

# MICRODADOS ENEM 2010



**Presidência da República Federativa do Brasil**

**Ministério da Educação**

**Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP)**

**Diretoria de Avaliação da Educação Básica (DAEB)**

**Diretoria de Tecnologia e Disseminação de Informações Educacionais (DTDIE)**

**Ministério da Educação**  
**Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira**

**Microdados do ENEM - 2010**

**Exame Nacional do Ensino Médio**

Manual do Usuário

**Diretoria de Avaliação da Educação Básica**

Maria Tereza Serrano Barbosa

**Coordenação-Geral do Sistema Nacional de Avaliação da Educação Básica**

Clara Etiene Lima de Souza

**Coordenação-Geral de Instrumentos e Medidas**

Camila Akemi Karino

**Coordenação Geral de Exames e Certificação**

Lenice Medeiros

**Equipe Técnica**

Joelson Severo dos Santos Azevêdo

José Dias Pereira

Luana Bergmann Soares

Rogério Diniz Junqueira

**Diretoria de Tecnologia da Informação**

Andrea de Miranda Ramos Kern

**Coordenação-Geral de Infraestrutura e Serviços**

Francisco Edilson de Carvalho Silva

**Coordenadora-Geral de Sistemas de Informação**

Ariene Mariano Rêgo Crispim

**Equipe Técnica**

Wesley Lourenço de Araújo

Anderson Santos Leal

# SUMÁRIO

APRESENTAÇÃO.....	6
O Exame Nacional de Avaliação do Ensino Médio.....	6
ARQUIVOS DOS MICRODADOS.....	8
1. Dados .....	8
2. Dicionário .....	8
3. Leia-me e documentos técnicos .....	8
4. Inputs .....	8
5. Provas e gabaritos .....	9
COMO CITAR OS MICRODADOS.....	9
CONSIDERAÇÕES FINAIS .....	9

## APRESENTAÇÃO

O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - Inep - tem como finalidade institucional desenvolver, na área educacional, sistemas de informação e documentação, estudos, pesquisas, estatísticas, avaliações e disseminação de informações educacionais para a sociedade como um todo. No cumprimento dessas funções, a Diretoria de Avaliação da Educação Básica tem como uma de suas responsabilidades, a realização do Exame Nacional do Ensino Médio – Enem, que tem como principal objetivo avaliar o Ensino Médio no Brasil.

A fim de garantir a transparência requerida pela sociedade e promover a participação dos diferentes agentes sociais envolvidos nas ações educativas, o Inep disponibiliza em seu [site \(www.inep.gov.br\)](http://www.inep.gov.br) os Microdados dos exames e avaliações realizados pela DAEB, entre outros. Dentre as informações educacionais produzidas pelo Inep, destacam-se as relativas ao Exame Nacional do Ensino Médio (Enem), realizado anualmente pela Diretoria de Avaliação da Educação Básica e destinado aos alunos que estão concluindo, que já concluíram ou que desejam obter certificação de conclusão no ensino médio.

Para agilizar e melhor atender às demandas de informações do Enem, o Inep está tornando acessível aos seus usuários os **Microdados do Enem 2010**. Além dos dados em formato .CSV (separado por virgula), foram elaborados inputs para a leitura dos arquivos utilizando o *software* SAS, R e o *software* SPSS.

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 2010 do ENEM. O objetivo é eliminar da base pública variáveis que facilitem a identificação indevida do participante.

## O Exame Nacional de Avaliação do Ensino Médio

Desde sua primeira realização, os pressupostos teórico-metodológicos do Exame Nacional do Ensino Médio (Enem) foram sendo cada vez mais explicitados e anunciados à comunidade educacional brasileira, que se debruçou com empenho e profissionalismo na tarefa de compreender a proposta do exame em suas múltiplas dimensões, avaliando-a com criterioso rigor, oferecendo valiosas contribuições ao modelo proposto.

O Enem caracteriza-se por ser uma avaliação centrada no indivíduo, mas que também promove uma reflexão sobre aquilo que se estuda nas escolas brasileiras da Educação Básica, pois, com base em seus resultados, o jovem passa a compreender melhor o horizonte de expectativas a respeito do Ensino Médio no país. Além disso, a utilização do exame como um dos critérios para seleção dos estudantes que concorrem a vagas do Ensino Superior deve-se, entre outros fatores, à credibilidade adquirida desde sua primeira edição, em 1998, consolidada ao longo desses doze anos de realização; credibilidade esta, evidenciada pelo aumento gradativo e constante do número de Instituições de Ensino Superior que vêm aderindo ao exame e aos programas desenvolvidos pelo MEC.

De 1998 a 2008, o Enem era realizado com a aplicação de uma única prova de 63 questões interdisciplinares. Durante esse período, algumas instituições utilizavam-no como estratégia de seleção para ingresso no Ensino Superior, mas tratava-se de ações isoladas e dispersas no país.

