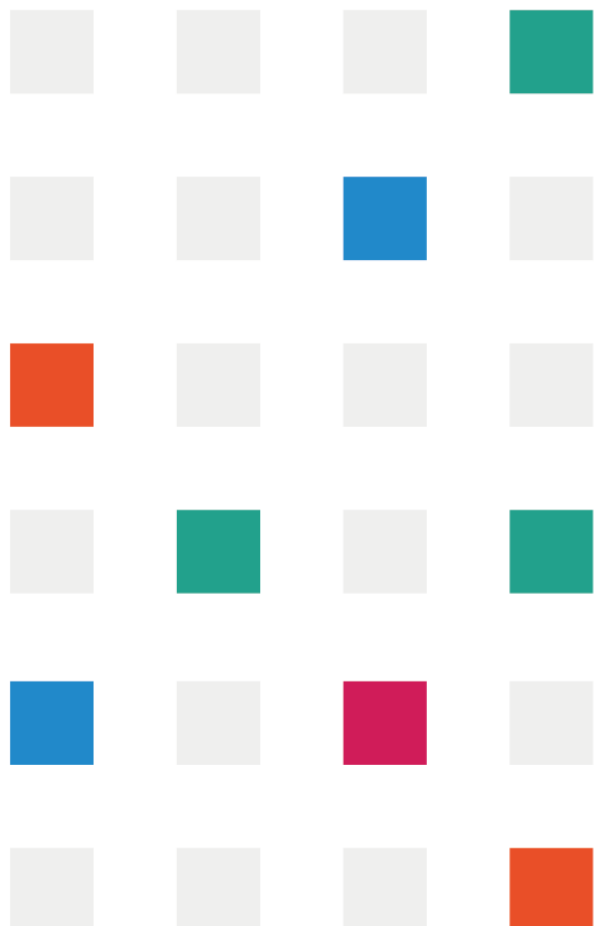


# ALUNOS



CENTRO DE ESTUDOS SOCIEDADE, UNIVERSIDADE E CIÊNCIA

MICRODADOS  
CENSO DA EDUCAÇÃO SUPERIOR - ALUNOS

Autor da metodologia: SoU_Ciência	Fonte Original: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep)
Versão: 01	Data da publicação: 12/07/2023





SoU\_Ciência

### **Coordenação Geral SoU\_Ciência:**

Soraya Soubhi Smaili (Unifesp)

### **Coordenação de Pesquisa:**

Maria Angélica Pedra Minhoto (Unifesp)

Pedro Arantes (Unifesp)

## **Grupo de Dados**

### **Coordenação:**

Thaís Cavalcante Martins

### **Pesquisadora Colaboradora:**

Daniela Leal Musa (Unifesp)

### **Equipe Técnica:**

Alexsandro Cardoso Carvalho

Ecila Alves de Oliveira Migliori

Antonio Guerreiro Silva Serrano


Meyrele Torres Nascimento

Fabiana Yumi Sato Kayanoki

Ricardo Ireno dos Santos



Nosso conteúdo está disponível sob a licença Creative Commons Atribuição 4.0 Internacional, e pode ser compartilhado e reutilizado para trabalhos derivados, desde que citada a fonte.





# INTRODUÇÃO

O conjunto de dados SoU\_Ciência foi construído a partir dos microdados do Censo da Educação Superior disponibilizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Este conjunto de dados surge como uma iniciativa para facilitar o acesso e uso dessas informações.


Os microdados originais (Inep) possuem informações sobre as Instituições de Ensino Superior do país, cursos ofertados, docentes e discentes. O SoU\_Ciência disponibiliza os microdados dos anos de 2009 a 2019 de forma categorizada e padronizada, permitindo que o público tenha acesso a informações precisas e confiáveis sobre as Instituições de Ensino Superior do país.

A equipe do SoU\_Ciência realizou um rigoroso processo de tratamento nos dados originais, visando facilitar a realização de estudos e análises mais precisas e confiáveis na área. Esse processo de tratamento incluiu, principalmente, a compatibilização de variáveis que se modificaram ao longo dos anos, sendo feita uma correspondência, ou seja, uma padronização, entre as variáveis equivalentes, mas que haviam sido expressas de modo diferente em algum ano da série histórica.

