Nota metodológica

Pesquisa VOZES do serviço público

ENAP

MINISTÉRIO DA GESTÃO E DA INOVAÇÃO EM SERVIÇOS PÚBLICOS

EXPEDIENTE

Governo Federal

Ministério da Gestão e da Inovação em Serviços Públicos – MGI
Esther Dweck – Ministra de Estado da Gestão e da Inovação em Serviços Públicos

Escola Nacional de Administração Pública – Enap
Betânia Peixoto Lemos – Presidenta

Secretaria de Gestão de Pessoas – SGP
José Celso Cardoso Jr. – Secretário de Gestão de Pessoas

Realização:
Diretoria de Altos Estudos – ENAP
Diretoria de Governança e Inteligência de Dados – DIGID/SGP/MGI

Apoio:
Secretaria de Gestão e Inovação – SEGES

Elaboração:

Coordenação da Pesquisa
Alexandre de Ávila Gomide
Mirian Lucia Bittencourt Guimarães
Pedro Masson Sesconetto Souza

Equipe de Estatística e de Análise de Dados
Camila Saeko Kobayashi de Pinho
César Augusto Fernandes Galvão
Iara Azevedo Vitelli Viana
Juliana Legentil Ferreira Faria
Leandro Del Grande Cláudio

Projeto gráfico
Samyra Lima

APRESENTAÇÃO

Este relatório apresenta os resultados da 1ª edição da **Pesquisa Vozes do Serviço Público Federal**, uma iniciativa do Ministério da Gestão e da Inovação em Serviços Públicos (MGI) em parceria com a Escola Nacional de Administração Pública (Enap). Lançada em 2024, a pesquisa busca coletar a percepção dos servidores federais sobre políticas e práticas no trabalho, possibilitando a geração de evidências e subsídios estratégicos para o aprimoramento da gestão de pessoas.

Com caráter inovador e comparável internacionalmente, a pesquisa Vozes reúne dados inéditos sobre dimensões como engajamento, satisfação, liderança, segurança psicológica, autonomia, motivação e intenção de permanência ou rotatividade. Diferentemente dos registros administrativos tradicionais, que já oferecem informações demográficas e funcionais, a Vozes preenche uma lacuna essencial ao coletar dados subjetivos e comportamentais que influenciam o desempenho individual e organizacional.

A metodologia foi construída com base em padrões de excelência internacional, assegurando robustez metodológica e confiabilidade estatística. A curadoria de perguntas uniu itens comparáveis internacionalmente, tendo como base principal a *Global Survey of Public Servants* e itens de escalas com evidência de validade no Brasil. A coleta foi realizada entre novembro de 2024 e janeiro de 2025 por meio do aplicativo SouGov, de forma censitária a todos os servidores federais ativos do Poder Executivo, incluindo temporários, comissionados e sem vínculo, alcançando 49.777 respondentes válidos.

Este documento está estruturado em seções que apresentam: i) o sumário executivo com os principais achados; ii) as questões metodológicas; iii) os resultados das escalas aplicadas; e iv) comparações internacionais que situam a experiência brasileira em um contexto global.

A Pesquisa Vozes inaugura uma nova instância de escuta contínua dos servidores federais, fortalecendo a transparência, a participação e a gestão baseada em evidências. Trata-se de um instrumento estratégico para a construção de ambientes de trabalho mais justos, engajadores e produtivos, que contribuam para a excelência do serviço público oferecido à sociedade brasileira.

Boa leitura!

LISTA DE FIGURAS

Figura 1 - Distribuição percentual de servidores públicos federais por geração - período de referência: outubro de 2024. 12

Figura 2 - Distribuição percentual de servidores públicos federais por escolaridade - período de referência: outubro de 2024. 13

Figura 3 - Comparativo da distribuição percentual de servidores públicos federais por escolaridade - população e respondentes, período de referência: outubro de 2024. 14

Figura 4 - Comparativo da distribuição percentual de servidores públicos federais por geração - população e respondentes. 14

Figura 5 - Comparação de desempenho dos desenhos amostrais testados utilizando o coeficiente de variação das médias amostrais. 24

LISTA DE TABELAS

Tabela 1 - Distribuição de servidores públicos federais por sexo e geração. 12

Tabela 2 - Número de questão, média, erro-padrão da média e CV percentual para cada questão da pesquisa Vozes. 24

SUMÁRIO

Introdução 8

A pesquisa Vozes do Serviço Público 9

Características da População de Pesquisa 11

Metodologia de expansão dos resultados 15

Método de expansão por pesos ajustados 15

Cálculo de pesos base 16

Ajuste para não-resposta 17

Ajuste por *raking* 18

Aplicação na Pesquisa Vozes 18

Método *propensity score matching* 19

Aplicação na Pesquisa Vozes 20

Método escolhido para a análise das respostas dos questionários 21

Considerações Finais 22

Apêndice 23

Referências Bibliográficas 28

Introdução

A pesquisa “Vozes do Serviço Público Federal: Evidências para a Gestão Pública” foi conduzida com o objetivo de captar percepções, demandas e pontos de vista dos(as) servidores(as) públicos(as) federais em atividade – incluindo servidores efetivos, temporários, ocupantes de cargos comissionados e colaboradores sem vínculo permanente – sobre uma diversidade de temas como: engajamento, motivação, segurança psicológica, liderança, satisfação salarial. A partir dos dados coletados e das subsequentes análises das respostas, a pesquisa visa proporcionar ferramentas para o aprimoramento das políticas de gestão de pessoas no âmbito da administração pública federal.

A iniciativa é fruto de parceria entre o Ministério da Gestão e da Inovação em Serviços Públicos (MGI) e a Escola Nacional de Administração Pública (ENAP), tendo alcançado cerca de 55 mil respondentes, no período de 4 de novembro de 2024 e 3 de janeiro de 2025.

Esta Nota Metodológica tem como objetivo apresentar as diretrizes metodológicas utilizadas para a análise dos resultados da pesquisa “Vozes do Serviço Público Federal: Evidências para a Gestão Pública”, coordenada pela Escola Nacional de Administração Pública (ENAP) e pelo Ministério da Gestão e Inovação em Serviços Públicos (MGI). Para obtenção dos resultados finais foram realizadas análises estatísticas para a expansão dos resultados obtidos para a população, sendo esta a metodologia explicitada nesta Nota Metodológica.

