



Relatório Metodológico

PESQUISA
TIC DOMICÍLIOS 2025

Relatório Metodológico

TIC Domicílios 2025

O Comitê Gestor da Internet no Brasil (CGI.br), por meio do Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), departamento do Núcleo de Informação e Coordenação do Ponto BR (NIC.br), apresenta a metodologia da Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros: TIC Domicílios.

A pesquisa TIC Domicílios incorpora, em seu processo de coleta de dados, o público-alvo da pesquisa TIC Kids Online Brasil, que compreende indivíduos de 9 a 17 anos de idade. Desse modo, as duas pesquisas compartilham a forma de seleção dos indivíduos respondentes, descrita em detalhes na seção “Plano amostral”. Ainda que os dados tenham sido coletados conjuntamente, os resultados relativos às duas pesquisas são divulgados de maneira específica para cada público.

Objetivo da pesquisa

A pesquisa TIC Domicílios tem como objetivo principal medir a posse e o uso das tecnologias de informação e comunicação (TIC) entre a população residente no Brasil com 10 anos de idade ou mais.

Conceitos e definições

SETOR CENSITÁRIO

Segundo definição do Instituto Brasileiro de Geografia e Estatística (IBGE) para o Censo Demográfico, setor censitário é a menor unidade territorial formada por área contínua e com limites físicos identificados, em área urbana ou rural, com dimensão apropriada à realização de coleta de dados. O conjunto de setores censitários de um país cobre a totalidade de seu território nacional.

ÁREA

O domicílio pode ser urbano ou rural, segundo sua área de localização, tomando-se por base a legislação vigente por ocasião da realização do Censo Demográfico. Como situação urbana, consideram-se as áreas correspondentes às cidades (sedes municipais), às vilas (sedes distritais) ou às áreas urbanas isoladas. A situação rural abrange toda a área que está fora desses limites.

GRAU DE INSTRUÇÃO

Refere-se ao nível de ensino que o indivíduo estava frequentando ou havia frequentado, ainda que não tenha completado todo o ciclo. A coleta do grau de instrução é feita em 20 subcategorias, variando de “não frequentou escola” até “doutorado”.

RENDA FAMILIAR MENSAL

A renda familiar mensal é dada pela soma da renda de todos os moradores do domicílio, incluindo o respondente. Para a divulgação dos dados, são estabelecidas seis faixas de renda, iniciando-se pelo salário mínimo (SM) definido pelo Governo Federal. A primeira faixa representa a renda total do domicílio de até um SM, enquanto a sexta faixa representa rendas familiares superiores a dez SM:

- até um SM,
- mais de um SM até dois SM,
- mais de dois SM até três SM,
- mais de três SM até cinco SM,
- mais de cinco SM até dez SM,
- mais de dez SM.

CLASSE SOCIAL

O termo mais preciso para designar o conceito seria “classe econômica”. Entretanto, mantém-se “classe social” para fins da publicação das tabelas e análises relativas a esta pesquisa. A classificação econômica é baseada no Critério de Classificação Econômica Brasil (CCEB), conforme definido pela Associação Brasileira de Empresas de Pesquisa (Abep). A entidade utiliza para tal classificação a posse de alguns itens duráveis de consumo doméstico, mais o grau de instrução do chefe do domicílio declarado. A posse dos itens estabelece um sistema de pontuação em que a soma para cada domicílio resulta na seguinte classificação: classes econômicas A1, A2, B1, B2, C, D e E. O Critério Brasil foi atualizado em 2015, resultando em classificação não comparável à anteriormente vigente (Critério Brasil 2008). Para os resultados divulgados a partir de 2016, foi adotado o Critério Brasil 2015.

CONDIÇÃO DE ATIVIDADE

Refere-se à condição do respondente de 10 anos ou mais em relação à sua atividade econômica. A partir de uma sequência de quatro perguntas, obtêm-se sete classificações referentes à condição de atividade do entrevistado. Essas opções são classificadas em duas categorias, como consta na Tabela 1.

TABELA 1

—

Classificação da condição de atividade

Alternativas no questionário		Classificação da condição
Código	Descrição	Descrição
1	Trabalha em atividade remunerada.	Na força de trabalho
2	Trabalha em atividade não remunerada, como ajudante.	
3	Trabalha, mas está afastado.	
4	Tomou providência para conseguir trabalho nos últimos 30 dias.	
5	Não trabalha e não procurou trabalho nos últimos 30 dias.	Fora da força de trabalho

DOMICÍLIO PARTICULAR PERMANENTE

Refere-se ao domicílio particular localizado em unidade que se destina a servir de moradia (casa, apartamento ou cômodo). O domicílio particular permanente é a moradia de uma pessoa ou de um grupo de pessoas, em que o relacionamento é ditado por laços de parentesco, dependência doméstica ou normas de convivência.

USUÁRIO DE INTERNET

São considerados usuários de Internet os indivíduos que utilizaram a rede ao menos uma vez nos três meses anteriores à entrevista, conforme definição da União Internacional de Telecomunicações (UIT, 2020).

População-alvo

A população-alvo da pesquisa é composta de domicílios particulares permanentes brasileiros e da população com 10 anos ou mais residente em domicílios particulares permanentes no Brasil.

Unidade de análise e referência

A pesquisa possui duas unidades de análise e referência: os domicílios particulares permanentes e a população residente com 10 anos ou mais.

Domínios de interesse para análise e divulgação

Para as unidades de análise e referência, os resultados são divulgados para domínios definidos com base nas variáveis e nos níveis descritos a seguir.

Para as variáveis relacionadas a domicílios:

- **área:** corresponde à definição de setor — segundo critérios do IBGE, classificada como rural ou urbana;
- **região:** corresponde à divisão regional do Brasil — segundo critérios do IBGE, nas macrorregiões Norte, Nordeste, Sudeste, Sul e Centro-Oeste;
- **renda familiar:** corresponde à divisão da renda total dos domicílios e da população residente em faixas de salários mínimos. As faixas consideradas são até um SM, mais de um SM até dois SM, mais de dois SM até três SM, mais de três SM até cinco SM, mais de cinco SM até dez SM ou mais de dez SM;
- **classe social:** corresponde à divisão em A, B, C e DE, conforme o Critério Brasil.

Em relação às variáveis sobre os indivíduos, acrescentam-se aos domínios mencionados acima as seguintes características:

- **sexo:** corresponde à divisão em masculino ou feminino;
- **cor ou raça:** corresponde à divisão em branca, preta, parda, amarela e indígena;
- **grau de instrução:** corresponde à divisão em analfabeto/até Educação Infantil, até Ensino Fundamental, até Ensino Médio e até Ensino Superior;
- **faixa etária:** corresponde à divisão nas faixas de 10 a 15 anos, de 16 a 24 anos, de 25 a 34 anos, de 35 a 44 anos, de 45 a 59 anos e de 60 anos ou mais;
- **condição de atividade:** corresponde à divisão em “na força de trabalho” ou “fora da força de trabalho”.

