47.313.283	0,3167
Total de domicílios particulares permanentes com mais de um banheiro	23.944.326	23.773.012	-0,7155
Total de pessoas em Terras Indígenas	689.532	690.121	0,0854
Total de pessoas em Territórios Quilomboas	203.574	203.629	0,0268
Total de pessoas em Favelas ou Comunidades Urbanas	16.393.840	16.395.550	0,0104

Fonte: IBGE, Diretoria de Pesquisas 2026.

Subamostra para microdados

Como mencionado anteriormente, o Censo Demográfico 2022 contou com frações diferenciadas para alguns territórios, chegando a aplicar o questionário da amostra em todos os domicílios de determinadas áreas.

Para a divulgação dos microdados de acesso controlado do questionário da amostra, foi necessária a seleção de uma subamostra de domicílios nessas áreas. A seleção foi realizada em cada setor que possuía domicílios com essa característica, estratificando por tipo de unidade domiciliar e tipo de área: Favela ou Comunidades Urbanas, Terras Indígenas ou Territórios Quilombolas. A fração para essa seleção foi de 50% que é a maior fração utilizada no Censo 2022.

Depois de selecionada a subamostra, uma nova calibração foi realizada, seguindo a mesma metodologia apresentada anteriormente.

Para a divulgação dos microdados públicos, além da subamostra, foi feita uma supressão global de 49 domicílios devido ao controle estatístico de confidencialidade¹ aplicado. Neste caso, também foi realizada a calibração por APOND, seguindo a mesma metodologia dos dados completos.

Na prática, isso resulta em termos pesos diferenciados para cada um desses conjuntos de microdados. E como o conjunto de restrições da calibração usada em cada APOND pode ser sido alterado, nem todas as estimativas para os indicadores calculados com os microdados disponibilizados coincidirão com as divulgadas oficialmente pelo IBGE.

No Anexo III é possível encontrar as tabelas com os totais de registros de amostra para cada uma dos microdados divulgados.

Estimação de totais

As estimativas de totais para domínios de interesse, devem ser feitas utilizando-se, para cada unidade (pessoa ou domicílio particulares permanentes), o peso correspondente, que foi determinado para cada unidade domiciliar da amostra e atribuído, também, a cada pessoa dessa unidade. Assim, para estimar o total de uma característica qualquer investigada pelo Questionário da Amostra do Censo Demográfico 2022, utiliza-se o estimador $\hat{Y}$, definido por:

$$\hat{Y} = \sum_{i=1}^n w_i^* y_i, \quad (8)$$

¹ Para maiores detalhes sobre o Controle Estatístico de Confidencialidade aplicado aos microdados do Censo Demográfico 2022, consultar a o documento “Notas Metodológicas 03/2026 - Controle Estatístico de Confidencialidade dos Microdados da Amostra do Censo Demográfico 2022”.

onde:

w_i^* é o peso calibrado da unidade i da amostra no domínio;

y_i é o valor de y da unidade i no domínio;

n é o número de unidades na amostra no domínio.

Os pesos calculados com a metodologia adotada não devem ser substituídos por pesos inteiros para não provocar a quebra na consistência das restrições efetivamente utilizadas no ajuste no modelo. O uso de pesos fracionários preserva o método de expansão da amostra e produz resultados mais precisos do ponto de vista estatístico. Assim, para o cálculo das estimativas das tabelas de divulgação do Censo Demográfico 2022 foi utilizado o peso fracionário com 13 casas decimais, sendo, então, arredondadas as estimativas resultantes.

Avaliação da precisão das estimativas

As estimativas de uma pesquisa por amostragem probabilística possuem um erro amostral associado que deve ser considerado para as análises a serem realizadas. Assim, a avaliação dos erros amostrais é um ponto fundamental, pois dele decorre o grau de confiança nas conclusões analíticas.

Os erros amostrais podem ser avaliados através das estimativas dos coeficientes de variação (cv) ou dos erros-padrão calculados a partir das estimativas das variâncias².

O método para estimar a variância das estimativas de totais considerando o processo de calibração encontra-se descrito em Särndal et al. (1992).

