



MINISTÉRIO DA EDUCAÇÃO
INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS
ANÍSIO TEIXEIRA

MICRODADOS ENEM 2018

LEIA-ME

Maio de 2019

DIRETORIA DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA

Coordenação Geral do Sistema Nacional de Avaliação da Educação Básica

Coordenação Geral de Instrumentos e Medidas

Coordenação Geral de Exames para Certificação

SUMÁRIO

1- APRESENTAÇÃO	4
2- O ENEM	4
3- MICRODADOS DO ENEM	6
4- ARQUIVOS DOS MICRODADOS	7
5- COMO CITAR OS MICRODADOS	9
6- CONSIDERAÇÕES FINAIS	9

1- APRESENTAÇÃO

O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), por intermédio da Diretoria de Avaliação da Educação Básica, em cumprimento da sua missão de desenvolver e disseminar informações sobre os exames e avaliações da educação básica, disponibiliza os **Microdados do Enem 2018**.

Devido à vigência da Lei Geral de Proteção de Dados (LGPD), incorporada ao ordenamento jurídico brasileiro por meio da Lei nº 13.709, de 14 de agosto de 2018, o INEP viu-se obrigado a realizar mudanças no modelo de microdados utilizado para o Exame Nacional do Ensino Médio (Enem). A LGPD promulgou o sistema protetivo dos dados pessoais no direito brasileiro, estabelecendo princípios norteadores da coleta, do compartilhamento e do tratamento de dados pessoais, além de um conjunto de obrigações aos controladores e responsáveis por essas atividades.

Os microdados se constituem no menor nível de desagregação de dados recolhidos por pesquisas, avaliações e exames realizados. No caso do ENEM, os dados estão por participante. Cabe lembrar que nunca constaram nos dados divulgados quaisquer variáveis que permitissem a identificação direta do participante, tais como nome, logradouro do endereço, e-mail, número do CPF, data de nascimento, dentre outras. Inclusive, o próprio número de inscrição real do participante é substituído por uma máscara gerada sequencialmente. Tratam-se, portanto, de dados anonimizados que, a priori, não são considerados dados pessoais e estão fora do escopo da legislação.

O problema é que o conceito previsto no art. 5º, III da LGPD estabelece que dado anonimizado é o “dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento”. Observa-se, portanto, que a lei não define objetivamente o que são esforços razoáveis. Por isso, o INEP optou pela cautela. Adotou-se um modelo simplificado de microdados, a ser utilizado para a edição 2018 do ENEM. O objetivo é eliminar da base pública variáveis que facilitem a identificação indevida do participante.

2- O ENEM

O Exame Nacional do Ensino Médio (Enem) tem como objetivo primordial aferir se aqueles que dele participam demonstram, ao final do ensino médio, individualmente, domínio dos princípios científicos e tecnológicos que presidem a produção moderna e se detêm conhecimento das formas contemporâneas de linguagem.

Os resultados do Enem deverão possibilitar:

- I - a constituição de parâmetros para a autoavaliação do participante, com vistas à continuidade de sua formação e a sua inserção no mercado de trabalho;
- II - a criação de referência nacional para o aperfeiçoamento dos currículos do ensino médio;
- III - a utilização do Exame como mecanismo único, alternativo ou complementar para acesso à educação superior, especialmente a ofertada pelas instituições federais de educação superior;
- IV - o acesso a programas governamentais de financiamento ou apoio ao estudante da educação

V - a sua utilização como instrumento de seleção para ingresso nos diferentes setores do mundo do trabalho; e

VI - o desenvolvimento de estudos e indicadores sobre a educação brasileira.

Desde sua primeira edição, em 1998, até 2008, o Enem era realizado anualmente, com a aplicação de uma única prova composta por 63 questões interdisciplinares. Durante esse período, algumas instituições de Ensino Superior passaram a utilizá-lo como instrumento de seleção para o ingresso de seus estudantes.

Em 2009, ocorreu a reformulação metodológica do Enem e de sua utilização como forma de seleção nos processos seletivos de muitas universidades federais. Reformularam-se as Matrizes de Referência do Exame, que passaram a ser estruturadas por competências em quatro áreas do conhecimento – Linguagens, Códigos e suas tecnologias, Matemática e suas tecnologias, Ciências Humanas e suas tecnologias e Ciências da Natureza e suas tecnologias –, apresentando, ainda, eixos cognitivos comuns a todas as áreas.