Instrumento de coleta

INFORMAÇÕES SOBRE O INSTRUMENTO DE COLETA

Os dados são coletados por meio de um questionário estruturado, com perguntas fechadas e respostas predefinida (respostas únicas ou múltiplas). Para mais informações a respeito do questionário, ver item “Instrumento de coleta” em “Relatório de Coleta de Dados”.

Plano amostral

CADASTROS E FONTES DE INFORMAÇÃO

Para incorporar a nova base cadastral derivada do Censo Demográfico de 2022 do IBGE, o planejamento amostral da pesquisa foi refeito, visando aprimorar o alcance dos objetivos tradicionais da pesquisa, aumentar a capacidade de produção de estimativas por unidades da federação (objetivo novo) e melhorar a eficiência da amostra.

Dois cadastros elaborados pelo IBGE foram utilizados para implementar a amostra do projeto:

- Cadastro de setores censitários da Base Operacional Geográfica (BOG) do Censo Demográfico 2022 do IBGE (IBGE, 2024a);
- Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) do Censo Demográfico 2022 do IBGE (IBGE, 2024b).

O primeiro cadastro foi divulgado pelo IBGE em 14 de novembro de 2024 e é composto por três conjuntos de informações: malha setorial definitiva do Censo 2022, com polígonos descrevendo setores censitários; dados tabulares agregados por setores censitários dessa malha setorial; e ponto inicial e perímetros dos setores censitários.

O segundo cadastro (CNEFE) contém a relação dos endereços de todas as edificações encontradas pelo IBGE durante a coleta do Censo Demográfico 2022, classificados como estabelecimentos ou domicílios. Esse cadastro servirá de base para a etapa de atualização de endereços nos setores da amostra selecionada antes da seleção de endereços de domicílios para a pesquisa. Depois de atualizados os endereços de cada um dos setores da amostra, ele servirá de base à seleção da amostra de domicílios para a pesquisa.

Um aspecto importante da delimitação da população de pesquisa decorreu da exclusão de um conjunto de setores censitários da malha setorial. Tal exclusão é operação habitual nas pesquisas por amostragem domiciliar, e seguiu padrões vigentes em pesquisas de natureza similar. Foram excluídos:

- a. 1.101 setores classificados pelo IBGE como “Massas de água”, conforme a variável “cd_sit”;
- b. 180 setores classificados pelo IBGE como “Quartel”, conforme a variável “cd_tipo”;
- c. 143 setores classificados pelo IBGE como “Alojamento/acampamento”, conforme a variável “cd_tipo”;
- d. 7.805 setores classificados pelo IBGE como “Agrupamento indígena”, conforme a variável “cd_tipo”;
- e. 837 setores classificados pelo IBGE como “Unidade prisional”, conforme a variável “cd_tipo”;
- f. 716 setores classificados pelo IBGE como “Convento/hospital/ILPI (Instituição de Longa Permanência para Idosos)/IACA (Instituição de Acolhimento para Crianças e Adolescentes)”, conforme a variável “cd_tipo”;

- g. 2.085 setores classificados pelo IBGE como ‘Agrovila do PA (Projeto de Assentamento)’, conforme a variável “cd_tipo”;
- h. 5.591 setores classificados pelo IBGE como ‘Agrupamento quilombola’, conforme a variável “cd_tipo”.

Feitas as exclusões indicadas, o conjunto de setores censitários considerados elegíveis mantidos no cadastro para seleção da amostra de setores ficou então com 449.639 deles. A Tabela 2 a seguir apresenta distribuição por situação e tipo.

TABELA 2

—

Total de setores elegíveis no cadastro e totais de domicílios particulares ocupados e moradores nesses setores, por situação do setor

Situação	Setores	Domicílios	Moradores
Área urbana de alta densidade de edificações de cidade ou vila	308 451	58 891 797	162 192 390
Área urbana de baixa densidade de edificações de cidade ou vila	31 975	3 685 672	10 650 356
Núcleo urbano	12 038	1 347 629	3 932 118
Aglomerado rural – Povoado	13 997	1 222 429	3 712 325
Aglomerado rural – Núcleo rural	286	16 211	51 261
Aglomerado rural – Lugarejo	1 371	61 780	181 060
Área rural (exclusive aglomerados)	81 521	6 673 098	19 669 294
Total	449 639	71 898 616	200 388 804

Fonte: Arquivo Agregado de Setores Censitários do Censo Demográfico 2022 (IBGE, 2024a).

Uma das dificuldades percebidas durante a preparação do cadastro para a pesquisa foi o fato de que há muitos setores censitários com pequena quantidade de domicílios recenseados. A título de exemplo, encontramos 21.558 setores elegíveis com menos de 15 domicílios (4,8%) e 74.367 deles com 15 a menos de 60 domicílios (16,5%).

Esta dificuldade motivou a construção de unidades primárias de amostragem (UPA) mediante agregação ou combinação de setores censitários, de modo a evitar a seleção de setores muito pequenos, nos quais a pesquisa poderia eventualmente não conseguir obter a amostra desejada. Essa operação está descrita na seção “Formação de UPA”.

FORMAÇÃO DE UPA

Esta seção descreve o processo de construção das UPA, tendo por base a nova malha de setores censitários do IBGE descrita na seção anterior. Para realizar tal construção, foi utilizada a linguagem R, na qual foram implementados três algoritmos de otimização (Goldberg *et al.*, 2015), que foram aplicados dentro de estratos geográficos e de situação previamente definidos. Respeitada a estratificação geográfica, foram aplicados critérios associados à formação, contiguidade e capacidade. Ao final, cada UPA formada correspondia a um agrupamento de setores censitários que estava de acordo com os critérios definidos.

Dentro de cada município e estrato de situação, construíram-se dois tipos de UPA: (I) UPA com, obrigatoriamente, um mínimo de 60 domicílios particulares permanentes ocupados (DPPO) — critério de capacidade — cujos setores censitários fossem contíguos; (II) UPA que tenham, preferencialmente, um mínimo de 60 DPPO, mas com setores censitários não contíguos.

Adicionalmente, para as UPA do tipo I, buscou-se a maximização do número de UPA formadas. Em relação ao tipo II, cada UPA formada foi constituída, no máximo, por dois setores censitários que tivessem alta dissimilaridade quanto à variável de estratificação correspondente à porcentagem de moradores em domicílios com até 1 banheiro em cada setor. Mais especificamente, se um estrato de situação de um município tiver n setores censitários, sendo n par, serão formadas $n/2$ UPA e, no caso de n ser ímpar, haverá $n/2 + 1$ UPA.

A definição de UPA contemplando os critérios de capacidade e contiguidade supracitados diz respeito a um problema de agrupamento capacitado e conexo (Brito & Montenegro, 2010), em que cada objeto corresponde a um setor censitário e as UPA referem-se aos grupos formados. Este problema pode ser formulado como um problema de otimização em grafos (Ahuja *et al.*, 1993) no qual busca-se maximizar a função objetivo associada ao número de UPA formadas. Para a construção das UPA do tipo I foram implementados dois algoritmos denominados, respectivamente, heurística baseada em árvore geradora mínima (HAGM) e heurística baseada em junção de nós (HNO), sendo ambos embasados em conceitos de grafos. No caso das UPA do tipo II, foi desenvolvido um algoritmo que utiliza os conceitos de mediana e de posição do setor na lista ordenada com base na variável de estratificação.