A pesquisa Vozes do Serviço Público

Na escolha do método de pesquisa, a Diretoria de Altos Estudos da ENAP, e a Diretoria de Governança e Inteligência de Dados da SGP/MGI, optaram pelo modelo tipo censitário. Isso é, todos os servidores que constam no SIAPE receberam a pesquisa para responder em seu aplicativo Sou.Gov e tiveram possibilidade de respondê-la. Tal escolha se deu em uma visão estratégica de médio e longo prazo, visando fortalecer a pesquisa Vozes como um canal de comunicação direta dos servidores públicos com o órgão Central de Gestão de Pessoas. Além de gerar estatísticas confiáveis sobre a opinião dos servidores em diversos assuntos, a Pesquisa Vozes busca instituir uma cultura de escuta dos mesmos para a tomada de decisão sobre Gestão de Pessoas no Governo Federal. Dessa forma, o método censitário deu a mesma oportunidade para todos os servidores opinarem e fazerem suas vozes serem ouvidas.

O método de coleta utilizado foi o CAWSI, ou, *Computer-Assisted Web Self-Interviewing*, no qual a entrevista foi respondida pela própria pessoa em formulário online disponibilizado via internet, por meio da plataforma Sou.Gov, a qual todos os elementos da população de interesse possuem acesso. As comunicações para divulgação e incentivo ao preenchimento foram feitas pelo e-mail de contato indicado pelo próprio servidor no portal Sou.Gov.

Tal método de coleta, assim como o conjunto de procedimentos metodológicos adotados para permitir que as informações coletadas pudessem ser expandidas para a população de pesquisa, são amplamente utilizados por instituições no Brasil e no exterior. Pesquisas como a "OPM FEVS (*U.S. Office of Personnel Management Federal Employee Viewpoint Survey*)", do Governo dos Estados Unidos, a "PNAES (Pesquisa Nacional de Avaliação dos Egressos)", do Senac, e, a pesquisa "Ética e Corrupção no Serviço Público Federal: a perspectiva dos servidores", realizada pelo Banco Mundial, são exemplos de levantamentos frequentes e com resultados validados e reconhecidos pela comunidade especializada.

A OPM FEVS é uma pesquisa anual de clima organizacional que coleta opiniões dos funcionários públicos sobre políticas, práticas e procedimentos em seus respectivos órgãos de lotação (OPM, 2022). O questionário *online* é enviado por *e-mail* a todos os funcionários elegíveis e, para a expansão dos resultados para a população, é utilizada uma técnica de pesos amostrais simulados. Os resultados ajudam a identificar áreas fortes e oportunidades de melhoria dentro dos órgãos federais, fornecendo insights valiosos para iniciativas de mudança organizacional e desenvolvimento. Já a PNAES (SENAC, 2024) é uma pesquisa realizada anualmente pelo Serviço Nacional de Aprendizagem Comercial (Senac), cujo objetivo é avaliar a situação ocupacional de todos alunos que concluíram

cursos da instituição. Para isso, é enviado um *e-mail* e/ou SMS no celular, conforme tipo de contato no registro do egresso, com instruções para a pesquisa a ser realizada via *web*. Para a expansão dos resultados para a população, o método utilizado é a do cálculo de pesos amostrais, simulados por modelos de propensão de resposta, para a determinação dos fatores de ajuste iniciais para expansão dos resultados, e, por fim, a técnica conhecida como rake para a calibração desses fatores de ajuste.

Na pesquisa realizada intitulada "Ética e Corrupção no Serviço Público Federal: a perspectiva dos servidores" (Banco Mundial, 2021) o método de coleta das informações foi tipo censitário, a partir do envio de questionário online para todos os servidores públicos do Governo Federal brasileiro. Para a expansão dos resultados para a população foi utilizado o método de estimação de pesos estratificados, uma das práticas utilizadas pelo *US Census Bureau*.

O questionário aplicado para a pesquisa "Vozes" seguiu, padrões internacionais de qualidade, o que permite uma comparabilidade com indicadores de outros países a partir do *framework* teórico do *Global Survey of Public Servants* (GSPS), além de outras escalas validadas em pesquisas acadêmicas que complementaram as perguntas originais do GSPS. No âmbito desta Nota Técnica serão explicitados apenas os aspectos relacionados ao processo de tratamento dos dados e à metodologia de expansão dos resultados obtidos pela pesquisa para a população.

As respostas foram coletadas, em aplicação própria desenvolvida pelo SERPRO, que foi disponibilizada na área restrita do Sou.Gov de cada elemento da população de pesquisa. Finalizado o período para responder ao questionário, as informações foram compiladas em formato de base de dados e disponibilizada para esta Coordenação-Geral de Ciência de Dados pela Diretoria de Governança e Inteligência de Dados da SGP/MGI.

Características da População de Pesquisa

Neste estudo foram considerados todos os servidores públicos federais incluindo temporários, comissionados e sem vínculo que estavam em atividade no poder Executivo Federal, resultando em 537.555 pessoas elegíveis no período de referência de outubro de 2024 para responderem o formulário.

