



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# Mineração de Dados Aplicada ao Exame Nacional de Desempenho dos Estudantes (ENADE)

Daniel Erick Sanchez Trujillo

Monografia apresentada como requisito parcial  
para conclusão do Curso de Engenharia da Computação

Orientador  
Prof. Dr. Jan Mendonca Correa

Brasília  
2024



# Dedicatória

Dedico este trabalho aos meus pais, cujo amor, sacrifício e apoio incondicional foram meu maior suporte durante a minha jornada acadêmica, e igualmente aos meus amigos, que compartilharam comigo os desafios vivenciados durante o curso.

# Agradecimentos

Primeiramente, expresso minha profunda gratidão a Deus, cuja presença foi constante e cuja força me guiou em cada passo desta longa jornada acadêmica. Aos meus pais Roberto e Cintia, dedico não só este trabalho, mas todos os frutos que esta jornada possa render. Estendo meu agradecimento ao meu irmão Marcelo, à minha tia Ines, à minha avó Celina e em memória ao meu avô Marcos, que sempre será uma parte essencial de quem sou. Agradeço imensamente aos meus amigos, Aline, Antony, Iago e Jhonatan, vocês compartilharam comigo não apenas o conhecimento, mas risadas, apoio e momentos inesquecíveis que amenizaram as pressões do percurso acadêmico. Um agradecimento aos professores da Universidade de Brasília, que contribuíram imensamente para o meu crescimento intelectual e profissional. Por último gostaria de agradecer meu orientador Jan Mendonça, por sua orientação, paciência e conhecimento inestimável, que foram cruciais na realização deste trabalho.

# Resumo

O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) fornece acesso público a uma ampla gama de dados obtidos por meio de suas pesquisas, avaliações e exames educacionais. Dentre esses, os dados relativos ao Exame Nacional de Desempenho dos Estudantes (ENADE) se destacam. Estes dados abrangem um leque variado de informações acerca dos estudantes que participaram do exame. Este estudo se concentrou na aplicação de métodos de visualização, análise e mineração de dados utilizando ferramentas avançadas como Python, Pandas e Jupyter Notebook. A análise do ENADE 2021 foi realizada com ênfase nos cursos de computação, com um foco especial no curso de Ciência da Computação (Bacharelado) e Computação (Licenciatura) nos alunos da Universidade de Brasília (UnB). Foram realizadas várias análises nos dados, incluindo correlação entre variáveis, classificação das notas com base em diferentes atributos, e exame detalhado das respostas dos alunos às questões do exame. O propósito dessas análises foi fornecer uma compreensão aprofundada do perfil dos estudantes de ensino superior no Brasil. Este trabalho proporciona uma compreensão aprofundada dos dados do ENADE e do desempenho dos alunos, incorporando tabelas e gráficos para análises detalhadas e identificação de fatores que influenciam o desempenho dos alunos no exame.

**Palavras-chave:** ENADE, Ciência da Computação, UnB, Mineração de Dados

# Abstract

The National Institute for Educational Studies and Research Anísio Teixeira (INEP) provides public access to a wide range of data obtained through its research, assessments, and educational exams. Among these, the data related to the National Exam of Student Performance (ENADE) stand out. These data encompass a varied range of information about the students who participated in the exam. This study focused on applying methods of visualization, analysis, and data mining using advanced tools such as Python, Pandas, and Jupyter Notebook. The analysis of ENADE 2021 was carried out with an emphasis on computing courses, with a special focus on the Computer Science (Bachelor's) and Computing (Teaching degree) and on the students of the University of Brasília (UnB). Several analyses were performed on the data, including the correlation between variables, classification of grades based on different attributes, and detailed examination of the students' responses to the exam questions. The purpose of these analyses was to gain a comprehensive understanding of the profile of higher education students in Brazil. This work provides a deep understanding of ENADE data and student performance, incorporating tables and graphs for detailed analysis and identification of factors influencing students' performance in the exam.