Os pesos calibrados foram utilizados para a estimação dos coeficientes de variação aproximados, utilizando o software SAS, para indicadores do plano tabular referentes a domicílios e pessoas, considerando três níveis geográficos: Brasil, Grandes Regiões e Unidades da Federação.

Esses resultados foram utilizados para ajustar um modelo linear, relacionando as estimativas de variância aos valores esperados dos totais estimados, seguindo o método de Função Generalizada de Variância – FGV descrito em Wolter (2007), para a obtenção dos coeficientes de variação dos resultados da amostra do Censo Demográfico 2022.

Desta maneira temos que:

$$\log(\widehat{cv}_i) = \alpha + \beta \log(\hat{t}_i), \quad (9)$$

onde $\hat{t}_i$ e $\widehat{cv}_i$ são, respectivamente, o total e o coeficiente de variação estimados do indicador i.

Um ajuste usual de mínimos quadrados é realizado para a obtenção das estimativas dos parâmetros do modelo e uma estimativa aproximada de cv^* pode ser predita avaliando o modelo FGV para um determinado total estimado de uma variável y, da seguinte forma:

$$\widehat{cv}^*(\hat{t}_y) = A \times \hat{t}_y^B \quad (10)$$

A Tabela 4 apresenta as estimativas dos coeficientes de regressão A e B, ajustados a partir de (9).

² Para maiores informações sugere-se a leitura do "Guia para Divulgação de Erros Amostrais nas Pesquisas por Amostragem Probabilística Realizadas pelo IBGE", Rio de Janeiro, IBGE, 2021.
<https://biblioteca.ibge.gov.br/visualizacao/livros/liv101898.pdf>

Tabela 4 – Estimativas dos coeficientes do modelo de Função Generalizada de Variância ajustado para os coeficientes de variação de totais de indicadores do plano tabular – Brasil – Censo Demográfico 2022

Tipo de estimativa	Coeficientes de regressão	
	A	B
Pessoas	4,0616	-0,5004
Domicílios	3,6106	-0,4959

Fonte: IBGE, Diretoria de Pesquisas, Coordenação de Métodos e Qualidade.

A Tabela 5 apresenta as estimativas de coeficientes de variação para Brasil, considerando ordens de grandeza de totais estimados³. Como exemplo, temos que uma estimativa de total de domicílios da ordem de grandeza 10.000 domicílios terá um coeficiente de variação aproximado de 3,8%.

³ As tabelas com as estimativas de coeficientes de variação e estimativas dos parâmetros A e B para os demais níveis geográficos podem ser encontradas no arquivo “Coeficientes de variação e de regressão”, disponibilizado juntamente com a documentação dos microdados do Censo Demográfico 2022.

Tabela 5 – Estimativas dos coeficientes de variação para alguns totais baseados na Função Generalizada de Variância – Brasil – Censo Demográfico 2022

Tamanho da estimativa	Coeficientes de variação (%)	
	Pessoas	Domicílios
100	40,5	36,8
500	18,1	16,6
1 000	12,8	11,7
2 000	9,1	8,3
5 000	5,7	5,3
10 000	4,0	3,8
20 000	2,9	2,7
50 000	1,8	1,7
100 000	1,3	1,2
150 000	1,0	1,0
200 000	0,9	0,8
500 000	0,6	0,5
1 000 000	0,4	0,4
2 000 000	0,3	0,3
3 000 000	0,2	0,2
4 000 000	0,2	0,2
5 000 000	0,2	0,2
6 000 000	0,2	0,2
7 000 000	0,2	0,1
8 000 000	0,1	0,1
9 000 000	0,1	0,1
10 000 000	0,1	0,1
15 000 000	0,1	0,1
20 000 000	0,1	0,1
25 000 000	0,1	0,1
30 000 000	0,1	0,1
35 000 000	0,1	0,1
40 000 000	0,1	0,1
45 000 000	0,1	0,1
50 000 000	0,1	0,1
60 000 000	0,1	0,1
70 000 000	0,0	0,0
80 000 000	0,0	0,0
90 000 000	0,0	0,0
100 000 000	0,0	0,0
120 000 000	0,0	0,0
130 000 000	0,0	0,0
140 000 000	0,0	0,0
150 000 000	0,0	0,0
160 000 000	0,0	0,0
170 000 000	0,0	0,0
180 000 000	0,0	0,0
190 000 000	0,0	0,0
200 000 000	0,0	0,0

Referências

Deville J-C, Sarndal C-E, Sautory O (1993) Generalized Raking Procedures in Survey Sampling. JASA 88:1013-1020.

