

# MICRODADOS ENEM 2012



**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)**

**Microdados do ENEM - 2012**

**Exame Nacional do Ensino Médio**

Manual do Usuário

Dezembro de 2013

**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 e Certificação**

**Diretoria de Tecnologia da Informação**  
**Coordenadora-Geral de Sistemas de Informação**  
**Coordenação-Geral de Infra-Estrutura e Serviços**

# SUMÁRIO

APRESENTAÇÃO .....	6
I. DIRETÓRIOS – Informações das pastas existentes no CD-ROM dos Microdados do Exame Nacional do Ensino Médio 2012.....	8

# APRESENTAÇÃO

O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), por intermédio da Diretoria de Avaliação da Educação Básica, em cumprimento a sua missão de desenvolver e disseminar avaliações e informações educacionais, disponibiliza os **Microdados do Enem 2012**.

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. Trata-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 2012 do ENEM. O objetivo é eliminar da base pública variáveis que facilitem a identificação indevida do participante.

O Exame Nacional do Ensino Médio (Enem) tem como finalidade precípua a avaliação do desempenho escolar e acadêmico ao fim do Ensino Médio, podendo seus resultados ser utilizados para: compor a avaliação de medição da qualidade do Ensino Médio no País; subsidiar a implementação de políticas públicas; criar referência nacional para o aperfeiçoamento dos currículos do Ensino Médio; desenvolver estudos e indicadores sobre a educação brasileira; estabelecer critérios de acesso do participante a programas governamentais; 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 unificada 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 LDB para a Educação de Jovens e Adultos.

Na edição do Enem de 2012, 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:

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 3 e 4 de novembro de 2012, 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 Redação.

Uma das metodologias utilizadas para a aplicação do Enem é a produção de diferentes cadernos para cada uma das provas. Assim, para cada prova aplicada no Enem, 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 fraudes durante a aplicação das provas.

Além das provas objetivas e da redação, os participantes do Enem respondem a um questionário socioeconômico que contempla questões sobre seu perfil socioeconômico, família, educação, atividades de trabalho e lazer, escola, interesses, expectativas etc.

Para atender a demanda de pesquisadores sobre informações específicas da prova e dos questionários, são disponibilizadas as estruturas das provas e dos questionários. Os dados são disponibilizados em formato .csv e foram elaborados *inputs* para a leitura dos arquivos utilizando os softwares SAS e SPSS.

Uma novidade trazida por esses *inputs* é a possibilidade de carregar os rótulos juntamente com os dados, o que permitirá aos pesquisadores terem acesso, de maneira mais intuitiva, aos enunciados das questões e das respectivas alternativas, o que pode facilitar o conhecimento e a sumarização dos dados. Todavia, isto não dispensa a consulta às estruturas das provas e dos questionários, para uma compreensão mais completa da organização do banco de dados.

Com a adoção do modelo simplificado de microdados para o Enem 2012, foram realizadas as seguintes alterações na tabela MICRODADOS\_ENEM\_2012, 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\_2012 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.



## **I. DIRETÓRIOS – Informações das pastas existentes no CD-ROM dos Microdados do Exame Nacional do Ensino Médio 2012.**

### **LEIA-ME**

Arquivo no formato pdf contendo o Manual do Usuário dos microdados do Enem 2012: Apresentação, Dicionário das Variáveis, INPUTS de leitura dos dados do tipo SAS, SPSS e R, Estrutura da Prova.

### **DADOS**

Contém o arquivo dos microdados do Exame Nacional do Ensino Médio 2012, em formato “csv”: MICRODADOS\_ENEM\_2012 e ITENS\_ENEM\_2012.

### **INPUTS\_SAS**

Contém os INPUTS para leitura dos microdados:

INPUT\_SAS\_MICRODADOS\_ENEM\_2012 – Programa de leitura das variáveis dos microdados do Enem 2012, do tipo SAS, com 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 Questionários. O programa abre os arquivos com os rótulos (labels) de cada variável.

INPUT\_SAS\_ITENS\_PROVA\_2012 – Programa de leitura dos dados referentes aos itens das provas aplicadas no ENEM 2012.

### **INPUTS\_SPSS**

Contém os INPUTS para leitura dos microdados:

INPUT\_SPSS\_ENEM\_2012 – Programa de leitura das variáveis dos microdados do Enem 2012, do tipo SPSS, com 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 Questionários. O programa abre os arquivos com os rótulos (labels) de cada variável.

INPUT\_SPSS\_ITENS\_PROVA\_2012 – Programa de leitura dos dados referentes aos itens das provas aplicadas no ENEM 2012.

### **INPUTS\_R**

Contém os INPUTS para leitura dos microdados:

INPUT\_R\_ENEM\_2012 – Programa de leitura das variáveis dos microdados do Enem 2012, do tipo SPSS, com 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 Questionários. O programa abre os arquivos com os rótulos (labels) de cada variável.

INPUT\_R\_ITENS\_PROVA\_2012 – Programa de leitura dos dados referentes aos itens das provas aplicadas no ENEM 2012.

### **PROVAS E GABARITOS**

Todos os modelos de prova e seus respectivos gabaritos do Enem 2012 podem ser encontrados na pasta PROVAS E GABARITOS. Estes arquivos também estão disponíveis na página <http://portal.inep.gov.br/web/enem/edicoes-anteriores/provas-e-gabaritos>

### **MATRIZ DE REFERÊNCIA**

A referência das habilidades listadas na estrutura da prova estão no arquivo Matriz\_Referencia\_Enem. Elas também podem ser acessadas na página

[http://download.inep.gov.br/educacao\\_basica/enem/downloads/2012/matriz\\_referencia\\_enem.pdf](http://download.inep.gov.br/educacao_basica/enem/downloads/2012/matriz_referencia_enem.pdf)

### **EDITAL DO ENEM**

Edital de publicação do ENEM 2012