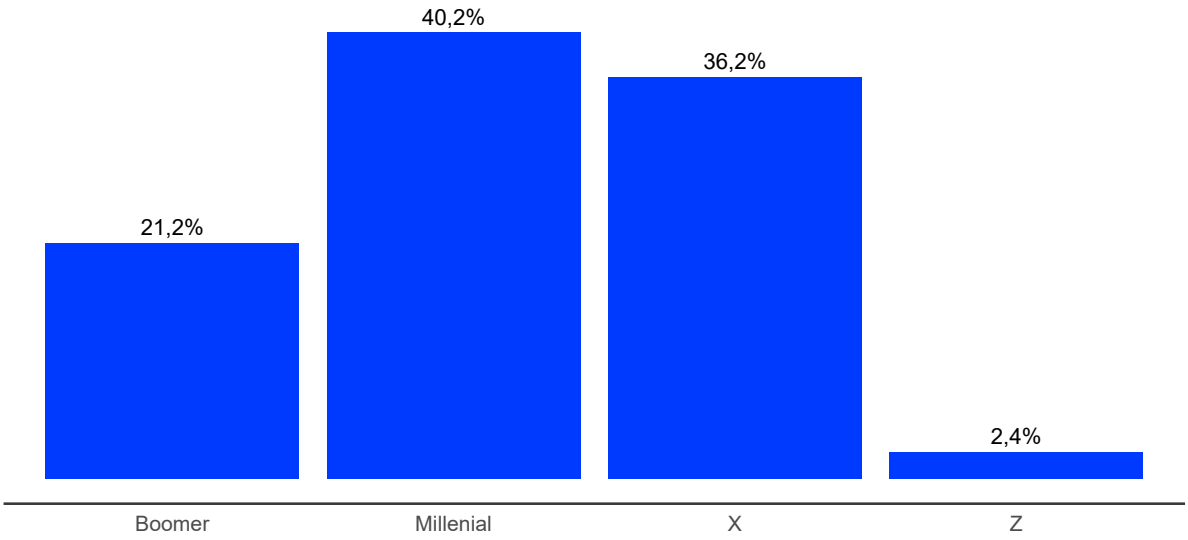
Servidores aposentados, pensionistas, empregados públicos, servidores do Banco Central, e servidores do Ministério das Relações Exteriores em exercício fora do país, não foram considerados elegíveis para esta pesquisa. Por meio da listagem constante no SIAPE (Sistema Integrado de Administração de Recursos Humanos) foram levantadas as pessoas elegíveis incluindo servidores temporários, comissionados e sem vínculo que estavam em atividade no poder Executivo Federal no mês de referência de outubro de 2024.

Com o objetivo de adicionar significado à idade dos respondentes, estes foram divididos em gerações da seguinte forma:

- Z – até 27 anos;
- *Millenial* – de 28 a 44 anos;
- X de 45 a 60 anos; e
- *Boomer* – acima de 60 anos.

A Figura 1 a seguir mostra a distribuição percentual dos servidores segundo recorte geracional. Observa-se que a maior parcela dos servidores é de nascidos entre 1981 e 1996, correspondendo a 40,2% da população de pesquisa. Já a segunda maior parcela é de pessoas entre 45 e 60 anos.

Figura 1 - Distribuição percentual de servidores públicos federais por geração - período de referência: outubro de 2024



Fonte: Elaboração própria.

Já a distribuição de servidores por sexo segundo faixa geracional mostra que a maioria (54%) dos servidores é do sexo masculino. Além disso, mulheres da geração Z, nascidas entre 1997 e 2010, são as que mais se aproximam da metade da quantidade de pessoas entre as faixas, equivalente a 48,3% do total.

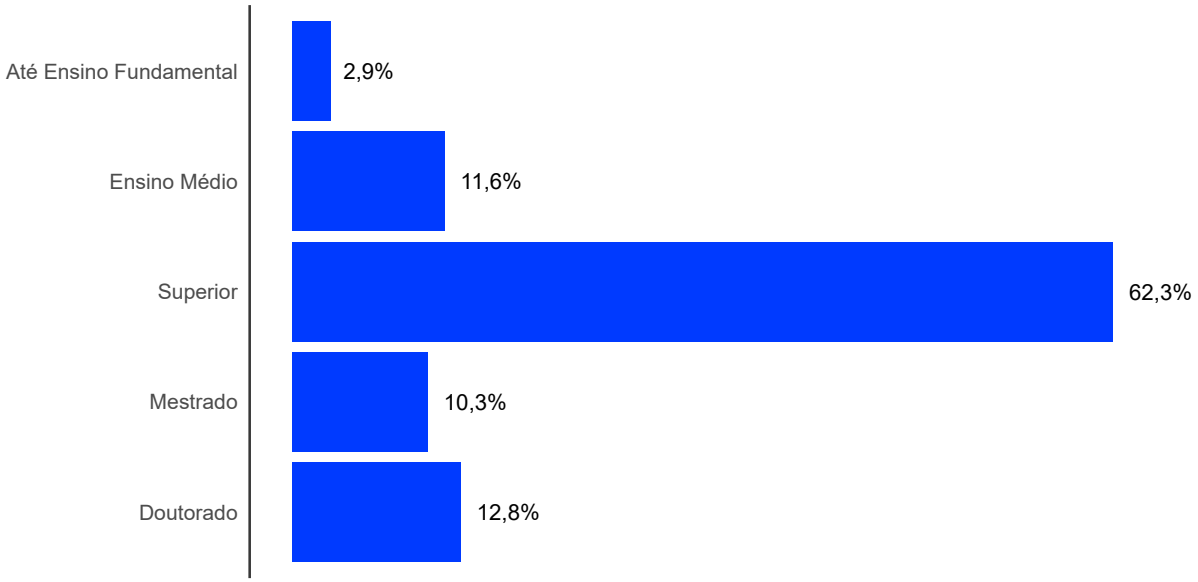
Tabela 1 - Distribuição de servidores públicos federais por sexo e geração

Sexo	Z		*Millenial*		X		*Boomer*		Total	
	N	%	N	%	N	%	N	%	N	%
Feminino	6.315	48.3	103.648	47.9	89.739	46.2	47.639	41.8	247.341	46.0
Masculino	6.753	51.7	112.618	52.1	104.628	53.8	66.215	58.2	290.214	54.0
Total Geral	13.068	100	216.266	100	194.367	100	113.854	100	537.555	100

Fonte: Elaboração própria.

Ao observar a distribuição dos servidores por escolaridade na Figura 2, verifica-se que 85,4% possuem algum grau educacional do Ensino Superior, sendo 62,3% destes com o nível superior completo. Servidores com até o nível de Ensino Médio completo correspondem a menos de 15% do total.