Para a aplicação desses três algoritmos, descritos resumidamente a seguir, foi desenvolvido um código em linguagem R contendo um conjunto de procedimentos — em particular, um procedimento que efetua a leitura e validação do arquivo de vizinhanças e do arquivo de cadastro de setores do censo 2022, que contém os valores da variável de estratificação considerada. O arquivo de vizinhanças compreende uma relação dos códigos dos setores vizinhos a cada setor, sendo considerados vizinhos apenas aqueles pertencentes ao mesmo município.

As informações provenientes desses dois arquivos foram separadas por unidade da federação e os algoritmos desenvolvidos foram aplicados em cada um de seus municípios, tendo como entrada os dois arquivos supracitados.

Em linhas gerais, após a aplicação dos procedimentos de leitura e validação, é realizado, por unidade federativa (UF), um segundo procedimento que efetua a construção de uma estrutura de dados que reflete a combinação dos estratos geográficos e de situação, produzindo os nove tipos de recorte listados nas Tabela 3 e 4 a seguir.

TABELA 3

—

Recortes considerados na formação das UPA

Tipo do recorte (por município)	Critérios considerados
Capital_Urbano_ nao_especial	Capacidade e contiguidade
Capital_Urbano_favela_comunidade	Capacidade e contiguidade
Capital_Rural	Capacidade e contiguidade
Entorno_RM_Capital_e_Urbano_ nao_especial	Capacidade e contiguidade
Entorno_RM_Capital_e_Urbano_favela_e_comunidade	Capacidade e contiguidade
Entorno_RM_Capital_Rural	Capacidade e contiguidade
Interior_Urbano_ nao_Especial	Capacidade e máxima dissimilaridade
Interior_Urbano_Favela_Comunidade	Capacidade e máxima dissimilaridade
Interior_Rural	Capacidade e máxima dissimilaridade

TABELA 4

—

Total de UPA formadas segundo os 9 tipos de recorte

Recorte	Total de UPA
Capital_Urbano_ nao_especial	70 386
Capital_Urbano_favela_comunidade	17 188
Capital_Rural	985
Entorno_RM_Capital_e_Urbano_ nao_especial	47 323
Entorno_RM_Capital_e_Urbano_favela_e_comunidade	7 023
Entorno_RM_Capital_Rural	3 974
Interior_Urbano_ nao_Especial	102 932
Interior_Urbano_Favela_Comunidade	4 566
Interior_Rural	46 949
Total	301 326

Fonte: Arquivo Agregado de Setores Censitários do Censo Demográfico 2022 (IBGE, 2024a).

Por conta do critério de contiguidade estabelecido para os recortes de 1 a 6, foi implementado um procedimento que constrói, para cada município, a partir da informação do arquivo de vizinhanças, um grafo G com n vértices e m arestas. Esse grafo contempla o critério de contiguidade do problema e é utilizado como estrutura de dados de entrada para os algoritmos HAGM e HNO. Nesse grafo, cada vértice v_i corresponde a um setor censitário i e cada aresta $a_{ij} = (v_i, v_j)$ expressa uma relação de vizinhança entre dois vértices v_i e v_j , correspondentes a dois setores i e j que fazem parte do mesmo município e são vizinhos (têm fronteira). Adicionalmente, atribuiu-se a cada vértice v_i de G o valor correspondente ao número de DPPO do setor i e atribuiu-se a cada aresta a_{ij} a diferença absoluta (dissimilaridade) entre o número de DPPO dos setores i e j .

O algoritmo HAGM é baseado na construção de uma árvore geradora mínima (AGM) (Ahuja *et al.*, 1993), que corresponde a um grafo T que contém todos os n vértices de G (setores) e as $n - 1$ arestas de G com os menores valores de diferença absoluta.

Para construir um conjunto de k UPA, sendo k no máximo igual a n , em que todos os setores tenham número de DPPO ≥ 60 , aplica-se um procedimento de particionamento sobre T que efetua iterativamente a remoção de $k - 1$ arestas de T , assegurando de imediato, pelas propriedades da AGM, o atendimento da restrição de contiguidade — ou seja, cada aresta a_{ij} removida produz duas subárvores conexas correspondentes a duas partições contíguas candidatas à formação de duas UPA.

De forma a obter o maior número de UPA formadas ao final do processo de particionamento, em cada iteração são avaliadas todas as remoções possíveis das arestas de T , removendo-se aquela que produzir duas subárvores com a maior diferença entre as somas dos DPPO associados aos setores correspondentes a seus vértices, de forma que as duas partições (subárvores) T_1 e T_2 sejam viáveis (soma dos DPPO ≥ 60), definindo-se assim duas possíveis UPA.

Em seguida, a partição viável (subárvore T_1 ou T_2) com menor soma de DPPO é definida como uma nova UPA, que é adicionada ao conjunto atual de UPA em construção, enquanto a partição viável com maior soma de DPPO é associada à árvore T atual. Tal processo é repetido nas iterações seguintes até que não seja possível produzir, a partir da remoção de alguma aresta da árvore T atual, novas partições (UPA) viáveis. O pseudocódigo abaixo apresenta os passos básicos do algoritmo HAGM.

BOX 1

—
ALGORITMO HAGM

Construa o grafo G a partir da relação de vizinhança dos setores que compõem o município processado.
Construa a AGM T a partir de G .

$T_{\text{atual}} \leftarrow T$

part_viavel $\leftarrow 1$

$C_{\text{UPA}} \leftarrow \emptyset$ (Conjunto de UPA formadas)

Enquanto (part_viavel=1) **Faça**

Remova de T_{atual} cada aresta a_{ij} e avalie o total de DPPO nas duas partições produzidas (candidatas às UPA)

Se houver pelo menos uma remoção de aresta que produza duas partições viáveis T_1 e T_2 **Então**

Remova de T_{atual} a aresta a_{ij} que produz 2 partições **com a maior diferença** entre as somas dos DPPO

$C_{\text{UPA}} \leftarrow C_{\text{UPA}} \cup T_{\text{menor}}$ (subárvore correspondente à partição viável com menor soma de DPPO)

$T \leftarrow T_{\text{maior}}$ (subárvore correspondente à partição viável com maior soma de DPPO)

Senão

part_viavel $\leftarrow 0$

Fim-Se

Fim-Enquanto

Retorne C_{UPA}

Por conta do tempo de processamento demandado pelo algoritmo HAGM — à medida que o número de setores analisados aumenta — e a fim de validar os resultados produzidos quanto à qualidade das soluções, medida pelo maior número possível de UPA formadas por município, foi implementado um segundo algoritmo, denominado HNO. Esse algoritmo também utiliza, como estrutura de dados de entrada, o grafo G definido anteriormente, mas diferencia-se do algoritmo HAGM quanto à forma de construção das UPA, sendo aplicado em dois passos básicos:

- **Passo 1:** aplica-se um procedimento que identifica todos os setores do município em processamento que têm o número de DPPO maior ou igual a 60, sendo cada um desses setores automaticamente definido como uma UPA.
- **Passo 2:** após o passo 1, se existirem setores com número de DPPO inferior a 60, busca-se a formação de novas UPA, efetuando-se junções de tais setores avaliando se são contíguos (tendo por base o grafo G) e objetivando a maximização do número de UPA. Nesse sentido, para formar novas UPA, são avaliadas diversas combinações desses setores para efetuar as junções, privilegiando-se aquelas com os menores quantitativos de setores e que atendam aos critérios de capacidade e contiguidade.