IBGE. Tipologia intraurbana: espaços de diferenciação socioeconômica nas concentrações urbanas do Brasil / IBGE, Coordenação de Geografia. - Rio de Janeiro: IBGE, 2017. 164p. ISBN 978-85-240-4429-8 1. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2101470>. Acesso: nov. /2024.

IBGE. *Guia para divulgação de erros amostrais nas pesquisas por amostragem probabilística realizadas pelo IBGE*. Rio de Janeiro, 2021. 36 p. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101898.pdf>. Acesso em: nov. 2024.

Särndal, C.-E., Swensson, B., & Wretman, J. (1992). *Model assisted survey sampling*. Springer-Verlag Publishing.

Silva, P. L. N., Bianchini, Z. M. e Dias, A. J. R. (2021). Amostragem: teoria e prática usando R. Disponível em: <https://amostragemcomr.github.io/livro/index.html>. Acesso em: nov. 2024.

Wolter, K. M. (2007). *Introduction to variance estimation* (2nd ed.). New York, NY: Springer-Verlag.

Anexo I

Tabela: Quantidade de questionários aplicados por tipo de questionário, segundo a área especial em que foram aplicados e a sua fração diferenciada – Censo Demográfico 2022.

Área especial	Fração amostral diferenciada	Questionário da Amostra		Questionário Básico	
		Domicílios	Pessoas	Domicílios	Pessoas
Favelas ou Comunidades Urbanas	Total	284.176	867.840	702.646	2.112.611
	100%	88.668	277.729	88.668	277.729
	50%	56.656	176.711	60818	189274
	33%	29.920	91.353	63.042	192.764
	20%	105.356	311.736	455.118	1.351.484
	10%	3.576	10.311	35.000	101.360
Terras Indígenas	Total	89.956	399.334	113.889	504.044
	100%	70.474	315.451	70.474	315.451
	50%	8.958	39.888	11.765	54.133
	33%	5.975	24.748	12.695	53.205
	20%	4.549	19.247	18.955	81.255
Territórios Quilombolas	Total	48.734	158.086	51.846	168.058
	100%	42.922	139.422	42.922	139.422
	50%	4.427	14.068	4.730	15.415
	33%	717	2.643	1.445	5.319
	20%	668	1.953	2.749	7.902
Distritos e Subdistritos	Total	17.076	44.377	103.974	279.257
	50%	2.133	4.728	2.367	5.414
	33%	3.509	8.862	8.352	21.696
	20%	3.587	9.829	15.760	43.034
	10%	7.847	20.958	77.495	209.113

Fonte: IBGE, Diretoria de Pesquisas 2026.

Anexo II

Tabela: Composição dos conjuntos de restrições utilizados no processo de calibração dos resultados do Censo Demográfico 2022.