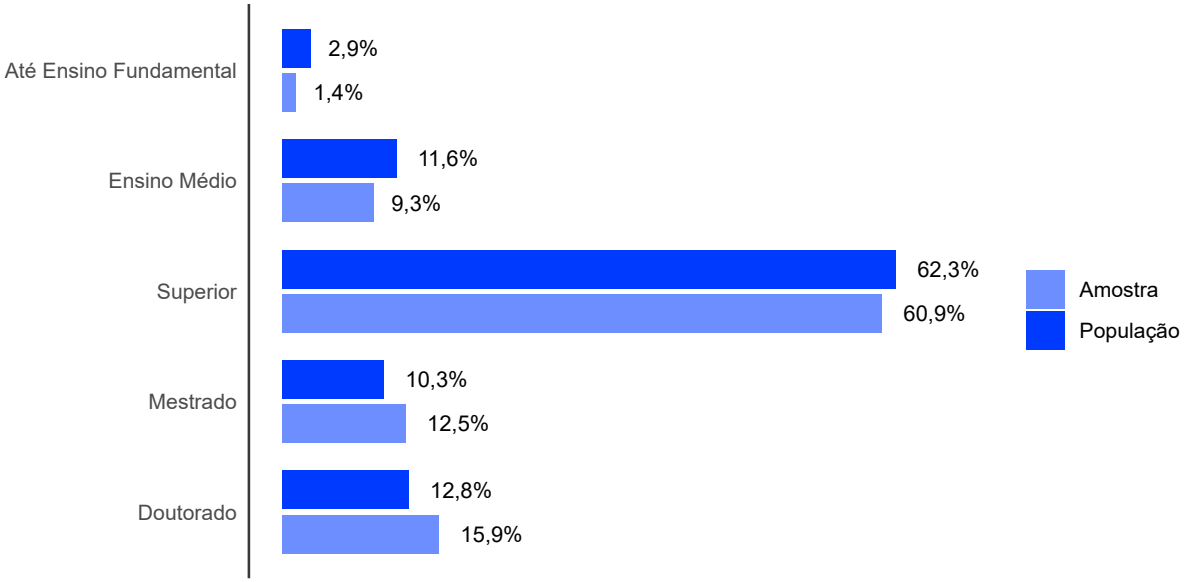
Figura 2 - Distribuição percentual de servidores públicos por escolaridade - período de referência: outubro de 2024



Fonte: Elaboração própria.

Um comparativo entre as proporções observadas em variáveis auxiliares da população e da amostra podem ser um indicativo de que há representatividade na amostra coletada. A Figura 3 mostra a distribuição entre as proporções da amostra e da população por escolaridade e verifica-se que estão bem próximas. Além disso, observa-se que pessoas com nível de escolaridade de mestrado e doutorado são as que mais tenderam a responder a pesquisa.

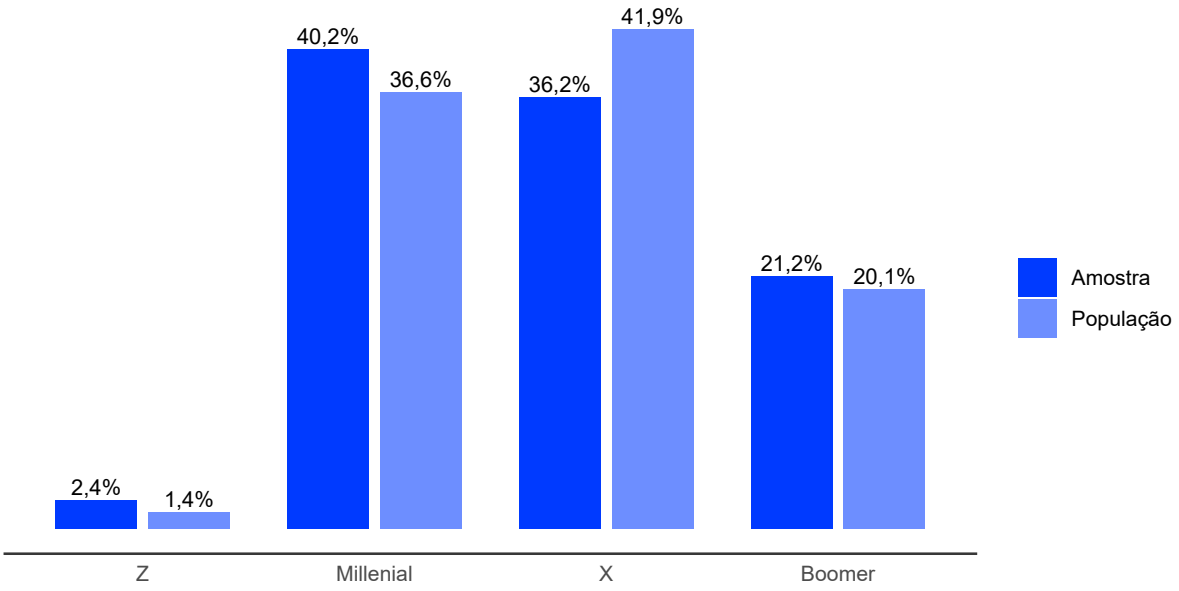
Figura 3 - Comparativo da distribuição percentual de servidores públicos federais por escolaridade - população e respondentes, período de referência: outubro de 2024



Fonte: Elaboração própria.

A distribuição geracional entre população e amostra é similar, conforme se pode ver na Figura 4.

Figura 4 - Comparativo da distribuição percentual de servidores públicos federais por geração - população e respondentes



Fonte: Elaboração própria.

Metodologia de expansão dos resultados

Realizada no período compreendido entre 4 de novembro de 2024 e 3 de janeiro de 2025, a pesquisa Vozes obteve ao todo 55.030 mil formulários preenchidos. Destes, 49.775 foram formulários completamente preenchidos, sendo o complementar, 5.252 o correspondente às recusas em responder o questionário.

Compreende-se como taxa de resposta a razão (em termos percentuais) entre o total de respostas obtidas e a soma de respondentes e não respondentes, o que, considerando-se os valores indicados no parágrafo anterior, corresponde a uma taxa de aproximadamente 9%.