A partir da edição de 2009, ocorreu a reformulação metodológica do Enem e de sua utilização como forma de seleção unificada nos processos seletivos das Universidades Federais. A partir de então, o Enem tornou-se uma das principais vias de acesso para o Ensino Superior, democratizando as oportunidades e possibilitando a mobilidade acadêmica e a reestruturação dos currículos do Ensino Médio, além de continuar sendo referência para a autoavaliação dos estudantes.

A partir de 2009, o Enem passou a possibilitar, também, a certificação para conclusão no Ensino Médio, obedecendo-se às exigências previstas na LDB para a Educação de Jovens e Adultos, como a necessidade do certificando possuir 18 anos completos na data da realização da prova.

A edição de 2010 do Enem foi estruturada a partir da nova Matriz de Referência, subdividida em quatro áreas do Conhecimento. O Exame está estruturado por 4 (quatro) provas objetivas, contendo cada uma 45 (quarenta e cinco) questões de múltipla escolha e uma redação.

As 4 (quatro) provas objetivas e a proposta de redação avaliam as seguintes áreas de conhecimento do Ensino Médio e respectivos componentes curriculares:

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

As provas do Enem em 2010 foram aplicadas em dois dias de realização. No primeiro dia os participantes realizam as provas de Ciências Humanas e suas Tecnologias, e Ciências da Natureza e suas Tecnologias; no segundo dia são realizadas as provas de Linguagens, Códigos e suas Tecnologias e Redação, e Matemática. Por essa razão, o participante que tenha comparecido aos dois dias de prova, terá recebido quatro cadernos de provas.

Uma das metodologias utilizadas para a aplicação do Enem é a produção de diferentes cadernos para cada uma das provas a serem realizadas. Essa diversidade de modelos de provas tem como objetivo dificultar fraudes durante a aplicação das provas. Assim, para cada prova aplicada no Enem, são produzidos quatro tipos de cadernos, sendo que todos possuem as mesmas questões, embora elas estejam organizadas em diferentes ordens, ou seja, para cada prova do Enem são gerados quatro cadernos com

diferentes rotações das mesmas quarenta e cinco questões por área do conhecimento.

Além das questões objetivas e da prova de redação, cada um dos participantes do Enem reponde, também, a um Questionário Socioeconômico. Os Questionários Socioeconômicos do Enem permitem, entre outras coisas, o entrecruzamento dos resultados obtidos com a realidade de cada participante enquanto sujeito da educação, isto é, os dados socioeconômicos e profissionais deles e de suas famílias, seus percursos escolares, suas avaliações acerca dos seus estudos, professores e escolas, suas finalidades ao participar do exame, suas atuações no mundo do trabalho, as dificuldades enfrentadas, os preconceitos e as discriminações vivenciados na escola, suas opiniões sobre assuntos gerais, seus interesses e planos para o futuro etc. Enfim, são dados indispensáveis e estratégicos para permitir melhor compreensão sobre as informações gerais fornecidas pelo próprio exame (escores, escala pedagógica, participação ao exame etc.) à medida que se estabelecessem cruzamentos entre aqueles dados e estas informações.

Ademais, o Questionário Socioeconômico respondido pelos participantes do Enem consiste em um valioso instrumento de que o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) dispõe para obter informações diretas sobre os egressos dos sistemas de ensino que participam do Exame. Além disso, é um instrumento em que são coletadas informações sobre a prática de leitura dos participantes, suas práticas de lazer, suas inserções no mundo do trabalho, suas opiniões sobre seus percursos escolares, seus professores, suas escolas, seus interesses, suas expectativas e suas experiências.

Portanto, para a análise das informações pertinentes aos Microdados, é necessário conhecer os instrumentos aplicados aos participantes, por isso, tanto os cadernos de provas, quanto os Questionários Socioeconômicos são documentos que compõem esta publicação e encontram-se compilados junto aos demais dados.

Com a adoção do modelo simplificado de microdados para o Enem 2010, foram realizadas as seguintes alterações na tabela MICRODADOS\_ENEM\_2010, 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.

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\_2010 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.

## **ARQUIVOS DOS MICRODADOS**

### **1. Dados**

Contém o arquivo dos Microdados do Exame Nacional do Ensino Médio 2010, em formato CSV: MICRODADOS\_ENEM\_2010. O arquivo contém informações relativas à caracterização do candidato; caracterização da escola frequentada pelo candidato; nota das Provas Objetivas, nota da Redação – global e por competências, respostas das Provas Objetivas, gabaritos das Provas Objetivas e questionário Socioeconômico.

### **2. Dicionário**

Informações gerais sobre as variáveis contidas nas bases e sobre as perguntas e alternativas do questionário aplicado, no formato XLSX e ODS.

### **3. Leia-me e documentos técnicos**

- **LEIA-ME:** Arquivo no formato pdf contendo descrição do Enem 2010, bem como das informações sobre as bases e os arquivos disponibilizados nos Microdados.
- **Questionario\_Socioeconomico:** Questionário Socioeconômico Exame do Nacional do Ensino Médio 2010, em formato PDF.
- **Matriz\_referencia:** Apresentação da Matriz de Referência, que compreende os eixos

cognitivos, as competências e as habilidades avaliadas em cada área de conhecimento do Ensino Médio.