**Keywords:** ENADE, Computer Science, UnB, Data Mining

# Sumário

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Introdução</b>   | <b>1</b> |
| 1.1      | Problema . . . . .  | 1        |
| 1.2      | Motivação . . . . .   | 2        |
| 1.3      | Objetivo . . . . .  | 2        |
| 1.3.1    | Objetivos Específicos . . . . .                               | 2        |
| 1.4      | Metodologia . . . . .   | 2        |
| 1.5      | Estrutura dos capítulos . . . . .                             | 3        |
| <b>2</b> | <b>Fundamentação teórica</b>                                  | <b>4</b> |
| 2.1      | Exame Nacional de Desempenho dos Estudantes (ENADE) . . . . . | 4        |
| 2.1.1    | Cálculo da nota final do estudante . . . . .                  | 5        |
| 2.1.2    | Cálculo da nota da formação geral . . . . .                   | 5        |
| 2.1.3    | Cálculo da nota de conhecimentos específicos . . . . .        | 6        |
| 2.2      | Principais conceitos . . . . .                                | 7        |
| 2.2.1    | Dados . . . . .   | 7        |
| 2.2.2    | Informação . . . . .  | 7        |
| 2.2.3    | Mineração de dados . . . . .                                  | 8        |
| 2.2.4    | Tratamento de dados . . . . .                                 | 8        |
| 2.2.5    | Análise de dados . . . . .                                    | 8        |
| 2.3      | Modelo CRISP-DM . . . . .                                     | 8        |
| 2.3.1    | Entendimento do negócio . . . . .                             | 9        |
| 2.3.2    | Entendimento dos dados . . . . .                              | 9        |
| 2.3.3    | Preparação dos dados . . . . .                                | 9        |
| 2.3.4    | Modelagem . . . . .   | 10       |
| 2.3.5    | Avaliação . . . . .   | 10       |
| 2.3.6    | Implantação . . . . .   | 10       |
| 2.4      | Ferramentas utilizadas . . . . .                              | 10       |
| 2.4.1    | Python e Jupyter Notebook . . . . .                           | 10       |
| 2.4.2    | Bibliotecas utilizadas . . . . .                              | 11       |

|          |   |           |
|----------|---|-----------|
| 2.5      | Impacto da LGPD . . . . .   | 11        |
| <b>3</b> | <b>Preparação dos dados</b>   | <b>13</b> |
| 3.1      | Coleta de dados . . . . .   | 13        |
| 3.1.1    | Microdados do ENADE . . . . .   | 14        |
| 3.1.2    | Cursos de Graduação do Brasil . . . . .                                       | 14        |
| 3.2      | Tratamento dos dados . . . . .  | 15        |
| 3.2.1    | Tratamento dos vetores . . . . .  | 15        |
| 3.2.2    | Arquivo auxiliar para mapeamento . . . . .                                    | 16        |
| 3.2.3    | Formatação do dados dos vetores . . . . .                                     | 17        |
| 3.2.4    | Formatação dos dados com a percepção da prova . . . . .                       | 17        |
| 3.2.5    | Tratamento dos questionários . . . . .  | 18        |
| <b>4</b> | <b>Análise dos dados</b>  | <b>20</b> |
| 4.1      | Análise geral dos dados . . . . .   | 20        |
| 4.1.1    | Heatmap dos dados do questionário do estudante . . . . .                      | 21        |
| 4.1.2    | Análise das notas . . . . .   | 29        |
| 4.2      | Análise das notas do curso Ciência da Computação . . . . .                    | 33        |
| 4.2.1    | Nota bruta na formação geral . . . . .  | 33        |
| 4.2.2    | Nota bruta nos conhecimentos específicos . . . . .                            | 34        |
| 4.2.3    | Nota bruta da prova . . . . .   | 34        |
| 4.2.4    | Boxplot das notas do curso Ciência da Computação . . . . .                    | 35        |
| 4.3      | Análise dos vetores de notas . . . . .  | 38        |
| 4.3.1    | Análise de número total de alunos . . . . .                                   | 39        |
| 4.3.2    | Análise da parte objetiva na formação geral . . . . .                         | 41        |
| 4.3.3    | Análise da parte objetiva de conhecimentos específicos . . . . .              | 45        |
| 4.3.4    | Comparação do desempenho dos alunos da UnB com outras universidades . . . . . | 55        |
| 4.4      | Análise do questionário do estudante . . . . .                                | 60        |
| 4.4.1    | Análise de gênero nos cursos de Computação da UnB . . . . .                   | 61        |
| 4.4.2    | Análise do estado civil nos cursos de Computação da UnB . . . . .             | 61        |
| 4.4.3    | Análise da cor/raça nos cursos de Computação da UnB . . . . .                 | 62        |
| 4.4.4    | Análise da nacionalidade nos cursos de Computação da UnB . . . . .            | 63        |
| 4.4.5    | Análise do ambiente familiar nos cursos de Computação da UnB . . . . .        | 64        |
| 4.4.6    | Análise do perfil educacional nos cursos de Computação da UnB . . . . .       | 69        |
| <b>5</b> | <b>Conclusões</b>   | <b>74</b> |
|          | <b>Referências</b>  | <b>76</b> |