O Exame, com 180 (cento e oitenta) questões objetivas de múltipla escolha e uma proposta de redação, passou a ser aplicado em dois dias seguidos (sábado e domingo), sendo que no primeiro dia os participantes recebiam um caderno de questões com as provas de Ciências da Natureza e suas Tecnologias e Ciências Humanas e suas Tecnologias e, no segundo, um caderno com as provas de Redação e Linguagens, Códigos e suas Tecnologias e Matemática e suas Tecnologias.

A partir de então, o Enem tornou-se uma das principais vias de acesso ao Ensino Superior público, democratizando as oportunidades e possibilitando a mobilidade acadêmica, além de continuar sendo referência para a autoavaliação dos estudantes.

A estrutura com 4 (quatro) provas objetivas, uma de cada área do conhecimento, contendo 45 (quarenta e cinco) questões de múltipla escolha e uma redação se mantém, mas as quatro áreas do conhecimento são reagrupadas entre os cadernos de questões, conforme o Quadro 1.

Quadro 1 – Descrição das Áreas de Conhecimento e Componentes Curriculares do Enem.

Área do conhecimento	Componentes Curriculares
Linguagens, Códigos e suas tecnologias	Língua Portuguesa, Literatura, Língua Estrangeira (Inglês ou Espanhol), Artes, Educação Física e Tecnologias da Informação e Comunicação.
Ciências Humanas e suas tecnologias	História, Geografia, Filosofia e Sociologia.
Ciências da Natureza e suas tecnologias	Química, Física e Biologia.
Matemática e suas tecnologias	Matemática.

A partir de 2017, as provas passaram a ser realizadas em dois domingos seguidos. Em 2018, as provas foram aplicadas nos dias 4 e 11 de novembro de 2018, sendo que no primeiro dia os participantes realizaram as provas de Linguagens, Códigos e suas tecnologias e Redação e de Ciências Humanas e suas tecnologias e, no segundo, as provas de Ciências da Natureza e suas

tecnologias e Matemática e suas tecnologias.

A segunda aplicação do Enem 2018, por sua vez, ocorreu nos dias 11 e 12 de dezembro de 2018, para Pessoas Privadas de Liberdade e Jovens sob Medida Socioeducativa que incluía privação de liberdade – PPL, bem como para os participantes com direito à reaplicação.

Os participantes surdos ou deficientes auditivos tiveram a possibilidade de solicitar, no momento da inscrição, o recurso para atendimento especializado videoprova em Libras. Esses participantes receberam, além do caderno de questões impresso, vídeos contendo a tradução correspondente na Língua Brasileira de Sinais - Libras. Esse recurso para atendimento especializado introduzido em caráter experimental em 2017 permaneceu na edição 2018.

Deste modo, além dos cadernos regulares foram disponibilizados os cadernos ampliados, superampliados, ledor (correspondente ao caderno em braille) e a videoprova, conforme a condição de acessibilidade necessária ao participante.

Na aplicação do Enem, um dos procedimentos utilizados é a produção de diferentes cadernos para cada uma das provas. Assim, para cada prova aplicada, são produzidos quatro tipos de cadernos, contendo as mesmas questões, mas organizadas em ordem diferente. Essa diversidade dos cadernos de prova tem como objetivo dificultar as fraudes durante a aplicação do Exame.

Além das provas objetivas e da redação, os participantes do Enem respondem a um questionário que contempla questões sobre seu nível socioeconômico, família, educação e trabalho. Em 2018 o questionário era composto por 27 questões.

3- MICRODADOS DO ENEM

Para atender a demanda dos usuários sobre informações específicas, são disponibilizadas as provas, os gabaritos, as informações sobre os itens, e as notas e o questionário respondido pelos inscritos no Enem. Estas informações estão disponibilizadas no “Microdados_Enem_2018” e aquelas referentes aos itens, em uma planilha específica. O Dicionário contém informações sobre as variáveis contidas em cada base.