Os algoritmos HAGM e HNO foram aplicados em todos os municípios que são definidos pelos recortes de 1 até 6.

Por fim, para a construção das UPA associadas aos municípios que são definidos pelos recortes de 7 até 9, em que a restrição de contiguidade não é considerada e a restrição de capacidade associada ao número de DPPO é desejável, mas não mandatária, foi implementado um algoritmo denominado heurística de maximização da mediana (HMAXMED). Esse algoritmo foi desenvolvido para possibilitar a formação de UPA com dois setores censitários cada, tal que os setores alocados a uma mesma UPA sejam o mais dissimilares entre si (distância euclidiana) quanto à variável de estratificação correspondente à porcentagem de moradores em domicílios com até 1 banheiro em cada setor.

O algoritmo foi aplicado considerando-se dois possíveis casos: (1) número par de setores e (2) número ímpar de setores. No caso 1, os setores são ordenados crescentemente, em relação aos valores da variável de estratificação, formando $n/2$ UPA mediante a união do setor na posição 1 com o da posição $n/2$, o setor da posição 2 com o da posição $(n/2) + 1$ e assim sucessivamente. No caso 2, sendo n ímpar, em um primeiro passo são avaliadas todas as combinações dos n setores tomados $(n - 1)$ a $(n - 1)$ (C_n^{n-1}), sendo produzidos n subconjuntos formados por $(n - 1)$ setores. Em cada subconjunto, ordenam-se crescentemente os seus setores pela variável de estratificação e são efetuadas suas junções de forma análoga àquela descrita no caso 1, calculando-se, em seguida, a mediana das dissimilaridades para os $n/2$ pares de setores. O subconjunto de n setores associados aos $n/2$ pares com a maior dissimilaridade mediana e que cumprem a restrição de capacidade é considerado conjunto solução das UPA. Por fim, o setor que não está contemplado neste conjunto de $n/2$ UPA é definido como uma nova UPA, separadamente.

ESTRATIFICAÇÃO DA AMOSTRA

O planejamento amostral proposto para a pesquisa considera uma estratégia baseada em amostragem estratificada e conglomerada em 3 ou 4 estágios — para compreensão dos termos técnicos aqui empregados, veja Silva *et al.* (2020).

A estratificação foi realizada em três etapas, sendo a primeira de natureza geográfica, associada à demanda de estimativas para domínios de interesse pré-definidos; a segunda baseou-se na separação de UPA de cada estrato natural conforme sua situação (com três categorias: “urbana não especial”, “urbana de comunidade ou favela” e “rural”); a terceira foi de natureza estatística, visando ampliar a eficiência do plano amostral mediante agrupamento de setores de nível socioeconômico similares.

Para a estratificação geográfica, foram definidos como domínios de interesse:

- a. Os municípios das capitais das 26 unidades da federação e mais o Distrito Federal, num total de 27 estratos naturais.
- b. Os conjuntos de municípios dos entornos de 11 regiões metropolitanas sediadas nas capitais — Manaus, Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba, Porto Alegre — e da RIDE (Região Integrada de Desenvolvimento) do Distrito Federal e Entorno, totalizando 12 estratos geográficos naturais, uma vez que, no caso da RIDE do DF e Entorno, há municípios em dois estados: Minas Gerais e Goiás.

- c. Os conjuntos de municípios do interior de 26 estados, obtidos por exclusão daqueles já considerados nos estratos naturais formados pelos domínios de interesse *a* e *b* acima.

Sendo assim, ao final dessa etapa de estratificação, foram formados 65 estratos naturais, cujas descrições podem ser vistas na Tabela 5 a seguir.

TABELA 5

—
Total de UPA na população, por estrato natural

Unidade da federação	Estrato natural	Número de UPA (população)
Rondônia	11 - Capital	708
Rondônia	11 - Interior	1 251
Acre	12 - Capital	638
Acre	12 - Interior	589
Amazonas	13 - Capital	3 098
Amazonas	13 - Interior	1 770
Amazonas	13 - RM de Manaus (AM)	991
Roraima	14 - Capital	479
Roraima	14 - Interior	283
Pará	15 - Capital	1 995
Pará	15 - Interior	5 555
Pará	15 - RM de Belém (PA)	1 776
Amapá	16 - Capital	647
Amapá	16 - Interior	363
Tocantins	17 - Capital	609
Tocantins	17 - Interior	1 777
Maranhão	21 - Capital	1 674
Maranhão	21 - Interior	6 573
Piauí	22 - Capital	1 437
Piauí	22 - Interior	2 976
Ceará	23 - Capital	4 041
Ceará	23 - Interior	6 605
Ceará	23 - RM de Fortaleza (CE)	2 664

CONTINUA ►

▶ CONTINUA

Unidade da federação	Estrato natural	Número de UPA (população)
Rio Grande do Norte	24 - Capital	1 068
Rio Grande do Norte	24 - Interior	3 013
Paraíba	25 - Capital	1 602
Paraíba	25 - Interior	4 383
Pernambuco	26 - Capital	2 739
Pernambuco	26 - Interior	5 738
Pernambuco	26 - RM de Recife (PE)	4 187
Alagoas	27 - Capital	1 475
Alagoas	27 - Interior	2 501
Sergipe	28 - Capital	1 087
Sergipe	28 - Interior	2 318
Bahia	29 - Capital	4 359
Bahia	29 - Interior	11 622
Bahia	29 - RM de Salvador (BA)	2 044
Minas Gerais	31 - Capital	4 878
Minas Gerais	31 - Interior	19 636
Minas Gerais	31 - RIDE do Distrito Federal e Entorno	192
Minas Gerais	31 - RM de Belo Horizonte (MG)	6 043
Espírito Santo	32 - Capital	621
Espírito Santo	32 - Interior	5 414
Rio de Janeiro	33 - Capital	13 022
Rio de Janeiro	33 - Interior	6 983
Rio de Janeiro	33 - RM do Rio de Janeiro (RJ)	12 718
São Paulo	35 - Capital	25 597
São Paulo	35 - Interior	27 989
São Paulo	35 - RM de São Paulo (SP)	18 299
Paraná	41 - Capital	3 125
Paraná	41 - Interior	8 775
Paraná	41 - RM de Curitiba (PR)	2 907
Santa Catarina	42 - Capital	956

► CONCLUSÃO

Unidade da federação	Estrato natural	Número de UPA (população)
Santa Catarina	42 - Interior	8 521
Rio Grande do Sul	43 - Capital	2 624
Rio Grande do Sul	43 - Interior	8 744
Rio Grande do Sul	43 - RM de Porto Alegre (RS)	5 202
Mato Grosso do Sul	50 - Capital	1 597
Mato Grosso do Sul	50 - Interior	2 144
Mato Grosso	51 - Capital	1 094
Mato Grosso	51 - Interior	4 085
Goiás	52 - Capital	2 242
Goiás	52 - Interior	4 839
Goiás	52 - RIDE do Distrito Federal e Entorno	1 297
Distrito Federal	53 - Capital	5 147
Total		301 326

A segunda etapa da estratificação usou a situação da UPA (com três categorias: “urbana não especial”, “urbana de comunidade ou favela” e “rural”). Em seguida, o tamanho total da amostra de cada estrato natural foi alocado aos estratos de situação correspondentes, utilizando uma alocação proporcional à potência 0,8 do número de UPA existentes no estrato, mas mantendo um mínimo de 3 UPA na amostra de cada estrato. O emprego dessa alocação visou reduzir situações de amostras muito pequenas nos menores estratos.