Dada a natureza voluntária da participação e a variação nas taxas de resposta entre órgãos e subgrupos demográficos, a aplicação de pseudopesos amostrais é essencial para corrigir distorções e proporcionar estimativas mais precisas sobre a população a partir da informação dos respondentes (Kott; Chang, 2010). Por isso, optou-se pela aplicação de dois métodos utilizados em pesquisas de grande relevância: o de pesos ajustados para não-resposta e o de pesos simulados via modelos de propensão de resposta. Após a obtenção dos pseudopesos pelos métodos, realizou-se a escolha da técnica que obteve melhor adequabilidade em termos de redução de viés.

Para a composição dos desenhos amostrais experimentados e cálculo dos pesos para recomposição da população (Kalton; Flores-Cervantes, 2003), foram usados pacotes disponíveis para a linguagem R (Lumley, 2023), que já contam com todos os calculos e funções necessários para execução da pesquisa.

Método de expansão por pesos ajustados

A *Federal Employee Viewpoint Survey* (FEVS), conduzida pelo Departamento de Gestão de Pessoal dos Estados Unidos (*Office of Personnel Management* - OPM), é uma pesquisa tipo censitária aplicada a servidores federais com o objetivo de aferir percepções sobre condições de trabalho, liderança, engajamento e outras dimensões relacionadas à gestão organizacional.

Em sua edição de 2022, a FEVS empregou uma metodologia de ponderação estatística em três etapas, visando corrigir possíveis vieses de não resposta e garantir que os resultados sejam representativos da população federal como um todo e por órgão. A metodologia de ponderação adotada na OPM FEVS 2022 está alinhada às melhores práticas internacionais

em pesquisas por levantamento, combinando correções de não resposta com ajustes marginais de composição da amostra, visando a produção de estatísticas confiáveis e representativas para a tomada de decisão em âmbito governamental.

A seguir são detalhadas as etapas de ajustamento de pesos.

Cálculo de pesos base

Neste passo é realizado o cálculo dos pesos base, definidos como o inverso da probabilidade de seleção de cada funcionário amostrado. Dentro de cada estrutura de órgão principal, os funcionários foram agrupados (estratificados) pela menor unidade de trabalho, definida pelo pesquisador e pelo status executivo. Assim, havia H grupos distintos (ou estratos) consistindo de $\mathrm{N_h}$ empregados em cada subgrupo, de modo que

$$N = \sum_{h=1}^{H} N_h$$

onde *N* é a contagem total de pessoas no recorte da estrutura do órgão – ou seja, o número total de funcionários listados no menor nível escolhido por órgão.

Em cada subgrupo, uma amostra aleatória foi selecionada sem reposição. A probabilidade de seleção variou por subgrupo para garantir uma representação adequada dos membros deste na amostra. Dado esse desenho, o peso base (w) do i-ésimo funcionário amostrado no subgrupo h foi calculado como

$$w_{hi} = \frac{N_h}{n_h}$$

onde n_h é o tamanho da amostra para o h-ésimo grupo, e N_h é o total de pessoas no menor nível do órgão para o h-ésimo subgrupo.

Dessa forma, para cada funcionário classificado no subgrupo h, o peso base correspondente é a razão entre o número total de funcionários no subgrupo e o tamanho da amostra do subgrupo (equivalente ao inverso da probabilidade de seleção). O peso base é atribuído a cada unidade amostral (funcionário) no arquivo de dados, ou seja, não apenas aos respondentes da pesquisa. Assim, todos recebem um peso base.

Ajuste para não-resposta

Na etapa seguinte, os pesos base são ajustados para compensar a não resposta em cada um dos subgrupos definidos por características da organização (ex.: órgão e subórgão). Esse ajuste redistribui os pesos base dos não respondentes entre os respondentes com perfis semelhantes, assegurando que os resultados ponderados reflitam melhor a estrutura da população de pesquisa, inclusive os indivíduos que não responderam o questionário (Bethlehem; Cobben; Schouten, 2011; Little; Rubin, 2002).

Para cada agência ou conjuntos de subagências do governo estadunidense foram calculados ajustes de não-resposta individuais. E, em cada uma delas, células de ponderação foram construídas para agrupar respondentes e não respondentes com características similares. As variáveis elencadas pelos pesquisadores do FEVS OPMS 2022 para formar essas células incluíam: identificador da subagência, função, sexo, status minoritário, faixa etária, tempo de serviço, jornada (integral ou reduzida) e localização (sede ou escritório de campo).

Quando houve subgrupos grandes os mesmos foram divididos em grupos menores para aumentar a variação. Um algoritmo de busca categórica, o CHAID (*Chi-squared Automatic Interaction Detector*), foi utilizado para dividir os dados em células menores com o objetivo de haver diferença nas taxas de resposta entre as células. Nos casos em que células tiveram taxas de resposta similares, combinaram-se algumas para garantir um tamanho mínimo de 30 respondentes.

Após a formação das células de ponderação, os pesquisadores calcularam dois fatores de ajuste para não resposta para cumprir a finalidade de distribuição dos pesos base entre os não respondentes e as outras unidades de elegibilidade conhecida, no caso do primeiro fator (f_1), e, um segundo fator (f_2) que distribuiu os pesos ajustados dos não respondentes entre os que responderam ao questionário.

Calculados os dois fatores de ajuste, procedeu-se ao cômputo dos valores dos pesos finais para o ajuste de não-resposta (w_{nr}), que resultaram da ponderação entre os pesos e fatores obtidos nas etapas anteriores para cada respondente, como pode ser observado a seguir:

$$w_{nr} = w \times f_1 \times f_2$$

Ajuste por *raking*

A etapa final consistiu na aplicação da técnica de raking, ou "ajuste por proporções iterativas", ou seja, os pesos dos respondentes foram calibrados para que a soma ponderada das variáveis de controle coincidisse com os totais populacionais (Deville; Särndal; Sautory, 1993) em duas dimensões principais: a primeira contemplando o cruzamento entre agência, sexo e status de minoria, e, a segunda, um identificador de estrato truncado reduzido a quatro caracteres.

Durante o processo, os pesos foram ajustados alternadamente (processo iterativo) para cada dimensão até que a soma dos pesos se aproximasse dos totais populacionais estabelecidos, dentro de um nível de precisão aceitável. O resultado final do raking foi a obtenção de um peso ajustado para cada respondente, representando o número de pessoas na população que aquele indivíduo representa estatisticamente. Esses pesos finais foram então incorporados ao banco de dados e utilizados para cálculo de estimativas ponderadas com maior precisão e representatividade.