Os dados são disponibilizados em formato “.csv” (formato de arquivo que contém valores separados por delimitador com ponto-e-vírgula) e os *inputs* para a leitura desses arquivos foram elaborados utilizando os softwares SAS, SPSS e R. Os *inputs* trazem a possibilidade de carregar os rótulos juntamente com os dados, o que facilita o seu manuseio pelo usuário, ao tornar sua utilização mais intuitiva e imediata. Todavia, isto não dispensa a consulta ao Dicionário de variáveis e às Provas, para obter uma compreensão mais completa da organização do banco de dados e da própria estrutura dos instrumentos utilizados. Não obstante, se o usuário desejar, poderá não usar os *inputs* para abrir as bases, pois o formato “.csv” também foi adotado por facilitar a leitura dos arquivos, independentemente do software estatístico utilizado.

Para facilitar a utilização dos resultados, o arquivo principal (MICRODADOS_ENEM_2018)

contém os questionários respondidos pelos participantes, congregando em um único arquivo as informações disponibilizadas sobre os participantes do Enem.

Com a adoção do modelo simplificado de microdados para o Enem 2018, foram realizadas as seguintes alterações na tabela MICRODADOS_ENEM_2018, em relação à edição anterior do exame:

- Excluir a variável CO_ESCOLA;
- Excluir dos microdados informações referentes aos pedidos de atendimento especializado e específico, recursos de atendimento especializado e específico para a realização da prova;
- Substituir a variável NU_IDADE por TP_FAIXA_ETARIA;
- Excluir informações referentes aos municípios de nascimento e residência do participante.

Foram excluídos da base do microdados os registros dos participantes que realizaram as provas: 487, 488, 489 e 490 de Ciências da Natureza; e 499, 500, 501 e 502 de Matemática. Isso porque o número de participantes que realizaram essas provas era muito pequeno, o que permitiria a sua identificação indevida.

O dicionário também foi disponibilizado em formato “.ods”, tendo em vista atender à política de dados abertos.

Foi criada também uma base contendo informações sobre as provas, chamada ITENS_PROVA_2018. Nela, para cada tipo de prova, foram carregadas as informações de posição do item na prova, habilidade, cor da prova, área de conhecimento e gabarito.

Os procedimentos de cálculo das notas dos participantes e de análise dos itens no Enem têm como base a Teoria de Resposta ao Item (TRI). A TRI é um conjunto de modelos matemáticos que busca representar a relação entre a probabilidade de o participante responder corretamente a uma questão, seu conhecimento na área em que está sendo avaliado e as características (parâmetros) dos itens.

Um aspecto importante da TRI é que ela considera a particularidade de cada item. Assim, as notas não dependem do total de itens da prova, mas de cada item que a compõe. Dessa forma, duas pessoas com a mesma quantidade de acertos na prova são avaliadas de forma distintas a depender de quais itens estão certos e errados e podem, assim, ter notas diferentes.

O modelo matemático da TRI usado no Enem possui três parâmetros que expressam as informações do item, essenciais para avaliar suas características e, conseqüentemente, a medida do conhecimento. São eles:

- Parâmetro de discriminação: é o poder de discriminação que cada questão possui para diferenciar os participantes que dominam dos participantes que não dominam a habilidade avaliada naquela questão.
- Parâmetro de dificuldade: associado à dificuldade da habilidade avaliada na questão, quanto maior seu valor, mais difícil é a questão. Ele é expresso na mesma escala da proficiência. Em uma prova de qualidade, devemos ter questões de diferentes níveis de dificuldade para avaliar adequadamente os participantes em todos os níveis de conhecimento.
- Parâmetro de acerto casual: em provas de múltipla escolha, um participante que não domina a

habilidade avaliada em uma determinada questão da prova pode responder corretamente a um item devido ao acerto casual. Assim, esse parâmetro representa a probabilidade de um participante acertar a questão não dominando a habilidade exigida.

Os parâmetros nunca foram divulgados desde que o exame passou a adotar a TRI, em 2009. Os dados eram acessíveis apenas por meio do Serviço de Acesso a Dados Protegidos (Sedap), do Inep, e somente após cinco anos da aplicação. Nesta edição dos microdados, incluímos, na tabela ITENS_PROVA_2018 os parâmetros dos itens do modelo de TRI.

Incluímos também os seguintes documentos para melhor entendimento dos critérios de correção das provas do ENEM: Entenda sua Nota no ENEM – Guia do Participante e ENEM Procedimentos de Análise.

Nesta edição também incluímos o Manual de Redação do Enem 2018 para melhor auxiliar os usuários.