# Lista de Figuras

|      |  |    |
|------|--|----|
| 2.1  | Diagrama CRISP-DM (baseado em [1]). . . . .  | 9  |
| 4.1  | Heatmap do sexo dos participantes. . . . .   | 22 |
| 4.2  | Heatmap da idade dos participantes. . . . .  | 23 |
| 4.3  | Heatmap do estado civil dos participantes. . . . .   | 24 |
| 4.4  | Heatmap da cor/raça dos participantes. . . . .   | 25 |
| 4.5  | Heatmap da renda familiar mensal dos participantes. . . . .                                | 26 |
| 4.6  | Heatmap de alunos por estado dos participantes. . . . .                                    | 28 |
| 4.7  | Heatmap do tipo de escola dos participantes. . . . .                                       | 30 |
| 4.8  | Boxplot de comparação de NT_FG. . . . .  | 36 |
| 4.9  | Boxplot de comparação de NT_CE. . . . .  | 37 |
| 4.10 | Boxplot de comparação de NT_GER. . . . .   | 38 |
| 4.11 | Questão 11 para Ciência da Computação (Bacharelado). . . . .                               | 49 |
| 4.12 | Questão 26 para Ciência da Computação (Bacharelado). . . . .                               | 50 |
| 4.13 | Questão 9 para Ciência da Computação (Licenciatura). . . . .                               | 54 |
| 4.14 | Gráfico do sexo dos alunos de Ciência Da Computação. . . . .                               | 61 |
| 4.15 | Gráfico do estado civil dos alunos de Ciência Da Computação. . . . .                       | 62 |
| 4.16 | Gráfico da cor/raça dos alunos de Ciência Da Computação. . . . .                           | 63 |
| 4.17 | Gráfico da nacionalidade dos alunos de Ciência Da Computação. . . . .                      | 64 |
| 4.18 | Gráfico da escolaridade do pai dos alunos de Ciência da Computação. . . . .                | 65 |
| 4.19 | Gráfico da escolaridade da mãe dos alunos de Ciência da Computação. . . . .                | 65 |
| 4.20 | Gráfico da moradia dos alunos de Ciência da Computação. . . . .                            | 66 |
| 4.21 | Gráfico da quantidade de membros que moram com os alunos de Ciência Da Computação. . . . . | 67 |
| 4.22 | Gráfico da renda familiar dos alunos de Ciência Da Computação. . . . .                     | 67 |
| 4.23 | Gráfico da situação de trabalho dos alunos de Ciência da Computação. . . . .               | 68 |
| 4.24 | Gráfico da situação financeira dos alunos de Ciência da Computação. . . . .                | 68 |
| 4.25 | Gráfico da unidade federativa onde o aluno concluiu o ensino médio (Bacharelado). . . . .  | 69 |

|  |    |
|--|----|
| 4.26 Gráfico da unidade federativa onde o aluno concluiu o ensino médio (Licenciatura). . . . .  | 70 |
| 4.27 Gráfico do tipo de escola dos alunos de Ciência da Computação. . . . .                      | 70 |
| 4.28 Gráfico da modalidade de ensino médio dos alunos de Ciência da Computação. . . . .          | 71 |
| 4.29 Gráfico da quantidade de livros lidos pelos alunos de Ciência da Computação. . . . .        | 71 |
| 4.30 Gráfico da quantidade de livros lidos pelos alunos de Ciência da Computação. . . . .        | 72 |
| 4.31 Gráfico do aprendizado de idiomas estrangeiros nos alunos de Ciência da Computação. . . . . | 73 |