- Relatorio\_pedagogico\_enem\_2010 – Apresenta aspectos relevantes do perfil e do desempenho dos participantes no Enem, além de sistematizar dados e informações gerais de cada aplicação.
- Edital\_enem\_2010 – Edital de publicação do ENEM 2010.

#### **4. Inputs**

- INPUT\_SAS\_MICRODADOS\_ENEM\_2010 – Programa de leitura das variáveis dos Microdados do Enem 2009, do tipo SAS, com rótulo (labels) de cada variável.
- INPUT\_SAS\_ITENS\_PROVA\_2010 - Programas para leitura das variáveis da base de Itens da prova no pacote estatístico SAS, com os rótulos (labels) de cada variável.
- INPUT\_SPSS\_MICRODADOS\_ENEM\_2010 – Programa de leitura das variáveis dos Microdados do Enem 2009, do tipo SPSS, com rótulo (labels) de cada variável.
- INPUT\_SPSS\_ITENS\_PROVA\_2010 - Programas para leitura das variáveis da base de Itens da prova no pacote estatístico SPSS, com os rótulos (labels) de cada variável.
- INPUT\_R\_MICRODADOS\_ENEM\_2010 – Programa de leitura das variáveis dos Microdados do Enem 2010, do tipo R.
- INPUT\_R\_ITENS\_PROVA\_2010 - Programas para leitura das variáveis da base de Itens da prova no pacote estatístico R.

#### **5. Provas e gabaritos**

Provas Objetivas do Exame Nacional do Ensino Médio 2010 e seus respectivos gabaritos, em formato PDF.

- ENEM\_2010\_PROVA\_GAB\_DIA\_2\_AZUL\_7 – Prova e gabarito do segundo dia – Cor Azul Caderno 7. Os cadernos possuem as mesmas questões independente do modelo, o que os diferenciam é a posição das questões;
- ENEM\_2010\_PROVA\_GAB\_DIA\_1\_AZUL\_1 – Prova e gabarito do primeiro dia – Cor Azul Caderno 1. Os cadernos possuem as mesmas questões independente do modelo, o que os diferenciam é a posição das questões;
- ENEM\_2010\_PROVA\_GAB\_DIA\_2\_AZUL\_7(2\_Aplicação) – Prova e gabarito do segundo dia de reaplicação – Cor Azul Caderno 7. Os cadernos possuem as mesmas questões independente do modelo, o que os diferenciam é a posição das questões;
- ENEM\_2010\_PROVA\_GAB\_DIA\_1\_AZUL\_1(2\_Aplicação) – Prova e gabarito do primeiro dia de reaplicação – Cor Azul Caderno 1. Os cadernos possuem as mesmas questões independente do modelo, o que os diferenciam é a posição das questões;

## COMO CITAR OS MICRODADOS

### **ISO:**

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do Enem 2010*. [online]. Brasília: Inep, 2010. [citado 2010-05-30]. Disponível em: <<http://portal.inep.gov.br/web/guest/microdados>>.

### **ABNT:**

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Microdados do Enem 2010**. Brasília: Inep, 2010. Disponível em: <<http://portal.inep.gov.br/web/guest/microdados>>. Acesso em: 01 nov. 2010.

### **APA:**

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2010). *Microdados do Enem 2010*. Recuperado em 01 de novembro de 2010, de <http://portal.inep.gov.br/basica-levantamentos-acessar>.

## CONSIDERAÇÕES FINAIS

Segue abaixo a lista das mudanças realizadas nesta segunda edição:

1. Unimos as tabelas DADOS\_ENEM\_2010 e QUESTIONARIO\_SOCIO\_ECONOMICO\_ENEM\_2010 e criamos a tabela MICRODADOS\_ENEM\_2010.csv.
2. Renomeamos as variáveis para o padrão definido. Lista de mudanças no dicionário;
3. Transformamos a base de Itens da Prova em arquivo .csv;
4. Adaptamos o dicionário para o novo modelo da base e adicionar uma planilha com o dicionário da base de itens;
5. Além do formato do dicionário salvo em .xlsx, também criamos um no formato .ODS;
6. Modificamos a estrutura do LEIA\_ME. Retiramos dicionário das bases e os programas de input. Adicionamos informações sobre Como Citar os Microdados e Considerações Finais;
- a. Adicionamos os seguintes documentos: Manual do Inscrito no Enem (contendo: Matriz de referência do Enem 2010); Edital do Enem 2010; Relatório Pedagógico do Enem 2010;
7. Ajustamos o *Input* do SAS e SPSS para importar a base dos Microdados em formato .csv e adicionar legenda para descrever as variáveis e suas categorias;
8. Criamos *input* no formato SAS, SPSS e R para a base de Itens da prova;
9. Criamos *input* no formato R para a base de Microdados;
10. Incluímos os modelos das provas e gabaritos do Enem 2010.

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)**