Depois de separadas por situação, as UPA foram ainda estratificadas com base na porcentagem de pessoas vivendo em domicílios sem banheiro ou com até 1 banheiro, por UPA. Essa variável de estratificação foi utilizada porque, até o momento da seleção da amostra, não havia dados sobre a renda média, por setor censitário, da pessoa responsável pelo domicílio. Considerando a experiência de censos anteriores, essa variável de estratificação é uma das que têm maior poder preditivo dos níveis de renda dos setores censitários.

Nessa etapa da estratificação, um estrato formado pela combinação do estrato natural e da situação da UPA poderia ou não ser subdividido em estratos socioeconômicos. Estratos com tamanho amostral menor que 8 UPA não foram subdivididos. Para estratos com 8 ou mais UPA na amostra, a subdivisão teve número de estratos que variou conforme a Tabela 6.

TABELA 6

—

Número de estratos formados segundo o tamanho da amostra de UPA nos estratos natural e de situação

Número de UPA na amostra	Número de estratos
8 a 11	2
12 a 18	3
19 a 24	4
25 ou mais	5

A formação dos estratos de nível socioeconômico utilizou os quantis da porcentagem de pessoas vivendo em domicílios sem banheiro ou com até 1 banheiro e a alocação igual da amostra nos estratos formados. Assim, por exemplo, em um estrato natural e de situação com tamanho de amostra de 20 UPA, foram formados quatro estratos com base nos quartis da distribuição da variável de estratificação, sendo alocadas 5 UPA em cada um dos estratos formados ao final. Esse processo resultou na formação de 378 estratos, nos quais foi então realizada a seleção da amostra de UPA.

ALOCÇÃO DA AMOSTRA

A alocação da amostra da UPA nos estratos naturais foi definida com base em dois critérios: (1) alocar amostras maiores nos estratos de capitais e do DF para permitir a elaboração de estimativas com precisão razoável para esses domínios; e (2) alocação de amostras iguais nos demais estratos naturais, visando à obtenção de estimativas de precisão similar para tais estratos. Os tamanhos amostrais foram definidos após análises da precisão de estimativas obtidas nas últimas edições da pesquisa TIC Domicílios. A Tabela 7 a seguir apresenta a alocação da amostra de UPA.

TABELA 7

—

Total de UPA na amostra por estrato natural

Unidade da federação	Estrato natural	Número de UPA (amostra)
Rondônia	11 - Capital	30
Rondônia	11 - Interior	30
Acre	12 - Capital	30
Acre	12 - Interior	30
Amazonas	13 - Capital	40
Amazonas	13 - Interior	30
Amazonas	13 - RM de Manaus (AM)	30
Roraima	14 - Capital	30
Roraima	14 - Interior	30
Pará	15 - Capital	40
Pará	15 - Interior	30
Pará	15 - RM de Belém (PA)	30
Amapá	16 - Capital	30
Amapá	16 - Interior	30
Tocantins	17 - Capital	30
Tocantins	17 - Interior	30
Maranhão	21 - Capital	30
Maranhão	21 - Interior	30
Piauí	22 - Capital	30
Piauí	22 - Interior	30
Ceará	23 - Capital	40
Ceará	23 - Interior	30
Ceará	23 - RM de Fortaleza (CE)	30
Rio Grande do Norte	24 - Capital	30
Rio Grande do Norte	24 - Interior	30
Paraíba	25 - Capital	30
Paraíba	25 - Interior	30
Pernambuco	26 - Capital	40

CONTINUA ►

▶ CONTINUA

Unidade da federação	Estrato natural	Número de UPA (amostra)
Pernambuco	26 – Interior	30
Pernambuco	26 – RM de Recife (PE)	30
Alagoas	27 – Capital	30
Alagoas	27 – Interior	30
Sergipe	28 – Capital	30
Sergipe	28 – Interior	30
Bahia	29 – Capital	40
Bahia	29 – Interior	30
Bahia	29 – RM de Salvador (BA)	30
Minas Gerais	31 – Capital	40
Minas Gerais	31 – Interior	30
Minas Gerais	31 – RIDE do Distrito Federal e Entorno	15
Minas Gerais	31 – RM de Belo Horizonte (MG)	30
Espírito Santo	32 – Capital	30
Espírito Santo	32 – Interior	30
Rio de Janeiro	33 – Capital	40
Rio de Janeiro	33 – Interior	30
Rio de Janeiro	33 – RM do Rio de Janeiro (RJ)	30
São Paulo	35 – Capital	40
São Paulo	35 – Interior	30
São Paulo	35 – RM de São Paulo (SP)	30
Paraná	41 – Capital	40
Paraná	41 – Interior	30
Paraná	41 – RM de Curitiba (PR)	30
Santa Catarina	42 – Capital	30
Santa Catarina	42 – Interior	30
Rio Grande do Sul	43 – Capital	40
Rio Grande do Sul	43 – Interior	30
Rio Grande do Sul	43 – RM de Porto Alegre (RS)	30
Mato Grosso do Sul	50 – Capital	30

► CONCLUSÃO

Unidade da federação	Estrato natural	Número de UPA (amostra)
Mato Grosso do Sul	50 - Interior	30
Mato Grosso	51 - Capital	30
Mato Grosso	51 - Interior	30
Goiás	52 - Capital	40
Goiás	52 - Interior	30
Goiás	52 - RIDE do Distrito Federal e Entorno	15
Distrito Federal	53 - Capital	70
Total		2 070

SELEÇÃO DA AMOSTRA

SELEÇÃO DE UPA

Em cada um dos estratos formados, a seleção de UPA foi feita usando Amostragem com Probabilidades Proporcionais ao Tamanho de Pareto (Rosén, 2000). Esse método também foi adotado pelo IBGE para realizar a amostragem do Sistema Integrado de Pesquisas Domiciliares (Freitas & Antonaci, 2014). A medida de tamanho considerada para a seleção de UPA foi definida como:

$$T_i = \max(10^{0,8}; D_i^{0,8})$$

Em que D_i representa o número de domicílios particulares ocupados na UPA, conforme apurado no Censo Demográfico 2022. O uso da potência 0,8 reduz a assimetria da distribuição dos tamanhos das UPA e o limite inferior foi especificado para evitar que UPA com probabilidades de inclusão muito pequenas acabassem tendo pesos demasiado grandes na hora da ponderação da amostra.