# Lista de Tabelas

|      |   |    |
|------|---|----|
| 4.1  | Tabela com os cursos participantes do ENADE . . . . .                           | 21 |
| 4.2  | Tabela das medianas das notas dos participantes do ENADE . . . . .              | 32 |
| 4.3  | Notas da formação geral . . . . .   | 33 |
| 4.4  | Notas dos conhecimentos específicos. . . . .                                    | 34 |
| 4.5  | Nota bruta da prova. . . . .  | 35 |
| 4.6  | Faculdades selecionadas. . . . .  | 39 |
| 4.7  | Faculdades selecionadas para Ciência da Computação (Bacharelado). . . . .       | 40 |
| 4.8  | Faculdades selecionadas para Ciência da Computação (Licenciatura). . . . .      | 41 |
| 4.9  | Percentual de acertos nas questões objetivas (FG) 1 a 4. . . . .                | 42 |
| 4.10 | Percentual de acertos nas questões objetivas (FG) 5 a 8. . . . .                | 42 |
| 4.11 | Comparação de desempenho (FG) questões objetivas 1 a 8. . . . .                 | 43 |
| 4.12 | Respostas da UnB para as questões objetivas (FG) Bacharelado. . . . .           | 44 |
| 4.13 | Percentual de acertos nas questões objetivas 1 a 4. . . . .                     | 44 |
| 4.14 | Percentual de acertos nas questões objetivas 5 a 8. . . . .                     | 44 |
| 4.15 | Comparação de desempenho (FG) questões objetivas 1 a 8. . . . .                 | 45 |
| 4.16 | Respostas da UnB para as questões objetivas (FG) Licenciatura. . . . .          | 45 |
| 4.17 | Percentual de acertos nas questões (CE) 1, 2, 3, 6, 7. . . . .                  | 46 |
| 4.18 | Percentual de acertos nas questões (CE) 8, 10, 11, 12, 14. . . . .              | 46 |
| 4.19 | Percentual de acertos nas questões (CE) 15, 16, 18, 19, 20. . . . .             | 47 |
| 4.20 | Percentual de acertos nas questões (CE) 22, 23, 26, 27. . . . .                 | 47 |
| 4.21 | Comparação de desempenho (CE) questões objetivas 1 a 27. . . . .                | 48 |
| 4.22 | Respostas da UnB para as questões objetivas (CE) Bacharelado. . . . .           | 51 |
| 4.23 | Percentual de acertos nas questões 1 a 6. . . . .                               | 52 |
| 4.24 | Percentual de acertos nas questões 7 a 12. . . . .                              | 52 |
| 4.25 | Percentual de acertos nas questões 14 a 21. . . . .                             | 52 |
| 4.26 | Percentual de acertos nas questões 22 a 27. . . . .                             | 52 |
| 4.27 | Comparação de desempenho (CE) questões objetivas 1 a 27 - Licenciatura. . . . . | 53 |
| 4.28 | Respostas da UnB para as questões objetivas (CE) Licenciatura. . . . .          | 55 |

|      |   |    |
|------|---|----|
| 4.29 | Comparação do desempenho (FG) da UnB - Ciência da Computação - Bacharelado com outras instituições. . . . . | 56 |
| 4.30 | Comparação do desempenho (CE) da UnB - Ciência da Computação - Bacharelado com outras instituições. . . . . | 57 |
| 4.31 | Comparação do desempenho (FG) da UnB - Computação - Licenciatura com outras instituições. . . . .           | 58 |
| 4.32 | Comparação do desempenho (CE) da UnB - Computação - Licenciatura com outras instituições. . . . .           | 59 |
| 4.33 | Disciplinas da UnB. . . . .   | 60 |

