



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

# **MICRODADOS ENEM 2016**

## **LEIA-ME**

**Novembro de 2017**

## **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

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

## 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 avaliações e informações educacionais, disponibiliza os **Microdados do Enem 2016**.

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 2016 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 finalidade precípua a avaliação do desempenho escolar ao fim do Ensino Médio, podendo seus resultados ser utilizados para: subsidiar a implementação de políticas públicas; criar referência nacional para o aperfeiçoamento dos currículos do Ensino Médio; desenvolver estudos sobre a educação brasileira; estabelecer critérios de acesso do participante a programas governamentais; selecionar candidatos para o ensino superior; e constituir parâmetros para a autoavaliação do participante, com vistas à continuidade de sua formação e à sua inserção no mercado de trabalho.

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.

A partir de 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.

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. Passou a possibilitar, também, a certificação para conclusão do Ensino Médio, obedecendo às exigências previstas na Lei de Diretrizes e Bases da Educação Nacional (Lei 9.394/1996) para a Educação de Jovens e Adultos.

Na edição do Enem de 2016, foram estruturadas 4 (quatro) provas objetivas, uma de cada área do conhecimento, contendo 45 (quarenta e cinco) questões de múltipla escolha e uma redação, englobando os seguintes componentes curriculares, 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.
Matemática e suas tecnologias	Matemática.
Ciências Humanas e suas tecnologias	História, Geografia, Filosofia e Sociologia.
Ciências da Natureza e suas tecnologias	Química, Física e Biologia.

As provas foram aplicadas nos dias 5 (sábado) e 6 (domingo) de novembro de 2016, sendo que no primeiro dia os participantes realizaram as provas de Ciências Humanas e suas tecnologias e Ciências da Natureza e suas tecnologias e, no segundo, as provas de Linguagens, Códigos e suas tecnologias, Matemática e suas tecnologias e a Redação.

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, trabalho, escola, interesses e expectativas, entre outros temas.

### 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\_2016” 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 e SPSS. 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\_2016) 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 2016, foram realizadas as seguintes alterações na tabela MICRODADOS\_ENEM\_2016, 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 dos microdados os registros dos participantes que realizaram as provas: 334 e 355 de Ciências da Natureza; 346, 347, 348, 349 e 370 de Matemática; 341, 342, 343, 344 e 365 de Línguas e Códigos; e 339 e 360 de Ciências Humanas. Isso porque o número de participantes que realizaram essas provas era muito pequeno, o que permitiria a sua identificação indevida.

O dicionário também foram disponibilizadas 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\_2016. 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, consequentemente, 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.

Cabe reportar aqui que há uma equivalência entre as provas da reaplicação que se dá da seguinte maneira:

ÁREA	CO_PROVA	
CN	331	351
	332	352
	333	353
	334	354
CH	336	356
	337	357
	338	358
	339	359
LC	341	361
	342	362
	343	363
	344	364
MT	346	366
	347	367
	348	368
	349	369

Portanto, por exemplo, a prova 331 é igual à prova 351 e assim sucessivamente.

Nesta edição também incluímos o Manual de Redação do Enem 2016 para melhor auxiliar os usuários.

#### 4- ARQUIVOS DOS MICRODADOS

Os arquivos que compõem os Microdados do Enem 2016 são apresentados, no Quadro 2, com uma breve descrição de seu conteúdo:

**Quadro 2 - Pastas e Arquivos dos Microdados do ENEM.**

<b>PASTAS</b>	<b>5 Pastas:</b>	
	DADOS	
	DICIONÁRIO	
	LEIA-ME e DOCUMENTOS TÉCNICOS	
	INPUTS	
	PROVAS e GABARITOS	
	<b>2 Arquivos:</b>	<b>Descrição:</b>
	MICRODADOS_ENEM_2016.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.



<b>Dados</b>	ITENS_PROVA_2016.csv	Informações gerais sobre os itens das provas.
--------------	----------------------	---

<b>Dicionário</b>	<b>2 Arquivos:</b>	<b>Descrição:</b>
	Dicionário_Microdados_Enem_2016.xlsx Dicionário_Microdados_Enem_2016.ods	Informações gerais sobre as variáveis contidas nas bases e sobre as perguntas e alternativas do questionário aplicado.

<b>Leia-Me e Documentos Técnicos</b>	<b>3 Arquivos:</b>	<b>Descrição:</b>
	Leia-me_Enem_2016.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_2016.pdf	Edital de publicação do ENEM 2016
	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_2016.pdf	Expõem a metodologia de avaliação da redação, bem como o que se espera do participante em cada uma das competências avaliadas.
	Manual_de_redacao_do_enem_2016	Expõem a metodologia de avaliação da redação, bem como o que se espera do participante em cada uma das competências avaliadas.