Vale notar que, a despeito da aplicação do algoritmo descrito para formar UPA que tivessem tamanho mínimo adequado, por motivos diversos algumas poucas UPA terminaram o processo com tamanhos ainda inferiores ao limite mínimo desejável. Assim, foi importante manter a ideia de que UPA que possuíssem menos de 10 domicílios particulares ocupados tivessem sua medida de tamanho imputada em valor correspondente a esse patamar inferior, após a transformação potência adotada.

Em UPA com até dois setores censitários, ambos foram pesquisados. Naquelas formadas por três ou mais setores, foi feito o sorteio de dois setores censitários para participar da pesquisa. O sorteio de setores censitários dentro das UPA também foi feito usando Amostragem com Probabilidades Proporcionais ao Tamanho de Pareto (Rosén, 2000). A medida de tamanho considerada para a seleção de setores utilizou a mesma transformação considerada para as UPA — nesse caso, o número de domicílios particulares ocupados foi o de cada setor censitário.

Após a aplicação desse processo, foram selecionadas UPA e setores censitários para a amostra da pesquisa. Uma distinção importante ocorreu na etapa de seleção de domicílios dentro dos setores censitários. Nas UPA dos estratos de capitais e das regiões metropolitanas de capitais consideradas, foram sempre selecionados 15 domicílios por UPA. Quando a UPA em questão tinha dois setores na amostra, foram selecionados 8 domicílios no maior dos dois setores e 7 domicílios no menor deles. Já nas UPA dos estratos naturais do interior, foram sempre selecionados 15 domicílios por setor na amostra, tenha a UPA um ou dois setores na amostra correspondente.

SELEÇÃO DOS DOMICÍLIOS E RESPONDENTES

A seleção de domicílios particulares permanentes foi feita por amostragem aleatória simples. Em uma primeira etapa de trabalho, os entrevistadores efetuaram o procedimento de listagem ou arrolamento de todos os domicílios existentes na UPA (aproximadamente dois setores censitários) para obter um cadastro completo e atualizado. Após esse levantamento atualizado da quantidade de domicílios por UPA selecionada, foram escolhidos aleatoriamente cerca de 30 domicílios por UPA que seriam visitados para entrevista.

Todos os domicílios da amostra responderam ao questionário TIC Domicílios – Módulo A: Acesso às tecnologias de informação e comunicação no domicílio.

Para a atribuição de qual pesquisa seria aplicada no domicílio (TIC Domicílios ou TIC Kids Online Brasil), todos os residentes de cada domicílio informante da pesquisa foram listados e a seleção ocorreu da seguinte maneira:

1. Quando não há residentes na faixa etária entre 9 e 17 anos, é realizada a entrevista da pesquisa TIC Domicílios com o morador de 18 anos ou mais selecionado aleatoriamente entre os residentes do domicílio.
2. Quando há residentes na faixa etária entre 9 e 17 anos, é gerado um número aleatório entre 0 e 1, e:
 - a. Se o número gerado é menor ou igual a 0,54, a entrevista da pesquisa TIC Kids Online Brasil é realizada com o residente de 9 a 17 anos selecionado aleatoriamente entre os moradores do domicílio dessa faixa etária e com o responsável por esse residente selecionado.
 - b. Se o número gerado é maior do que 0,54 e menor ou igual a 0,89, a entrevista da pesquisa TIC Domicílios é realizada com residente de 10 a 17 anos selecionado aleatoriamente entre os moradores do domicílio dessa faixa etária.
 - Em domicílios selecionados para a realização da pesquisa TIC Domicílios (com um residente de 10 a 17 anos) que só tenha residentes de 9 anos ou menos, além de maiores de 18 anos, a pesquisa TIC Domicílios é aplicada ao morador de 18 anos ou mais selecionado aleatoriamente.
 - c. Se o número gerado é maior do que 0,89, a entrevista da pesquisa TIC Domicílios é realizada com residente de 18 anos ou mais selecionado aleatoriamente entre os moradores do domicílio dessa faixa etária.

A seleção do morador em cada domicílio selecionado para responder à pesquisa é realizada após a listagem de todos os moradores do domicílio.

Coleta de dados em campo

MÉTODO DE COLETA

A coleta dos dados é realizada com o método CAPI (do inglês *computer-assisted personal interviewing*), que consiste em ter o questionário programado em um *software* para *tablet* e aplicado por entrevistadores em interação face a face.

Processamento de dados

PROCEDIMENTOS DE PONDERAÇÃO

O processo de seleção de cada domicílio e morador, como descrito anteriormente, estabelece uma probabilidade de seleção inicial para cada UPA. Com base no resultado da coleta de dados, são realizadas correções de não resposta para cada etapa do processo de seleção. Essas etapas são descritas a seguir.

PONDERAÇÃO DAS UPA

Cada UPA possui uma probabilidade de seleção, como descrito em Seleção de UPA. O inverso dessa probabilidade corresponde ao peso básico de cada UPA selecionada. A partir da coleta dos dados, pode ocorrer de não serem coletadas respostas de domicílios para uma UPA. Nesse caso, a correção de não resposta é feita considerando-se que a não resposta é aleatória dentro do estrato. A correção dos pesos das UPA respondentes por estrato é dada pela Fórmula 1.

FÓRMULA 1

$$w_{ih}^r = w_{ih} \times \frac{\sum_{h=1}^H w_{ih}}{\sum_{h=1}^H w_{ih} \times I_h^r}$$

w_{ih}^r é o peso da UPA i do estrato h corrigido para não resposta

w_{ih} é o peso básico do desenho amostral da UPA i do estrato h

I_h^r é uma variável indicadora que recebe valor 1 se a UPA i do estrato h teve ao menos um domicílio respondente e 0, caso contrário

PONDERAÇÃO DOS DOMICÍLIOS NAS UPA

Da mesma forma que cada UPA possui uma probabilidade de seleção inicial, cada domicílio também possui uma probabilidade de seleção inicial. Essa probabilidade é determinada como a razão entre 15 (número de domicílios selecionados por setor censitário) e o número de domicílios elegíveis em cada setor censitário que compõe a UPA.

O primeiro fator da construção de pesos dos domicílios corresponde à estimativa do total de domicílios elegíveis no setor censitário. Consideram-se elegíveis os domicílios particulares permanentes que possuem população apta a responder às pesquisas (excluem-se apenas domicílios com indivíduos que não se comunicam em português ou que apresentam outras condições que impossibilitam a realização da pesquisa), conforme Fórmula 2.

FÓRMULA 2

$$E_{jih} = d_{jih} \times \frac{d_{jih}^E}{d_{jih}^A}$$

E_{jih} é a estimativa do total de domicílios elegíveis no setor censitário j da UPA i do estrato h

d_{jih}^E é o total de domicílios elegíveis abordados no setor censitário j da UPA i do estrato h

d_{jih}^A é o total de domicílios abordados no setor censitário j da UPA i do estrato h

d_{jih} é o total de domicílios arrolados no setor censitário j da UPA i do estrato h

O segundo fator da fórmula corresponde à proporção de domicílios elegíveis entre os domicílios abordados no setor censitário. O peso do domicílio em um setor censitário é dado pela Fórmula 3.