# Lista de Abreviaturas e Siglas

**CNCST** Catálogo Nacional de Cursos Superiores de Tecnologia.

**CNPq** Conselho Nacional de Desenvolvimento Científico e Tecnológico.

**ENADE** Exame Nacional de Desempenho dos Estudantes.

**INEP** Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.

**LGPD** Lei Geral de Proteção de Dados Pessoais.

**MEC** Ministério da Educação.

**SINAES** Sistema Nacional de Avaliação da Educação Superior.

**UnB** Universidade de Brasília.

# Capítulo 1

## Introdução

Com o aumento exponencial na geração de dados a cada dia [2], torna-se um desafio crescente e uma necessidade de extrair e analisar essas informações valiosas. Neste contexto, o tratamento e a análise de dados emergem como ferramentas indispensáveis para a extração e interpretação das informações relevantes na área em avaliação.

No contexto brasileiro, o Exame Nacional de Desempenho dos Estudantes (ENADE) [3] é uma iniciativa do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) que avalia estudantes concluintes de cursos de graduação, analisando seus conhecimentos, habilidades e competências adquiridas durante a formação. Esse exame não apenas mensura o desempenho dos estudantes, mas também serve como base para os indicadores de qualidade do ensino superior no país. Os microdados disponíveis no ENADE contêm uma riqueza de informações sobre os alunos, incluindo dados demográficos como etnia, renda, tipo de ensino médio, notas obtidas, e o curso correspondente, entre outros. Muitas dessas informações podem passar despercebidas aos olhos humanos devido ao seu grande volume.

Ao realizar o tratamento e a análise de dados nesses microdados, é possível descobrir padrões relacionados aos atributos e perfis dos estudantes, bem como identificar características que possam influenciar suas notas. Esse processo revela informações valiosas que podem ser fundamentais para melhorar políticas educacionais e práticas pedagógicas, contribuindo significativamente para o avanço do ensino superior no país.

### 1.1 Problema

Os microdados do ENADE reúnem informações detalhadas sobre cerca de 489 mil estudantes divididos em 30 cursos que participaram do exame. No entanto, a análise direta desses dados brutos pode ser desafiadora para a extração de interpretações significativas. A implementação de técnicas de tratamento de dados nesse vasto conjunto de informações

tem como principais objetivos: a criação de conhecimento acessível e compreensível para os usuários e a identificação de padrões complexos e não óbvios em meio a essa extensa quantidade de dados.

## 1.2 Motivação

A análise da qualidade do ensino superior no Brasil representa uma questão de grande relevância. Contudo, a maneira como os dados estão formatados inicialmente não facilita uma avaliação direta dos microdados. Depois de fazer o tratamento desses dados e utilizar algoritmos de mineração de dados, é possível facilitar a extração de informações valiosas e tornar essa tarefa mais eficaz e direta.

## 1.3 Objetivo

O principal objetivo deste trabalho é verificar o desempenho dos alunos da Universidade de Brasília (UnB) nos cursos em relação às questões individuais, comparando-o com o desempenho de outros alunos do mesmo curso utilizando técnicas de mineração de dados.

### 1.3.1 Objetivos Específicos

Para atingir o objetivo geral deste trabalho, foram estabelecidos os seguintes objetivos:

1. Extrair os microdados do ENADE.
2. Compreender a estrutura e o conteúdo dos microdados extraídos.
3. Colocar os dados em um formato apropriado para facilitar a posterior análise na mineração de dados.
4. Utilizar técnicas e algoritmos de mineração de dados no conjunto de dados tratado.
5. Examinar e interpretar os resultados obtidos.

## 1.4 Metodologia

A metodologia adotada neste trabalho envolveu a extração criteriosa e a curadoria dos microdados do ENADE por meio de *scripts* em Python[4]. Foram realizadas operações de mapeamento e transformação de dados, o que facilitou a interpretação e a análise subsequente. Esta abordagem possibilitou uma preparação eficiente dos microdados, essencial para a aplicação de algoritmos de mineração de dados e interpretações relevantes sobre o desempenho dos estudantes.