<b>Inputs</b>	<b>4 Arquivos:</b>	<b>Descrição:</b>
	<b>SAS:</b>	Programas para leitura dos arquivos de dados no pacote estatístico SAS, com os rótulos ( <i>labels</i> ) de cada variável.
	INPUT_SAS_MICRODADOS_ENEM_2016.sas	
	INPUT_SAS_ITENS_PROVA_2016.sas	
	<b>SPSS:</b>	Programas para leitura dos arquivos de dados no pacote estatístico SPSS, com os rótulos ( <i>labels</i> ) de cada variável.
	INPUT_SPSS_MICRODADOS_ENEM_2016.sps	
	INPUT_SPSS_ITENS_PROVA_2016.sps	
	<b>R:</b>	Programas para leitura dos arquivos de dados no pacote estatístico R, com os rótulos ( <i>labels</i> ) de cada variável.
	INPUT_R_MICRODADOS_ENEM_2016.r	
	INPUT_R_ITENS_PROVA_2016.r	

<b>Provas e Gabaritos</b>	<b>32 Arquivos:</b>	<b>Descrição</b>
	CAD_ENEM_2016_DIA_1_01_AZUL.pdf	Todos os cadernos de prova aplicados no Enem 2016 ( <b>PRIMEIRA APLICAÇÃO</b> )
	CAD_ENEM_2016_DIA_1_02_AMARELO.pdf	
	CAD_ENEM_2016_DIA_1_03_BRANCO.pdf	
	CAD_ENEM_2016_DIA_1_04_ROSA.pdf	
	CAD_ENEM_2016_DIA_2_05_AMARELO.pdf	
	CAD_ENEM_2016_DIA_2_06_CINZA.pdf	
	CAD_ENEM_2016_DIA_2_07_AZUL.pdf	
	CAD_ENEM_2016_DIA_2_08_ROSA.pdf	
	GAB_ENEM_2016_DIA_1_01_AZUL.pdf	Todos os gabaritos dos cadernos de prova aplicados no Enem 2016 ( <b>PRIMEIRA APLICAÇÃO</b> )
	GAB_ENEM_2016_DIA_1_02_AMARELO.pdf	
	GAB_ENEM_2016_DIA_1_03_BRANCO.pdf	
	GAB_ENEM_2016_DIA_1_04_ROSA.pdf	
	GAB_ENEM_2016_DIA_2_05_AMARELO.pdf	
	GAB_ENEM_2016_DIA_2_06_CINZA.pdf	
	GAB_ENEM_2016_DIA_2_07_AZUL.pdf	
	GAB_ENEM_2016_DIA_2_08_ROSA.pdf	
	CAD_ENEM_2016_DIA_1_01_AZUL_2.pdf	Todos os cadernos de prova aplicados no Enem 2016 ( <b>SEGUNDA APLICAÇÃO</b> )
	CAD_ENEM_2016_DIA_1_02_AMARELO_2.pdf	
	CAD_ENEM_2016_DIA_1_03_BRANCO_2.pdf	
	CAD_ENEM_2016_DIA_1_04_ROSA_2.pdf	
	CAD_ENEM_2016_DIA_2_05_AMARELO_2.pdf	
	CAD_ENEM_2016_DIA_2_06_CINZA_2.pdf	
	CAD_ENEM_2016_DIA_2_07_AZUL_2.pdf	
	CAD_ENEM_2016_DIA_2_08_ROSA_2.pdf	
	GAB_ENEM_2016_DIA_1_01_AZUL_2.pdf	Todos os gabaritos dos cadernos de prova aplicados no Enem 2016 ( <b>SEGUNDA APLICAÇÃO</b> )
	GAB_ENEM_2016_DIA_1_02_AMARELO_2.pdf	
	GAB_ENEM_2016_DIA_1_03_BRANCO_2.pdf	
	GAB_ENEM_2016_DIA_1_04_ROSA_2.pdf	
	GAB_ENEM_2016_DIA_2_05_AMARELO_2.pdf	

	GAB_ENEM_2016_DIA_2_06_CINZA_2.pdf	<b>APLICAÇÃO)</b>
	GAB_ENEM_2016_DIA_2_07_AZUL_2.pdf	
	GAB_ENEM_2016_DIA_2_08_ROSA_2.pdf	

## 5- COMO CITAR OS MICRODADOS

### **ISO:**

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do Enem 2016*. [online]. Brasília: Inep, 2017. [citado 2017-05-30]. Disponível em: <<http://portal.inep.gov.br/web/guest/microdados>>.

### **ABNT:**

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Microdados do Enem 2016**. Brasília: Inep, 2017. Disponível em: <<http://portal.inep.gov.br/web/guest/microdados>>. Acesso em: 01 nov. 2017.

### **APA:**

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2016). *Microdados do Enem 2016*. Recuperado em 01 de novembro de 2017, de <http://portal.inep.gov.br/basica-levantamentos-acessar>.

## 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](mailto:microdados.daeb@inep.gov.br)**