Cabe salientar que os dados originais foram indisponibilizados pelo INEP a partir de fevereiro de 2022. Entretanto, o SoU\_Ciência se utilizou da cópia dos dados que haviam sido previamente salvos em arquivos locais. Além disso, em todo o processo a equipe do SoU\_Ciência manteve a fidedignidade dos dados originais.

Com o objetivo de garantir a proteção do dado pessoal e a garantia do sigilo das pessoas a quem os dados se referem, os dados aqui disponibilizados estão em conformidade com a Lei Geral de Proteção de Dados Pessoais, Lei nº 13.709/2018. É importante salientar que mesmos os dados pseudoanonimizados, que permitissem uma mínima identificação, foram retirados, como, por exemplo: data e local de nascimento.

## Objetivo

- permitir o acesso e facilitar o uso dos microdados *padronizados e categorizados* do Censo da Educação Superior (arquivo alunos);
  - disponibilizar dados para o desenvolvimento de estudos e análises na área.
- 

## Características

O Censo da Educação Superior é realizado e disponibilizado anualmente pelo Inep e apresenta dados sobre o ensino superior do país. O conjunto de dados aqui disponibilizado contém informações sobre os alunos de graduação de todas as Instituições de Ensino Superior no Brasil como: raça, sexo, faixa etária, nacionalidade, curso, situação no curso, tipo de ingresso, apoios recebidos, deficiências, entre outros.

Os dados estão disponíveis no site do SoU\_Ciência, no painel da Expansão do Ensino Superior Privado (<https://souciencia.unifesp.br/eesp>), e estão em formato CSV delimitados por ponto e vírgula (;). Os arquivos encontram-se compactados (em formato .zip) e devem ser descompactados no seu disco rígido.

## Registros

A tabela 1 apresenta o número de registros em cada ano do *Censo da Educação Superior - Alunos* de 2009 a 2019.

Tabela 1 - Número de registros por ano do arquivo *SoU\_censo\_aluno*

Ano	Registros	Tamanho do arquivo csv (Kb)
2009	6.982.018	2.066.191.777
2010	8.337.219	2.473.008.178
2011	8.961.724	2.656.592.417
2012	9.565.483	2.835.410.326
2013	9.929.289	2.946.321.114
2014	10.793.935	3.202.243.626
2015	11.187.296	3.402.157.401
2016	11.449.222	3.480.667.135
2017	11.589.194	3.525.633.466
2018	12.043.993	3.541.735.064
2019	12.350.832	3.767.748.756
Total	113.190.205	33.897.709.260

*Esses dados são disponibilizados em arquivos separados por ano com extensão .csv, sendo o tamanho total ocupado de 33.897.709,260 Mb.*

O Dicionário de Dados *SoU\_censo\_aluno*, disponibilizado no arquivo compactado (.zip), contém, em formato .ods (OpenDocument Format), uma descrição detalhada de cada variável, bem como um campo de observação que informa alguma particularidade dos dados em algum ano. Há também duas abas auxiliares com a descrição do código e nome dos municípios, das unidades de federação e das regiões do Brasil, que seguem a classificação do Instituto Brasileiro de Geografia e Estatística (IBGE).

### **Observações:**

Importante informar que esses dados estão em constante processo de atualização e aperfeiçoamento pela equipe do SoU\_Ciência. A versão disponibilizada no repositório de dados trata-se da versão mais atualizada. Para consultar o versionamento acesse o arquivo MD5 para conferência do dataset.

Data da publicação: 12/07/2023  
Data da atualização: 12/07/2023

### **Contato:**

Agradecemos todas as críticas e sugestões recebidas de pesquisadores e usuários que estão colaborando para a melhoria na qualidade da prestação das informações.

Qualquer sugestão ou dúvida pode ser encaminhada para o email [souciencia@unifesp.br](mailto:souciencia@unifesp.br).