## 1.5 Estrutura dos capítulos

- Capítulo 1: Introdução

O capítulo introduz o tema central do trabalho, abordando o problema a ser investigado, detalha a motivação para a escolha do tema e define os objetivos específicos que o estudo pretende alcançar e uma breve introdução à metodologia utilizada.

- Capítulo 2: Fundamentação teórica

O capítulo apresenta uma revisão das ferramentas utilizadas para a realização do estudo, contextualiza o tema dentro do campo acadêmico e fornece as bases teóricas necessárias para esta análise dos dados.

- Capítulo 3: Preparação dos dados

O capítulo descreve detalhadamente o processo de mineração de dados, incluindo a coleta, preparação e tratamento dos dados, detalha as técnicas e procedimentos adotados para tornar os dados aptos para análise.

- Capítulo 4: Análise dos dados

O capítulo foca na análise dos dados tratados, utilizando técnicas de mineração de dados, gerando gráficos e tabelas para explorar as relações descobertas nos dados e discute o significado desses resultados no contexto do trabalho, focando a análise nos alunos da Universidade de Brasília.

- Capítulo 5: Conclusões

O capítulo é dedicado a consolidar as descobertas cruciais emergidas ao longo do trabalho, são sintetizados os principais *insights* e resultados obtidos.



# Capítulo 2

## Fundamentação teórica

Este capítulo fornece uma revisão abrangente de conceitos teóricos, das ferramentas e métodos essenciais que foram aplicados para a realização do trabalho.

### 2.1 Exame Nacional de Desempenho dos Estudantes (ENADE)

O ENADE [3] representa um componente crítico na avaliação da qualidade do ensino superior no Brasil. Instituído em 2004 pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), o ENADE é parte integrante do Sistema Nacional de Avaliação da Educação Superior (SINAES). Tem como objetivo principal avaliar o desempenho dos estudantes concluintes de cursos de graduação em relação a um conjunto diversificado de competências e habilidades. Este exame se concentra na verificação do domínio dos conteúdos necessários para a formação profissional e geral.

O exame é obrigatório para estudantes ingressantes e concluintes dos cursos de graduação, sendo sua participação um pré-requisito para a obtenção do diploma. Os cursos avaliados pelo ENADE são determinados pelo ciclo avaliativo, que define as áreas de conhecimento e os cursos a elas vinculados. As áreas para cursos de bacharelado e licenciatura são baseadas na tabela de áreas do conhecimento do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), e os eixos tecnológicos seguem o Catálogo Nacional de Cursos Superiores de Tecnologia (CNCST) do Ministério da Educação. A prova do ENADE tem os seguintes dois componentes:

- **Formação geral (FG):** Composto de 10 questões, sendo 8 de múltipla escolha e 2 discursivas.
- **Conhecimentos específicos (CE):** Composto de 30 questões, sendo 27 de múltipla escolha e 3 discursivas.

### 2.1.1 Cálculo da nota final do estudante

A nota final do estudante ( $NT_{GER}$ ) assume valor de 0 a 100, arredondada com uma casa decimal [5]. É calculada como uma soma ponderada das notas da componente de formação geral ( $NT_{FG}$ , com peso 25%) e da componente de conhecimentos específicos ( $NT_{CE}$ , com peso 75%), isto é:

A fórmula para o cálculo é:

$$NT_{GER} = 0,25 \cdot NT_{FG} + 0,75 \cdot NT_{CE} \quad (2.1)$$

em que:

- $NT_{GER}$  é a nota final do estudante (de 0 a 100);
- $NT_{FG}$  é a nota da formação geral;
- $NT_{CE}$  é a nota dos conhecimentos específicos.