FÓRMULA 3

$$w_{jih} = \frac{E_{jih}}{\sum_{k=1}^{n_{jih}} I_{kjih}^r}$$

w_{jih} é o peso dos domicílios no setor censitário j da UPA i do estrato h corrigido para não resposta no setor censitário

E_{jih} é a estimativa do total de domicílios elegíveis no setor censitário j da UPA i do estrato h

I_{kjih}^r é uma variável indicadora que recebe valor 1 se o domicílio k do setor censitário j da UPA i do estrato h respondeu à pesquisa e 0, caso contrário

n_{jih} corresponde ao número de domicílios selecionados no setor censitário j da UPA i do estrato h

Assim como ocorre com as UPA, existem domicílios selecionados que se recusam a participar da pesquisa. Em alguns casos, um setor censitário de uma UPA pode não ter domicílios respondentes. Assim, faz-se necessário corrigir a não resposta do setor censitário dentro da UPA.

A correção de não resposta para os domicílios dentro da UPA é realizada após o cálculo dos pesos dos domicílios nos setores censitários, como apresentado anteriormente. Essa correção é realizada pela Fórmula 4.

FÓRMULA 4

$$w_{jih}^r = w_{jih} \times \frac{SC_{ih}}{\sum_{j=1}^{SC_{ih}} I_{ih}^r}$$

w_{jih}^r é o peso dos domicílios no setor censitário j da UPA i do estrato h corrigido para não resposta na UPA

w_{jih} é o peso dos domicílios no setor censitário j da UPA i do estrato h corrigido para não resposta no setor censitário

SC_{ih} é o total de setores censitários que compõem a UPA i do estrato h

I_{ih}^r é uma variável indicadora que recebe valor 1 se o setor censitário j da UPA i do estrato h teve ao menos um domicílio respondente e 0, caso contrário

O peso final de cada domicílio, corrigido para não resposta, é dado por:

$$w_{jih}^d = w_{ih}^r \times w_{jih}^r$$

Calibração dos domicílios

A partir do peso domiciliar corrigido para não resposta (w_{jih}^d), é feita a calibração desses pesos para totais conhecidos de domicílios e da população em geral obtidos com base em estimativas na Pesquisa Nacional por Amostra de Domicílios Contínua (Pnad Contínua) mais recente disponível (IBGE, 2023).

O método de calibração considera características de totais domiciliares e da população separadamente. O método utilizado é o *iterative proportional updating* (IPU) (Ye *et al.*, 2009). Esse algoritmo permite estabelecer pesos iguais para as pessoas moradoras de um mesmo domicílio, respeitando totais marginais domiciliares e da população. A metodologia é aplicada ao conjunto de residentes que compõem a amostra e são listados no quadro de moradores, com todos eles recebendo inicialmente o mesmo peso domiciliar calculado w_{jih}^d .

As características utilizadas na calibração são listadas a seguir.

Para domicílios:

- UF (2021 a 2024),
- área (rural ou urbana),
- tamanho do domicílio (1, 2, 3, 4, 5 e 6 ou mais residentes).

Para pessoas:

- grande região,
- área (rural ou urbana),
- sexo,
- faixa etária (0 a 2 anos, 3 a 5 anos, 6 a 8 anos, 9 anos, 10 a 15 anos, 16 a 24 anos, 25 a 34 anos, 35 a 44 anos, 45 a 59 anos, 60 anos ou mais).

Como resultado, obtém-se um peso final para cada domicílio, denotado por w_{jih}^c , que corresponde ao peso do domicílio no setor censitário j da UPA i do estrato h , corrigido para não resposta e calibrado para totais populacionais de domicílios e de pessoas.

A calibração dos pesos é implementada utilizando-se o pacote *mlfit*¹ do *software* estatístico livre R.

PONDERAÇÃO DOS INFORMANTES EM CADA DOMICÍLIO

Em cada domicílio selecionado, a pesquisa TIC Domicílios é aplicada de acordo com a composição do domicílio, por meio de um processo aleatório de seleção da pesquisa a ser aplicada e do respondente. O peso básico de cada respondente em cada pesquisa é dado pelas Fórmulas 5 ou 6.

Morador de 10 a 17 anos

FÓRMULA 5

$$w_{l/kjih}^T = \frac{1}{0,35 \times (1-p^*)} \times P_{kjih}^T$$

$w_{l/kjih}^T$ é o peso do respondente de 10 a 17 anos no domicílio k do setor censitário j da UPA i do estrato h

P_{kjih}^T é o número de pessoas na faixa etária de 10 a 17 anos no domicílio k do setor censitário j da UPA i do estrato h

Morador de 18 anos ou mais

FÓRMULA 6

$$w_{l/kjih}^A = \frac{1}{0,11 \times (p^* \times 0,35)} \times P_{kjih}^T$$

$w_{l/kjih}^A$ é o peso do respondente de 18 anos ou mais no domicílio k do setor censitário j da UPA i do estrato h

P_{kjih}^T é o número de pessoas na faixa etária de 18 anos ou mais no domicílio k do setor censitário j da UPA i do estrato h

p^* é a estimativa da proporção de domicílios com apenas moradores de 9 anos de idade em relação ao total de domicílios com moradores de 9 a 17 anos, obtida a partir dos microdados mais recentes disponíveis da PNAD Contínua. Nos domicílios selecionados para a realização da pesquisa TIC Domicílios – Indivíduos com morador de 10 a 17 anos que possuam apenas moradores de 9 anos ou menos, além de moradores de 18 anos ou mais, deve-se realizar a pesquisa TIC Domicílios – Indivíduos com um morador de 18 anos ou mais selecionado aleatoriamente.

¹ Ver <https://cran.r-project.org/web/packages/mlfit/>

PESO DE CADA INFORMANTE

O peso final de cada indivíduo entrevistado na pesquisa é dado pela multiplicação dos pesos de cada etapa da construção da ponderação.

- a. Peso do informante da pesquisa TIC Domicílios (com morador de 10 a 17 anos):

$$w_{lkjih} = w_{jih}^c \times w_{l/kjih}^T$$

- b. Peso do informante da pesquisa TIC Domicílios (com morador de 18 anos ou mais):

$$w_{lkjih} = w_{jih}^c \times w_{l/kjih}^A$$

CALIBRAÇÃO DO PESO DE CADA INFORMANTE

Os pesos das entrevistas são calibrados de forma a refletir algumas estimativas de contagens populacionais conhecidas ou estimadas com boa precisão, obtidas a partir da Pnad Contínua mais recente disponível, como também é feito para os domicílios. Esse procedimento visa, com a correção de não resposta, corrigir vieses associados à não resposta diferencial de grupos específicos da população, para o conjunto de informantes selecionados nos domicílios para responder à pesquisa.