4- ARQUIVOS DOS MICRODADOS

Os arquivos que compõem os Microdados do Enem 2018 são apresentados, no Quadro 2, com uma breve descrição de seu conteúdo:

Quadro 2 - Pastas e Arquivos dos Microdados do ENEM.

PASTAS	5 Pastas:	
	DADOS	
	DICIONÁRIO	
	LEIA-ME e DOCUMENTOS TÉCNICOS	
	INPUTS	
	PROVAS e GABARITOS	
Dados	2 Arquivos:	Descrição:
	MICRODADOS_ENEM_2018.csv	Informações gerais sobre a realização das provas, a caracterização do participante e da escola que ele declarou ter frequentado, e as notas das provas objetivas e da redação.
	ITENS_PROVA_2018.csv	Informações gerais sobre os itens das provas.
Dicionário	2 Arquivos:	Descrição:
	Dicionário_Microdados_Enem_2018.xlsx Dicionário_Microdados_Enem_2018.ods	Informações gerais sobre as variáveis contidas nas bases e sobre as perguntas e alternativas do questionário aplicado.

Leia-Me e Documentos Técnicos	3 Arquivos:	Descrição:
	Leia-me_Enem_2018.pdf	Breve descrição do Enem, bem como das informações sobre as bases e os arquivos disponibilizados nos Microdados.
	Matriz_referencia_enem.pdf	Apresentação da Matriz de Referência, que compreende os eixos cognitivos, as competências e as habilidades avaliadas em cada área de conhecimento do Ensino Médio.
	Edital_enem_2018.pdf	Edital de publicação do ENEM 2018
	entenda_a_sua_nota_no_enem_guia_do_participante.pdf	Publicação acerca de orientação ao participante sobre o cálculo de notas do ENEM
	enem_procedimentos_de_analise.pdf	Publicação acerca dos procedimentos de análise de itens e cálculo de resultados do ENEM
	Manual_de_redacao_do_enem_2018.pdf	Expõem a metodologia de avaliação da redação, bem como o que se espera do participante em cada uma das competências avaliadas.

Inputs	4 Arquivos:	Descrição:
	SAS:	Programas para leitura dos arquivos de dados no pacote estatístico SAS, com os rótulos (<i>labels</i>) de cada variável.
	INPUT_SAS_MICRODADOS_ENEM_2018.sas	
	INPUT_SAS_ITENS_PROVA_2018.sas	
	SPSS:	Programas para leitura dos arquivos de dados no pacote estatístico SPSS, com os rótulos (<i>labels</i>) de cada variável.
	INPUT_SPSS_MICRODADOS_ENEM_2018.sps	
	INPUT_SPSS_ITENS_PROVA_2018.sps	
	R:	Programas para leitura dos arquivos de dados no pacote estatístico R, com os rótulos (<i>labels</i>) de cada variável.
	INPUT_R_MICRODADOS_ENEM_2018.r	
	INPUT_R_ITENS_PROVA_2018.r	

Provas e Gabaritos	32 Arquivos:	Descrição
	ENEM_2018_P1_CAD_01_DIA_1_AZUL.pdf	Todos os cadernos de prova aplicados no Enem 2018
	ENEM_2018_P1_CAD_01_DIA_1_AZUL_AMPLIADA.pdf	
	ENEM_2018_P1_CAD_01_DIA_1_AZUL_SUPERAMPLIADA.pdf	
	ENEM_2018_P1_CAD_02_DIA_1_AMARELO.pdf	
	ENEM_2018_P1_CAD_03_DIA_1_BRANCO.pdf	
	ENEM_2018_P1_CAD_04_DIA_1_ROSA.pdf	
	ENEM_2018_P1_CAD_09_DIA_1_LARANJA_LEDOR.pdf	
	ENEM_2018_P1_CAD_10_DIA_1_VERDE_LIBRAS.pdf	
	ENEM_2018_P1_CAD_05_DIA_2_AMARELO.pdf	