### 2.1.2 Cálculo da nota da formação geral

A nota da formação geral do estudante ( $NT_{FG}$ ) assume valor de 0 a 100. A composição da nota de formação geral no ENADE é realizada por meio de uma média ponderada entre as avaliações objetivas e discursivas. Os pesos atribuídos para as questões objetivas e discursivas são de 60% e 40%, respectivamente, isto é:

$$NT_{FG} = 0,60 \cdot NT_{Obj\_FG} + 0,40 \cdot NT_{Dis\_FG} \quad (2.2)$$

em que:

- $NT_{FG}$  é a nota correspondente ao componente de formação geral (de 0 a 100);
- $NT_{Obj\_FG}$  é a nota referente às questões objetivas do componente de formação geral (de 0 a 100);
- $NT_{Dis\_FG}$  é a nota referente às questões discursivas do componente de formação geral (de 0 a 100).

Para calcular a nota bruta das questões objetivas de FG ( $NT_{Obj\_FG}$ ), são 8 questões objetivas de múltipla escolha com peso idêntico, perfazendo 100%. Assim a nota é a proporção de acertos dessas questões em relação ao número de questões válidas.

A nota das questões discursivas de formação geral ( $NT_{Dis\_FG}$ ) é a média simples das notas das duas questões discursivas:

$$NT_{Dis\_FG} = \frac{NT_{FG\_D1} + NT_{FG\_D2}}{2} \quad (2.3)$$

em que:

- $NT_{Dis\_FG}$  é a média das notas das questões discursivas do componente de formação geral;
- $NT_{FG\_D1}$  é a nota da 1ª questão discursiva do componente de formação geral (de 0 a 100);
- $NT_{FG\_D2}$  é a nota da 2ª questão discursiva do componente de formação geral (de 0 a 100).

Para as duas questões discursivas ( $D_1$  e  $D_2$ ), são considerados os seguintes pesos, para o conteúdo (peso de 80%) e aspectos referentes à língua portuguesa (peso de 20%) que os parâmetros da língua portuguesa são distribuídos da seguinte maneira: aspectos ortográficos (30%), aspectos textuais (20%) e aspectos morfosintáticos e vocabulares (50%). Desta forma, a nota da questão discursiva 1 é assim calculada:

$$NT_{FG\_D1} = 0,20 \cdot NT_{FG\_D1\_PT} + 0,80 \cdot NT_{FG\_D1\_CT} \quad (2.4)$$

em que:

- $NT_{FG\_D1}$  é a nota da 1ª questão discursiva do componente de formação geral (de 0 a 100);
- $NT_{FG\_D1\_PT}$  é a nota dos aspectos referentes à língua portuguesa da 1ª questão discursiva do componente de formação geral;
- $NT_{FG\_D1\_CT}$  é a nota dos aspectos referentes ao conteúdo da 1ª questão discursiva do componente de formação geral.

De igual forma é calculada a nota da questão  $D_2$ .

### 2.1.3 Cálculo da nota de conhecimentos específicos

A nota de conhecimentos específicos do estudante ( $NT_{CE}$ ) assume valor de 0 a 100. Onde é a média ponderada das duas notas, objetivas e discursivas com pesos de 85% e 15%, respectivamente.

$$NT_{CE} = 0,85 \cdot NT_{Obj\_CE} + 0,15 \cdot NT_{Dis\_CE} \quad (2.5)$$

em que:

- $NT_{CE}$  é a nota correspondente ao componente de conhecimentos específicos;
- $NT_{Obj\_CE}$  é a nota referente às questões objetivas do componente de conhecimentos específicos (de 0 a 100);
- $NT_{Dis\_CE}$  é a média das notas das questões discursivas do componente de conhecimentos específicos (de 0 a 100).

Para as 27 questões objetivas de múltipla escolha com peso idêntico, a nota das questões de conhecimentos específicos ( $NT_{Obj\_CE}$ ) é a proporção de acertos dessas questões em relação ao número de questões válidas.

Finalmente para as 3 questões discursivas nas quais 100% da nota refere-se ao conteúdo de conhecimentos específicos ( $NT_{Dis\_CE}$ ) é a média simples das notas dessas 3 questões.

$$NT_{Dis\_CE} = \frac{NT_{CE\_D1} + NT_{CE\_D2} + NT_{CE\_D3}}{3} \quad (2.6)$$

em que:

- $NT_{Dis\_CE}$  é a média das notas das questões discursivas do componente de conhecimentos específicos;
- $NT_{CE\_D1}, NT_{CE\_D2}, NT_{CE\_D3}$  são as notas referentes às questões discursivas do componente conhecimentos específicos (de 0 a 100).