Restrição	Conjunto de Restrição					
	1	2	3	4	5	6
Total de pessoas	x	x	x	x	x	x
Total de domicílios	x	x	x	x	x	x
Total de homens	x	x	x	x	x	x
Pessoas de 0 a 4 anos	x	x	x			
Pessoas de 5 a 9 anos	x	x	x			
Pessoas de 10 a 14 anos	x	x	x			
Pessoas de 15 a 19 anos	x	x	x			
Pessoas de 20 a 24 anos	x	x	x			
Pessoas de 25 a 29 anos	x	x	x			
Pessoas de 30 a 34 anos	x	x	x			
Pessoas de 35 a 39 anos	x	x	x			
Pessoas de 40 a 44 anos	x	x	x			
Pessoas de 45 a 49 anos	x	x	x			
Pessoas de 50 a 54 anos	x	x				
Pessoas de 55 a 59 anos	x	x				
Pessoas de 60 a 64 anos	x	x				
Pessoas de 65 a 69 anos	x	x				
Pessoas de 70 a 74 anos	x	x				
Pessoas de 75 a 79 anos	x	x				
Pessoas de 80 ou mais anos	x	x	x	x		
Homens de 0 a 4 anos	x	x	x			
Homens de 5 a 9 anos	x	x	x			
Homens de 10 a 14 anos	x	x	x			
Homens de 15 a 19 anos	x	x	x			
Homens de 20 a 24 anos	x	x	x			
Homens de 25 a 29 anos	x	x	x			
Homens de 30 a 34 anos	x	x	x			
Homens de 35 a 39 anos	x	x	x			
Homens de 40 a 44 anos	x	x	x			
Homens de 45 a 49 anos	x	x	x			
Homens de 50 a 54 anos	x	x				
Homens de 55 a 59 anos	x	x				
Homens de 60 a 64 anos	x	x				
Homens de 65 a 69 anos	x	x				
Homens de 70 a 74 anos	x	x				
Homens de 75 a 79 anos	x	x				
Homens de 80 ou mais anos	x	x	x	x		
Pessoas na situação urbana	x	x	x	x	x	
Pessoas em domicílios particulares	x	x	x	x	x	
Total de domicílios particulares	x	x	x	x	x	
Total de domicílios particulares urbanos	x					
Total de domicílios particulares permanentes sem banheiro	x	x	x	x		
Total de domicílios particulares permanentes com um banheiro	x					
Total de domicílios particulares permanentes com mais de um banheiro	x					
Total de pessoas em Terras Indígenas	x	x	x	x	x	
Total de pessoas em Territórios Quilomboas	x	x	x	x	x	
Total de pessoas em Favelas ou Comunidades Urbanas	x	x	x	x	x	

Fonte: IBGE, Diretoria de Pesquisas, Coordenação de Métodos e Qualidade.

Anexo III

Tabela – Fração amostral, número total de municípios, população total, número de unidades domiciliares e número de pessoas por tipo de questionário, segundo as classes de população dos municípios – microdados de acesso controlado.

Classe de população dos municípios	Fração amostral	Nº de Municípios	População total	Questionário da Amostra		Questionário Básico	
				Unidades Domiciliares	Pessoas	Unidades Domiciliares	Pessoas
Até 2500	50%	289	606.617	112.996	300.389	114.473	303.721
Mais de 2.500 até 8.000	33%	1.808	8.814.725	1.058.445	2.909.443	2.143.265	5.879.050
Mais de 8.000 até 20.000	20%	1.673	21.076.063	1.474.121	4.214.089	5.878.193	16.757.921
Mais de 20.000 até 500.000	10%	1.751	109.956.835	3.952.428	11.179.394	35.089.889	98.503.958
Mais de 500.000	5%	49	62.626.516	1.091.973	2.936.264	22.281.075	59.684.054
Total	11%	5.570	203.080.756	7.689.963	21.539.579	65.506.895	181.128.704

Nota: A classificação dos municípios por população foi realizada baseada na estimativa populacional de 2019.

Tabela – Fração amostral, número total de municípios, população total, número de unidades domiciliares e número de pessoas por tipo de questionário, segundo as classes de população dos municípios – microdados de acesso público.

Classe de população dos municípios	Fração amostral	Nº de Municípios	População total	Questionário da Amostra		Questionário Básico	
				Unidades Domiciliares	Pessoas	Unidades Domiciliares	Pessoas
Até 2500	50%	289	606.617	112.995	300.371	114.473	303.721
Mais de 2.500 até 8.000	33%	1.808	8.814.725	1.058.439	2.909.296	2.143.265	5.879.050
Mais de 8.000 até 20.000	20%	1.673	21.076.063	1.474.100	4.213.615	5.878.193	16.757.921
Mais de 20.000 até 500.000	10%	1.751	109.956.835	3.952.408	11.178.978	35.089.889	98.503.958
Mais de 500.000	5%	49	62.626.516	1.091.972	2.936.248	22.281.075	59.684.054
Total	11%	5.570	203.080.756	7.689.914	21.538.508	65.506.895	181.128.704

Nota: A classificação dos municípios por população foi realizada baseada na estimativa populacional de 2019.