ENEM_2018_P1_CAD_05_DIA_2_AMPLIADA.pdf	(PRIMEIRA APLICAÇÃO)
ENEM_2018_P1_CAD_05_DIA_2_SUPERAMPLIADA.pdf	
ENEM_2018_P1_CAD_06_DIA_2_CINZA.pdf	
ENEM_2018_P1_CAD_07_DIA_2_AZUL.pdf	
ENEM_2018_P1_CAD_08_DIA_2_ROSA.pdf	
ENEM_2018_P1_CAD_11_DIA_2_LARANJA_LEDOR.pdf	
ENEM_2018_P1_CAD_12_DIA_2_VERDE_LIBRAS.pdf	
ENEM_2018_P1_GAB_01_DIA_1_AZUL.pdf	Todos os gabaritos dos cadernos de prova aplicados no Enem 2018 (PRIMEIRA APLICAÇÃO)
ENEM_2018_P1_GAB_01_DIA_1_AZUL_AMPLIADA.pdf	
ENEM_2018_P1_GAB_01_DIA_1_AZUL_SUPERAMPLIADA.pdf	
ENEM_2018_P1_GAB_02_DIA_1_AMARELO.pdf	
ENEM_2018_P1_GAB_03_DIA_1_BRANCO.pdf	
ENEM_2018_P1_GAB_04_DIA_1_ROSA.pdf	
ENEM_2018_P1_GAB_09_DIA_1_LARANJA_LEDOR.pdf	
ENEM_2018_P1_GAB_10_DIA_1_VERDE_LIBRAS.pdf	
ENEM_2018_P1_GAB_05_DIA_2_AMARELO.pdf	
ENEM_2018_P1_GAB_05_DIA_2_AMARELO_AMPLIADA.pdf	
ENEM_2018_P1_GAB_05_DIA_2_AMARELO_SUPERAMPLIADA.pdf	
ENEM_2018_P1_GAB_06_DIA_2_CINZA.pdf	
ENEM_2018_P1_GAB_07_DIA_2_AZUL.pdf	
ENEM_2018_P1_GAB_08_DIA_2_ROSA.pdf	
ENEM_2018_P1_GAB_11_DIA_2_LARANJA_LEDOR.pdf	
ENEM_2018_P1_GAB_12_DIA_2_VERDE_LIBRAS.pdf	
ENEM_2018_P2_CAD_01_DIA_1_AZUL.pdf	Todos os cadernos de prova aplicados no Enem 2018 (SEGUNDA APLICAÇÃO)
ENEM_2018_P2_CAD_02_DIA_1_AMARELO.pdf	
ENEM_2018_P2_CAD_03_DIA_1_BRANCO.pdf	
ENEM_2018_P2_CAD_04_DIA_1_ROSA.pdf	
ENEM_2018_P2_CAD_05_DIA_2_AMARELO.pdf	
ENEM_2018_P2_CAD_06_DIA_2_CINZA.pdf	
ENEM_2018_P2_CAD_07_DIA_2_AZUL.pdf	
ENEM_2018_P2_CAD_08_DIA_2_ROSA.pdf	
ENEM_2018_P2_GAB_01_DIA_1_AZUL.pdf	
ENEM_2018_P2_GAB_02_DIA_1_AMARELO.pdf	
ENEM_2018_P2_GAB_03_DIA_1_BRANCO.pdf	
ENEM_2018_P2_GAB_04_DIA_1_ROSA.pdf	
ENEM_2018_P2_GAB_05_DIA_2_AMARELO.pdf	
ENEM_2018_P2_GAB_06_DIA_2_CINZA.pdf	
ENEM_2018_P2_GAB_07_DIA_2_AZUL.pdf	
ENEM_2018_P2_GAB_08_DIA_2_ROSA.pdf	

5- COMO CITAR OS MICRODADOS

ISO:

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do Enem 2018*. [online]. Brasília: Inep, 2018. [citado 2019-04-30]. Disponível em: <<http://portal.inep.gov.br/web/guest/microdados>>.

ABNT:

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Microdados do Enem 2018**. Brasília: Inep, 2018. Disponível em: <<http://portal.inep.gov.br/web/guest/microdados>>. Acesso em: 30 abr. 2017.

APA:

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2018). *Microdados do Enem 2018*. Recuperado em 30 de abril de 2019, de <http://portal.inep.gov.br/web/guest/microdados>.

6- CONSIDERAÇÕES FINAIS

Tendo em vista o aperfeiçoamento constante da documentação e da disseminação das bases de dados do Inep, contamos com a colaboração dos usuários para analisar e tecer considerações a respeito dos microdados ora disponibilizados. Para tanto, disponibilizamos um e-mail para receber comentários, críticas e sugestões: **microdados.daeb@inep.gov.br**