



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

MICRODADOS ENEM 2014

LEIA-ME

Novembro de 2015

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.....	5
4- ARQUIVOS DOS MICRODADOS	6
5- COMO CITAR OS MICRODADOS	8
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 avaliações e informações educacionais, disponibiliza os **Microdados do Enem 2014**.

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 2014 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 e acadêmico ao fim do Ensino Médio, podendo seus resultados ser utilizados para: compor a avaliação 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; 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 2014, 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 08 (sábado) e 09 (domingo) de novembro de 2014, 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 “Dicionário_Microdados_Enem_2014” e as referentes aos itens, em uma planilha específica. O Dicionário também contém informações sobre as demais variáveis das bases disponibilizadas.

Os dados são disponibilizados em formato “.csv” (formato de arquivo que contém valores separados por delimitador com vírgula) e os *inputs* para a leitura desses arquivos foram elaborados utilizando os softwares SAS e SPSS. Além disso, a novidade, trazida na edição de 2012 dos Microdados do Enem, foi mantida nesses *inputs*, qual seja, 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.

Além disso, as inovações apresentadas na edição dos Microdados do Enem 2013 também foram mantidas. Uma delas diz respeito à integração entre as bases de resultados com a do questionário respondido pelos participantes. Esta mudança facilitou o uso da base, pois congregou em uma base única as informações disponibilizadas sobre os participantes do Enem.

Com a adoção do modelo simplificado de microdados para o Enem 2014, foram realizadas as seguintes alterações na tabela MICRODADOS_ENEM_2014, 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.

Na edição dos microdados de 2014, as planilhas também foram disponibilizadas em formato “.ods”, tendo em vista atender à política de dados abertos.

Por fim, é preciso ressaltar que foram eliminados, da base MICRODADOS_ENEM_2014, 37 alunos que participaram da reaplicação do Enem, a fim de evitar que tais alunos pudessem ser identificados. Com esse mesmo intuito, a variável NU_INSCRICAO do aluno é uma máscara e não o seu número de inscrição original no Enem. Devido a eliminação desses alunos, os resultados do Enem 2014 por Escola de 4 (quatro) escolas não poderão ser reproduzidos.

Foi criada também uma base contendo informações sobre as provas, chamada ITENS_PROVA_2014. 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_2014 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.

4- ARQUIVOS DOS MICRODADOS

Os arquivos que compõem os Microdados do Enem 2014 são apresentados, no Quadro 2, com uma breve descrição de seu conteúdo:

Quadro 2 - Pastas e Arquivos dos Microdados do ENEM.

PASTAS	6 Pastas:
	DADOS
	DICIONÁRIO
	LEIA-ME e DOCUMENTOS TÉCNICOS
	INPUTS
	PROVAS e GABARITOS

Dados	2 Arquivos:	Descrição:
	MICRODADOS_ENEM_2014.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_2014.csv	Informações gerais sobre os itens das provas.

Dicionário	2 Arquivos:	Descrição:
	Dicionário_Microdados_Enem_2014.xlsx Dicionário_Microdados_Enem_2014.ods	Informações gerais sobre as variáveis contidas nas bases e sobre as perguntas e alternativas do questionário aplicado.

Leia-Me e Documentos Técnicos	4 Arquivos:	Descrição:
	Leia-me_Enem_2014.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.
	Portaria_Inep_Enem_por_Escola_2014.pdf	Documento que estabelece os procedimentos para o cálculo e os critérios para a divulgação dos resultados do Exame Nacional do Ensino Médio por Escola 2014.
	enem_procedimentos_de_analise	Publicação acerca do cálculo de notas do ENEM
	entenda_a_sua_nota_no_enem_guiado_participante	Publicação acerca de orientação ao participante sobre o cálculo de notas do ENEM
	Nota_Explicativa_Enem_Por_Escola_2014.pdf	Nota que explica os critérios utilizados para a divulgação dos resultados do Enem 2014 por Escola.

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_SPSS_MICRODADOS_ENEM_2014.sas	
	INPUT_SPSS_CONSISTENCIA_CENSO_2014.sas	
	SPSS:	Programas para leitura dos arquivos de dados no pacote estatístico SPSS, com os rótulos (<i>labels</i>) de cada variável.
	INPUT_SAS_MICRODADOS_ENEM_2014.sps	
	INPUT_SAS_CONSISTENCIA_CENSO_2014.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_2014.r	
	INPUT_R_ITENS_PROVA_2014.r	

Provas e Gabaritos	16 Arquivos:	Descrição
	Caderno1_Azul_Sab.pdf	Todos os cadernos de prova aplicados no Enem 2014.
	Caderno2_Amarelo_Sab.pdf	
	Caderno3_Branco_Sab.pdf	
	Caderno4_Rosa_Sab.pdf	
	Caderno5_Amarelo_Dom.pdf	
	Caderno6_Cinza_Dom.pdf	
	Caderno7_Azul_Dom.pdf	
	Caderno8_Rosa_Dom.pdf	
	Gab_Caderno1_Azul_Sab.pdf	Todos os gabaritos dos cadernos de prova aplicados no Enem 2014.
	Gab_Caderno2_Amarelo_Sab.pdf	
	Gab_Caderno3_Branco_Sab.pdf	
	Gab_Caderno4_Rosa_Sab.pdf	
	Gab_Caderno5_Amarelo_Dom.pdf	
	Gab_Caderno6_Cinza_Dom.pdf	
	Gab_Caderno7_Azul_Dom.pdf	
	Gab_Caderno8_Rosa_Dom.pdf	

5- COMO CITAR OS MICRODADOS

ISO:

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do Enem 2014*. [online]. Brasília: Inep, 2015. [citado 2015-05-30]. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>.

ABNT:

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Microdados do Enem 2014**. Brasília: Inep, 2015. Disponível em: <<http://portal.inep.gov.br/basicas-levantamentos-acessar>>. Acesso em: 30 nov. 2015.

APA:

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2015). *Microdados do Enem 2014*. Recuperado em 30 de novembro de 2015, de <http://portal.inep.gov.br/basicas-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**