As variáveis consideradas para a calibração dos pesos dos indivíduos da pesquisa TIC Domicílios são: sexo, faixa etária (seis categorias: de 10 a 15 anos, de 16 a 24 anos, de 25 a 34 anos, de 35 a 44 anos, de 45 a 59 anos e de 60 anos ou mais), área (duas categorias: urbana ou rural), estratos TIC, condição de atividade (duas categorias: na força de trabalho e fora da força de trabalho) e grau de instrução (quatro categorias: analfabeto/Educação Infantil, Ensino Fundamental, Ensino Médio ou Ensino Superior).

A calibração dos pesos é implementada utilizando-se a função *calibrate* da biblioteca *survey* (Lumley, 2010) do *software* estatístico livre R.

ERROS AMOSTRAIS

As estimativas das margens de erro levam em consideração o plano amostral estabelecido para a pesquisa. Para os indivíduos respondentes, foi utilizado o método de replicação, por meio da função *as.svrepdesign* do pacote *survey* do R. Nesse método, são gerados 200 pesos de réplica, que correspondem a 200 amostras com reposição da amostra original, mantendo o desenho amostral (estratificado e conglomerado).

O método de replicação também foi utilizado para a estimação das margens de erro dos domicílios respondentes da pesquisa. Nesse caso, como o procedimento de calibração adotado não está disponível no pacote *survey* do R, as réplicas foram geradas na base do quadro de moradores a partir do algoritmo descrito a seguir:

1. São geradas 200 réplicas com pesos corrigidos apenas para não resposta, resultando em uma base com 201 pesos.
2. O peso original, corrigido para não resposta, é calibrado para os totais de domicílios e de pessoas por meio do método IPU.
3. Os 200 pesos de réplica são então calibrados utilizando os 200 pesos de réplica disponíveis na PNAD Contínua.

Como resultado, obtém-se uma base de dados de domicílios com 201 pesos: o peso que fornece as estimativas pontuais e 200 pesos de réplica, utilizados para o cálculo dos erros das estimativas pontuais. Essa metodologia de ajuste está descrita em Opsomer e Erciulescu (2021).

A partir das variâncias estimadas, optou-se por divulgar os erros amostrais expressos pela margem de erro. Para a divulgação, as margens de erro foram calculadas com nível de confiança de 95%. Assim, se a pesquisa fosse repetida diversas vezes, em aproximadamente 19 de cada 20 vezes o intervalo estimado conteria o verdadeiro valor populacional.

Normalmente, também são apresentadas outras medidas derivadas dessa estimativa de variabilidade, tais como erro padrão, coeficiente de variação e intervalo de confiança.

O cálculo da margem de erro considera o produto do erro padrão (a raiz quadrada da variância) por 1,96, valor crítico da distribuição normal padrão associado a um nível de confiança de 95%. Esses cálculos foram realizados para cada variável em todas as tabelas. Portanto, todas as tabelas de indicadores incluem as margens de erro associadas a cada estimativa apresentada em cada célula da tabela.

Disseminação de dados

Os resultados desta pesquisa são apresentados de acordo com as variáveis descritas no item Domínios de interesse para análise e divulgação.

Arredondamentos fazem com que, em alguns resultados, a soma das categorias parciais difira de 100% em questões de resposta única. O somatório de frequências em questões de respostas múltiplas usualmente é diferente de 100%. Vale ressaltar que, nas tabelas de resultados, o hífen (-) é utilizado para representar a não resposta ao item. Por outro lado, como os resultados são apresentados sem casa decimal, as células com valor zero indicam que houve resposta ao item, mas ele é explicitamente maior do que zero e menor do que um.

Os resultados desta pesquisa são publicados em formato *online* e disponibilizados no *website* (<https://www.cetic.br/>) e no portal de visualização de dados do Cetic.br|NIC.br (<https://data.cetic.br/>). As tabelas de proporções, totais e margens de erro calculadas para cada indicador estão disponíveis para *download* em português, inglês e espanhol. Mais informações sobre a documentação, os metadados e as bases de microdados estão disponíveis na página de microdados (<https://www.cetic.br/microdados/>).

Referências

Ahuja, R. K., Magnanti, T. L., & Orlin, J. (1993). *Network flows: Theory, algorithms, and applications*. Pearson.

Brito, J. A. M., & Montenegro, F. M. T. (2010). Um algoritmo VNS aplicado ao problema de definição de áreas de ponderação. *Simpósio de pesquisa operacional e logística da marinha*. <https://www.marinha.mil.br/spolm/sites/www.marinha.mil.br/spolm/files/74163.pdf>

Freitas, M. P. S., & Antonaci, G. A. (2014). *Sistema integrado de pesquisas domiciliares: amostra mestra 2010 e amostra da Pnad Contínua* (Texto para discussão, n. 50). IBGE. <https://biblioteca.ibge.gov.br/visualizacao/livros/liv86747.pdf>

Goldbarg, M. C., Goldbarg, E. G., & Luna, H. P. L. (2015). *Otimização combinatória e meta-heurísticas: algoritmos e aplicações*. GEN LTC.

Instituto Brasileiro de Geografia e Estatística. (2023). *Pesquisa nacional por amostra de domicílios contínua (Pnad Contínua)*. <https://www.ibge.gov.br/estatisticas/sociais/habitacao/17270-pnad-continua.html>

Instituto Brasileiro de Geografia e Estatística. (2024a). *Censo Demográfico 2022: agregados por setores censitários preliminares: população e domicílios: resultados do universo (Nota metodológica n. 02)*. <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2102071>

Instituto Brasileiro de Geografia e Estatística. (2024b). *Censo Demográfico 2022: Cadastro Nacional de Endereços para Fins Estatísticos - CNEFE (Nota metodológica n. 04)*. <https://biblioteca.ibge.gov.br/visualizacao/livros/liv102091.pdf>

Lumley, T. (2010). *Complex surveys: A guide to analysis using R*. John Wiley & Sons.

Opsomer, J. D., & Erциulescu, A. L. (2021). Replication variance estimation after sample-based calibration. *Survey Methodology*, 47(2), 265–277. <http://www.statcan.gc.ca/pub/12-001-x/2021002/article/00006-eng.htm>

Rosén, B. (2000). *A user's guide to Pareto π ps sampling*. Statistics Sweden.

Silva, P. L. N., Bianchini, Z. M., & Dias, A. J. R. (2020). *Amostragem: teoria e prática usando R*. <https://amostragemcomr.github.io/livro/>

União Internacional de Telecomunicações. (2020). *Manual for measuring ICT access and use by households and individuals, 2020 edition*. https://www.itu.int/en/ITU-D/Statistics/Documents/publications/manual/ITUManualHouseholds2020_E.pdf

Ye, X., Konduri, K., Pendyala, R., Sana, B., & Waddell, P. (2009). *Methodology to match distributions of both household and person attributes in generation of synthetic populations* [Apresentação]. 88th Annual Meeting of the Transportation Research Board, Seattle, WA, Estados Unidos. https://www.researchgate.net/publication/228963837_Methodology_to_match_distributions_of_both_household_and_person_attributes_in_generation_of_synthetic_populations