**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL**  
**COLETA PREPARAÇÃO E ANÁLISE DE DADOS**

Alexya Ungaratti, Filipe Roman, Lorenzo Schavinski e Rennan Wictor

**TRABALHO FINAL EXTENSIONISTA**  
PRIMEIRA ENTREGA PARCIAL

Porto Alegre, RS  
2025

**1) Pergunta a ser respondida:**

Qual o desempenho das diferentes faculdades do Sul do brasil em qualificação dos seus cursos superiores da área da saúde (medicina, nutrição etc.) com base na nota do ENADE da avaliação mais recente destes cursos.

**2) Bases de Dados Escolhidas:**

* Instituições de Educação superior no brasil
* Cursos de Graduação no Brasil
* Avaliação ENADE

**2.1) Breve descrição das bases de dados escolhidas:**  
  
1-Instituições de Educação superior no Brasil:  
Dataset contendo as informações das diferentes instituições de ensino superior do brasil. Possui informações como: código da Instituição de Educação Superior (IES); nome da IES; sigla; categoria da IES; IES comunitária (Sim/Não); IES confessional (Sim/Não); IES filantrópica; organização.  
  
2-Cursos de Graduação no Brasil:  
Lista de todos os cursos de graduação ofertados pelas IES do Brasil, contém informações como: código da Instituição de Educação Superior (IES); nome da IES; categoria da IES; organização acadêmica; código do curso; nome do curso; grau; área OCDE; modalidade de ensino (presencial ou EaD); situação do curso (ativo ou inativo); vagas autorizadas; carga horária; segmentadas por código do município (IBGE); município; UF; região.  
  
3-Microdados do ENADE :  
Base de dados contento informações a respeito de estudantes de diferentes cursos e faculdades e seus desempenhos em uma prova para avaliar as devidas competências nos conteúdos de seus cursos de graduação, categorizada por IES, curso etc.  
  
  
**3) Justificativa:**   
Acreditamos que a pergunta de pesquisa em questão a respeito da avaliação de diferentes cursos da área da saúde, pode ser respondida. Por meio do uso dos datasets acima partimos do pressuposto que as informações que precisamos são e entre parênteses o/os dataset(s) que contém essa informação:  
-IES do brasil e seus nomes (1);  
-Cursos de saúde ofertados por essas IES que realizam o ENADE (2);  
-Nota do ENADE dos cursos de saúde de cada faculdade (3);

Com acesso a esses dados oferecidos pelos datasets escolhidos, acreditamos ser possível chegar a uma resposta para a pergunta desenvolvida, afinal, todas informações necessárias estão contidas dentro deles e precisam apenas serem extraídas e filtradas.  
Ademais, os datasets 1 e 2 são os grandes responsáveis pela união desses dados por conterem as informações das IES juntamente ligadas a seu código, dado em comum em todos os datesets.