

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

ALISSON DA SILVA VIEIRA

**ENRIQUECIMENTO DO CENSO DA EDUCAÇÃO SUPERIOR PARA O ESTUDO
DAS CARACTERÍSTICAS DOS CURSOS DE GRADUAÇÃO DO BRASIL**

CAMPO MOURÃO

2023

ALISSON DA SILVA VIEIRA

**ENRIQUECIMENTO DO CENSO DA EDUCAÇÃO SUPERIOR PARA O ESTUDO
DAS CARACTERÍSTICAS DOS CURSOS DE GRADUAÇÃO DO BRASIL**

**Enrichment of the Higher Education Census to study the characteristics of
undergraduate courses in Brazil**

Trabalho de conclusão de curso de graduação
apresentado como requisito parcial para a obtenção
do título de Bacharel em Ciência da Computação,
Universidade Tecnológica Federal do Paraná
(UTFPR).

Orientador(a): Prof. Dr. Andre Luis Schwerz

CAMPO MOURÃO

2023

ALISSON DA SILVA VIEIRA

**ENRIQUECIMENTO DO CENSO DA EDUCAÇÃO SUPERIOR PARA O ESTUDO
DAS CARACTERÍSTICAS DOS CURSOS DE GRADUAÇÃO DO BRASIL**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do título
de Bacharel em Ciência da Computação da
Universidade Tecnológica Federal do Paraná
(UTFPR).

Data de aprovação: 20/junho/2023

Andre Luis Schwerz
Doutorado
<http://lattes.cnpq.br/4954414332524750>
Universidade Tecnológica Federal do Paraná

Eduardo Henrique Monteiro Pena
Doutorado
<http://lattes.cnpq.br/3795697672633760>
Universidade Tecnológica Federal do Paraná

Rafael Liberato Roberto
Doutorado
<http://lattes.cnpq.br/7833556006473564>
Universidade Tecnológica Federal do Paraná

CAMPO MOURÃO

2023

RESUMO

VIEIRA, Alisson da Silva. **Enriquecimento do Censo da Educação Superior para o estudo das características dos cursos de graduação do Brasil**. 2023. Trabalho de conclusão de curso (Bacharelado em Ciência da Computação) – Universidade Tecnológica Federal do Paraná, Campo Mourão, 2023.

No Brasil, houve uma grande expansão das Instituições de Ensino Superior (IESs) a partir da década de 90. No último Censo da Educação Superior realizado pelo INEP, foram registradas 2457 IESs, 41953 cursos de ensino superior e 3,680 milhões de matrículas. Os dados do Censo da Educação Superior são abertos e publicados pelo Inep anualmente. Eles possuem um imenso valor uma vez que retratam o cenário acadêmico brasileiro e podem ser estudados para evidenciar tendências e padrões. Entretanto, acredita-se que os indicadores e as métricas do cenário acadêmico podem ser influenciados por outros fatores externos às IESs como o cenário social, demográfico e econômico. Esta pesquisa propõe o enriquecimento dos dados do Censo da Educação Superior com o intuito de realizar análises sobre os mais diversos indicadores sociais, econômicos e demográficos do país obtidos a partir de bases de dados abertas. Para o cenário econômico, foram levantados dados do CAGED e, para o cenário social, foram levantados dados provenientes do IBGE. Durante o processo do enriquecimento, foi realizada uma filtragem pelos cursos de rotulados como Ciência da Computação no Censo, pelas áreas econômicas de computação no CAGED e pelos municípios que possuem IESs que ofertam cursos de Ciência da Computação no IBGE. Com os dados enriquecidos, foi realizada uma análise descritiva do conjunto e discutido uma série de correlações que não eram possíveis de serem analisadas de forma isolada em trabalhos anteriores. Entre os resultados, destaca-se que os cursos de computação apresentam a segunda maior média de evasão dos cursos do ensino superior. A análise do conjunto propiciou também identificar correlações fortes entre métricas e indicadores como, por exemplo, média salarial e evasão. Para cada uma delas, discutiu-se seus possíveis fatores e sugeriu-se hipóteses que podem ser estudadas em trabalhos futuros.

Palavras-chave: censo da educação superior; enriquecimento de dados; métricas do ensino superior; caged; ibge.

() Não autorizo a disponibilização de endereço de correio eletrônico para contato.

(X) Autorizo a disponibilização do seguinte correio eletrônico para contato:

alisson.v3@hotmail.com

ABSTRACT

VIEIRA, Alisson da Silva. **Enrichment of the Higher Education Census to study the characteristics of undergraduate courses in Brazil.** 2023. Trabalho de conclusão de curso (Bacharelado em Ciência da Computação) – Universidade Tecnológica Federal do Paraná, Campo Mourão, 2023. Título original: Enriquecimento do Censo da Educação Superior para o estudo das características dos cursos de graduação do Brasil.

In Brazil, there has been a significant expansion of Higher Education Institutions (IESs) since the 1990s. In the latest Higher Education Census conducted by INEP, 2,457 IESs, 41,953 higher education courses, and 3.68 million enrollments were recorded. The data from the Higher Education Census are open and published annually by INEP. They hold a huge value as they illustrate the Brazilian academic scenario and can be studied to highlight trends and patterns. However, it is believed that the indicators and metrics of the academic scenario can be influenced by external factors to IESs, such as the social, demographic, and economic scenario. This research proposes the enrichment of the data from the Higher Education Census in order to conduct analyses on various social, economic, and demographic indicators of the country obtained from open databases. For the economic scenario, data from the General Registry of Employees and Unemployed (CAGED) were collected, and for the social scenario, data from the Brazilian Institute of Geography and Statistics (IBGE) were collected. During the enrichment process, filtering was applied to the Computer Science courses from the Census, the computing-related economic areas from CAGED, and the cities with IESs offering Computer Science courses from IBGE. With the enriched data, a descriptive analysis of the dataset was conducted, and a series of correlations that were not possible to analyze in isolation in previous works were discussed. Among the results, it brings out that computer science courses have the second highest dropout rate among higher education courses. The analysis of the dataset enabled the identification of several strong correlations between metrics and indicators, such as average salary and dropout rate. For each of them, their possible factors were discussed, and hypotheses that can be studied in future works were suggested.

Keywords: higher education census; data enrichment; higher education metrics; caged; ibge.