Aplicação na Pesquisa Vozes

Para a aplicação da metodologia da OPM FEVS 2022 nesta pesquisa, buscou-se ao máximo seguir os mesmos passos para garantir fidedignidade ao método. Alguns aspectos não puderam ser adotados, como a separação entre sede e filiais para composição de estratos amostrais e a divisão entre órgãos de acordo com a quantidade de servidores. No entanto, as mesmas técnicas de clusterização e pós-estratificação foram adotadas.

O desenho amostral construído a posteriori para expansão dos resultados consiste em um desenho de dois estágios: um primeiro nível de estratos e um segundo nível de conglomerados.

Para a construção das células de ponderação, assim como feito no OPM FEVS 2022, foi utilizado o algoritmo CHAID. Essas células foram construídas para cada órgão e constituem um conglomerado, cada um com um número identificador. Foram testadas diversas combinações de variáveis utilizadas para a execução do algoritmo, as quais são listadas a seguir. O desempenho de cada desenho amostral, assim como da utilização do *propensity score* para construção de pesos amostrais, foi comparado graficamente utilizando a densidade dos coeficientes de variação das médias amostrais, conforme a figura disposta no Apêndice, utilizando as seguintes combinações de variáveis sociodemográficas:

- Faixa salarial e geração;
- Sexo e raça/cor;
- Sexo e geração;
- Faixa salarial, sexo geração e raça cor.

Conforme o relatório OPM FEVS 2022, é nesta etapa em que se aplicam os fatores de correção para não resposta. Dado que o tamanho da amostra é exatamente o tamanho da população e que não houve casos em que a elegibilidade para responder ao questionário fosse conhecida ou desconhecida (ou seja, esse fator não foi considerado no momento de coleta dos dados), tem-se

$$w_i = 1, \forall i = 1, \dots, H, \text{ e } f_1 = 1$$

Desse modo, apenas o fator de correção f_2 foi aplicado aos dados, definido como

$$f_c^2 = \frac{\sum_{i \in ER_c} w_i + \sum_{j \in ENR_c} w_j}{\sum_{i \in ER_c} w_i}$$

onde f^2_c é definido para cada célula de ponderação c e cada peso w_i ou w_j pertence aos respondentes (ER ou *Elegible Repondents*) ou aos não respondentes (ENR ou *Elegible Non-Respondents*) daquela célula.

Método *propensity score matching*

A modelagem de propensão de resposta é uma abordagem consolidada para mitigar o efeito de estimativas enviesadas em pesquisas tipo censitárias que enfrentam o problema da não-resposta. É um método inspirado em estudos de inferência causal em estudos observacionais (Rosenbaum; Rubin, 1983) e tem sido utilizado pela equipe técnica do Senac para a expansão de resultados da pesquisa PNAES para a população de egressos em análise.

O ajuste por este tipo de modelagem resulta no *propensity score*, que é a probabilidade condicional de uma unidade observacional (pessoa, empresa, etc.) responder à pesquisa, dado um conjunto de variáveis observadas sobre a população de pesquisa após a aplicação dos questionários. A lógica é que, ao estimar essa probabilidade, pode-se corrigir o viés

de não resposta, balanceando as características dos respondentes e não respondentes, de forma semelhante ao que seria esperado em uma amostra aleatória. Desse modo, cada unidade recebe um score (nota) que indica a chance de ela ter respondido, com base em suas características observadas.

No cenário de amostragem, o método baseia-se na construção de um modelo paramétrico (tipicamente uma regressão logística) para a variável indicadora de resposta ao formulário, utilizando covariáveis disponíveis, como as de perfil demográfico e de salário e/ou lotação, por exemplo, para todas as unidades da amostra. O resultado é um escore de propensão estimado

$$e(x_i) = P\ (\delta_i = 1/x_i)$$

que é então utilizado para ajustar os pesos das unidades respondentes, aplicando-se o inverso da propensão como fator de ponderação. Ajustando os pesos das respostas observadas pelo inverso do propensity score, tenta-se reconstruir o que seria a amostra total sem viés (Kim; Riddles, 2012).

O método tem por objetivo o equilíbrio na distribuição das covariáveis entre respondentes e não respondentes, e assume que, condicionalmente ao escore de propensão, os dados estão ausentes. Essa propriedade de "balanceamento" é central na teoria de Rosenbaum; Rubin (1983) e fundamenta a validade da inferência sob o uso do escore como um redutor de viés.

Após a estimação dos pesos pelo modelo de propensão de *score* é aplicada, também neste caso, a técnica de *raking*, na qual os pesos dos respondentes são calibrados para que a soma ponderada das variáveis de controle coincida com os totais populacionais, obtendo-se assim os pesos finais para a expansão dos resultados para a população.

Aplicação na Pesquisa Vozes

A aplicação do *propensity score* para composição de pesos amostrais tem execução trivial. No entanto, Gelman; Hill (2007), assim como Demidenko (2013) apontam para a necessidade de representar adequadamente as estruturas de variância inerentes aos dados para construção de modelos estatísticos. Ou seja, é necessário considerar a estrutura de grupos existente nos dados coletados para que se possa construir um modelo adequado.

No entanto, há impedimentos algorítmicos e de capacidade de processamento para se executar o plano ótimo de análise, conforme o parágrafo anterior. Dessa forma, para o ajuste de regressões logísticas e assim compor o *propensity score*, optou-se por uma regressão logística por órgão – o impacto dessa opção também é descrito por Gelman; Hill (2007). Entende-se que, mediante a falta de opções ótimas, o método escolhido é melhor que a desconsideração absoluta da estrutura de grupos presente nos dados.

Finalmente, o procedimento pode ser resumido aos seguintes passos:

1. Para cada orgão, ajustar um modelo logístico considerando como variáveis explica- tivas sexo, raça/cor, geração, faixa salarial. O inverso das respostas estimadas pelo modelo compõe o peso amostral base;

2. Um desenho amostral é construído utilizando órgão como estrato;

3. Aplica-se raking para a distribuição final dos pesos amostrais.

Método escolhido para a análise das respostas dos questionários

Para a seleção final do método de expansão dos resultados para a população a ser utilizado para as análises posteriores, foram considerados os seguintes aspectos: a avaliação das densidades dos coeficientes de variação em casa questão nos dois métodos (figura no Apêndice), a possibilidade de extrapolação dos resultados para a maior quantidade de órgãos possíveis e a qualidade e variedade de literatura específica disponível.

Neste caso, optou-se pelos resultados obtidos por meio do método de modelagem de propensão de respostas, uma vez que, de acordo com o gráfico no Apêndice, foi o que mostrou uma menor dispersão das densidades observadas para os valores de coeficiente de variação (variabilidade da amostra relativa à sua média, definido como o desvio-padrão amostral dividido pela média) para as questões do formulário, além de possibilitar que todos os órgãos que tiveram respondentes tenham seus resultados expandidos para a totalidade dos servidores ativos no período. A vasta e diversa literatura disponível sobre este método, em comparação ao outro relatado nesta nota técnica, permite também uma maior liberdade de análise e melhor avaliação da qualidade dos resultados obtidos. Além disso, o método não gera rótulos arbitrários para grupos, o que é observado no método utilizado pela pesquisa OPMS FEVS 2022 para a criação das células de ponderação e que pode gerar complicações em um estágio inferencial da pesquisa.

Considerações Finais

Como um esforço inicial e inédito na história da Administração Pública Federal, a Pesquisa Vozes logrou sucesso ao permitir que os gestores públicos ouçam as Vozes de mais de 55 mil servidores públicos federais. A partir dela é possível pela primeira vez traçar decisões estratégicas para melhorar a gestão de pessoas no serviço público federal como um todo.

A presente Nota Metodológica apresenta as técnicas mais utilizadas para expansão de uma amostra coletada de forma similar ao Vozes, sua aplicação na amostra coletada na pesquisa Vozes, e a escolha da melhor opção para a expansão dos resultados para a população.

Realizar pesquisas, seja ela de natureza quantitativa ou qualitativa, exigem da equipe de pesquisadores escolhas metodológicas. Em que pese a larga utilização de pesquisas amostrais que fazem seleção probabilística estratificada, foi imperioso no caso da "Pesquisa Vozes do Serviço Público: Evidências para Gestão de Pública", a utilização de um método de pesquisa capaz de permitir a todo servidor público federal participar da iniciativa.

O objetivo maior da pesquisa Vozes é criar uma cultura de escuta ativa dos servidores e de tomada de decisão pelos gestores informada por evidência. Assim constitui-se um binômio a cultura e qualidade das evidências coletadas.

A escolha do método tipo censitário foi respaldada por outras pesquisas, nacionais e internacionais, que se valem da mesma metodologia de coleta de dados, e também por diversas metodologias para expansão dos resultados para a população, com vasta literatura relacionada.

Os resultados que serão apresentados em relatórios posteriores estarão respaldados pela melhor técnica estatística e pelo rigor científico que tal iniciativa requer.

Apêndice

Como forma de avaliar a qualidade de cada metodologia para composição dos pesos amostrais, optou-se por calcular os coeficientes de variação para cada metodologia. Este coeficiente é dado por

$$CV = \frac{SE}{\bar{x}}$$

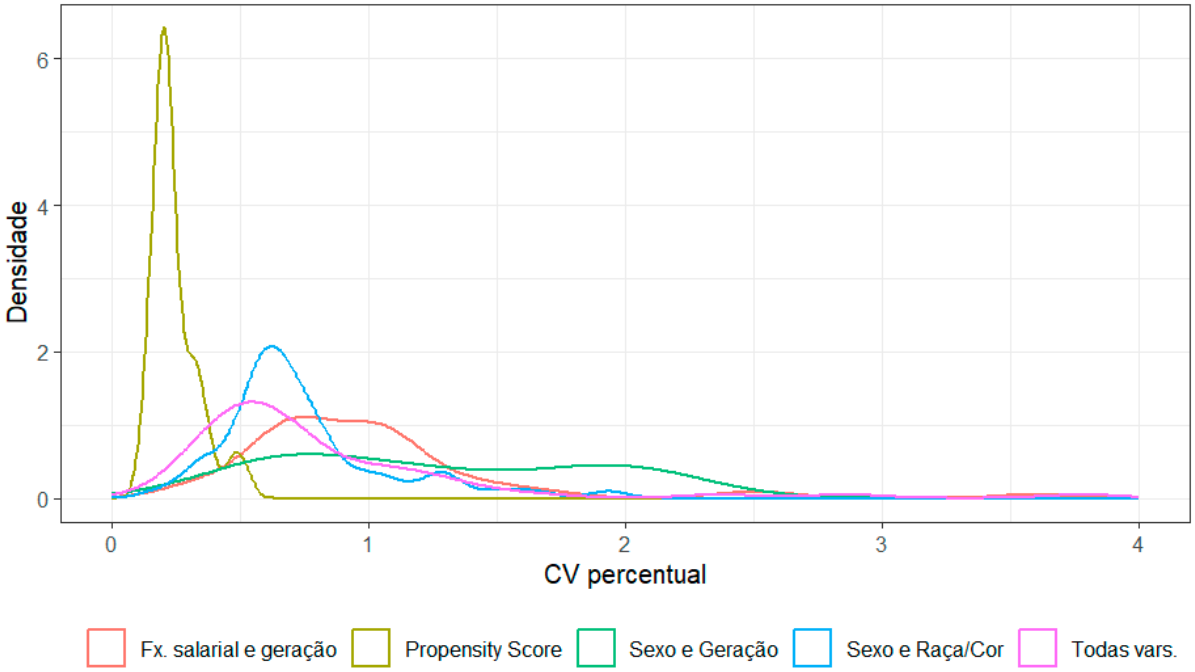
onde SE é o erro-padrão do estimador e $\bar{x}$ a média amostral e o CV é calculado para cada questão da pesquisa. Seu valor indica quanto uma determinada estimativa está variando, tomando como referência o valor da sua média. Por exemplo, se um CV tem valor 0,01, significa que o desvio-padrão corresponde a 1% da média e valores de estimativas dessa média devem variar, em sua maior parte, em mais ou menos 1 ponto percentual em torno da média. Nesse sentido, entende-se que quanto menor o CV, melhor.

A Figura 5 a seguir exibe a distribuição empírica do coeficiente de variação percentual obtido na estimação das médias de todos os itens do questionário considerados para a aná- lise. Cada cor do gráfico representa um desenho amostral ou metodologia para composição dos pesos amostrais.

A região abaixo de cada curva sempre equivale a 1, de modo que a sua forma evidenciará a sua dispersão e concentração em pontos específicos. Pesos amostrais obtidos via propensity score matching, a curva de cor mostarda e metodologia escolhida para expansão da amostra, possui a maior concentração em torno de um valor menor, com menos picos pequenos afastados de seu valor máximo.

Considerando ainda que o CV está dado em percentual e que todas as metodologias apresentaram concentrações em baixos valores, esse pico distinto foi utilizado como um indicador de qualidade por fornecer maior uniformidade na variação da estimação das estatísticas de interesse.

Figura 5 - Comparação de desempenho dos desenhos amostrais testados utilizando o coeficiente de variação das médias amostrais



Fonte: Elaboração própria.

A tabela a seguir apresenta os dados utilizados para compor a Figura 5, especificamente para a curva referente ao CV percentual das estimativas usando *Propensity Score*.

Tabela 2 - Número de questão, média, erro-padrão da média e CV percentual para cada questão da pesquisa Vozes

Questão	Média	Erro-Padrão	CV(%)
2	4,25	0,008	0,18
3	4,34	0,008	0,18
4	4,21	0,01	0,24
5	3,14	0,011	0,35
6	2,07	0,011	0,53
7	3,62	0,008	0,23
8	3,46	0,007	0,19

Questão	Média	Erro-Padrão	CV(%)
9	3,93	0,006	0,16
13	3,91	0,007	0,18
14	3,82	0,008	0,21
15	2,83	0,009	0,32
16	4,45	0,006	0,14
17	3,28	0,008	0,26
18	4,03	0,007	0,18
19	4,02	0,008	0,19
20	3,8	0,008	0,22
21	4,28	0,007	0,17
22	4,78	0,005	0,11
23	4,09	0,008	0,19
24	4	0,008	0,2
28	3,97	0,008	0,2
29	3,67	0,008	0,21
30	3,55	0,008	0,22
31	3,65	0,008	0,21
32	3,68	0,008	0,22
33	3,74	0,008	0,21
34	4,2	0,007	0,17
35	4,19	0,007	0,17
36	3,52	0,008	0,22
37	4,16	0,007	0,16

Questão	Média	Erro-Padrão	CV(%)
38	3,84	0,007	0,19
39	4,21	0,006	0,15
40	4,27	0,006	0,15
41	4,31	0,006	0,14
42	1,35	0,004	0,28
43	4,04	0,008	0,19
44	3,86	0,008	0,22
45	3,99	0,01	0,25
46	3,85	0,008	0,21
47	3,98	0,01	0,24
48	3,82	0,008	0,22
49	3,76	0,009	0,23
50	3,91	0,008	0,21
51	3,85	0,009	0,22
52	3,95	0,01	0,25
53	1,93	0,009	0,45
54	3,47	0,009	0,27
55	1,74	0,008	0,48
56	3,05	0,009	0,28
57	2,13	0,011	0,49
58	3,96	0,008	0,21
59	2,87	0,01	0,34
60	2,34	0,009	0,38

Questão	Média	Erro-Padrão	CV(%)
61	3,2	0,01	0,3
62	2,76	0,009	0,34
63	2,7	0,009	0,34
64	3,05	0,009	0,31
65	2,91	0,009	0,32
66	2,47	0,01	0,39

Fonte: Elaboração própria.

Referências Bibliográficas

BANCO MUNDIAL. *Ética e corrupção no serviço público federal: a perspectiva dos servidores*. Brasília: Banco Mundial, 2021.

BETHLEHEM, J.; COBBEN, F.; SCHOUTEN, B. *Handbook of Nonresponse in Household Surveys*. Hoboken: John Wiley & Sons, 2011.

DEMIDENKO, E. *Mixed models: theory and applications with R*. Hoboken: John Wiley & Sons, 2013.

DEVILLE, J. C.; SÄRNDAL, C. E.; SAUTORY, O. Generalized raking procedures in survey sampling. *Journal of the American Statistical Association*, v. 88, n. 423, p. 1013–1020, 1993.

GELMAN, A.; HILL, J. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press, 2007.

KALTON, G.; FLORES-CERVANTES, I. Weighting methods. *Journal of Official Statistics*, v. 19, n. 2, p. 81–97, 2003.

KIM, J. K.; RIDDLES, M. K. Some theory for propensity-score-adjustment estimators in survey sampling. *Survey Methodology*, v. 38, n. 2, p. 157–165, dez. 2012.

KOTT, P. S.; CHANG, T. Using calibration weighting to adjust for nonignorable unit nonresponse. *Journal of the American Statistical Association*, v. 105, n. 491, p. 1265–1275, 2010.

LITTLE, R. J. A.; RUBIN, D. B. *Statistical analysis with missing data*. Hoboken: John Wiley & Sons, 2002.

LUMLEY, T. *Survey: analysis of complex survey samples*. R package version 4.2-1, 2023. Disponível em: http://r-survey.r-forge.r-project.org/survey/ . Acesso em: 31 mar. 2025.

OFFICE OF PERSONNEL MANAGEMENT (OPM). *2022 Federal Employee Viewpoint Survey: Technical Report*. Washington, D.C.: U.S. Office of Personnel Management, 2022.

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, v. 70, n. 1, p. 41–55, 1983.

SENAC (Brasil). Departamento Nacional. *Pesquisa Nacional de Avaliação dos Egressos do Senac: síntese PNAES 2022*. Rio de Janeiro: Serviço Nacional de Aprendizagem Comercial – Senac, Departamento Nacional, 2024.