## 2.2 Principais conceitos

Neste trabalho revisaremos vários conceitos fundamentais, estes incluem:

### 2.2.1 Dados

De acordo com KITCHIN e LAURIAULT (2014) [6], “dados” referem-se a qualquer conjunto de informações ou observações, geralmente apresentados em forma numérica, textual ou visual. Eles são a matéria-prima fundamental para análises, modelagens e tomada de decisões.

### 2.2.2 Informação

Segundo DAVENPORT e PRUSAK (1998) [7], dados tornam-se informação quando são processados, organizados ou estruturados de forma a serem interpretáveis e significativos. A informação é descrita como dados que foram conferidos com significado através do contexto, análise e interpretação.

### 2.2.3 Mineração de dados

HAN, KAMBER e PEI (2011) [8], definem a Mineração de dados como o processo de identificação de padrões, correlações e *insights* valiosos em grandes conjuntos de dados, através da utilização de algoritmos e técnicas estatísticas e de aprendizado de máquina. Este método é primordial para descobrir conhecimento útil e não evidente à primeira vista, o que o torna essencial para extrair informações significativas dos microdados do ENADE.

### 2.2.4 Tratamento de dados

Conforme descrito por PYLE (1999) [9], o Tratamento de dados envolve a limpeza e a preparação dos dados para a análise subsequente. Este processo inclui o tratamento de valores ausentes, a normalização e a transformação dos dados, para garantir sua adequação para análise. O tratamento de dados é uma etapa preliminar crucial, assegurando a qualidade, padronização e a precisão dos resultados analíticos.

### 2.2.5 Análise de dados

De acordo com SHMUELI e PATEL (2010) [10], a Análise de dados é um processo central neste trabalho. Eles a definem como a inspeção, transformação e modelagem de dados com o objetivo de descobrir informações úteis. Ao contrário da mineração de dados, a análise de dados frequentemente se concentra em questões e conjuntos de dados mais específicos.

## 2.3 Modelo CRISP-DM

O modelo CRISP-DM [11][12] (*Cross-Industry Standard Process for Data Mining*) é um processo padrão da indústria para mineração de dados que oferece uma abordagem estruturada para planejar e executar projetos de *data mining*. O modelo é amplamente aceito e utilizado devido à sua flexibilidade e adaptabilidade a diferentes tipos de projetos de mineração de dados, o modelo é dividido em seis fases principais representadas na Figura 2.1 que são detalhadas a seguir:

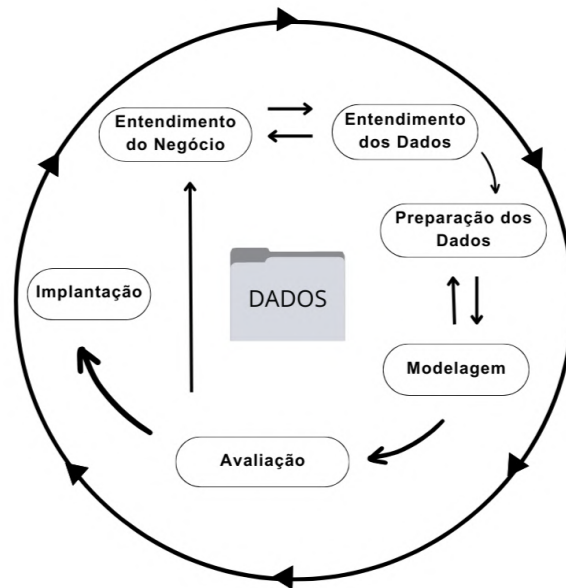


Figura 2.1: Diagrama CRISP-DM (baseado em [1]).

### 2.3.1 Entendimento do negócio

Entendimento do negócio é compreender os objetivos e requisitos do projeto a partir de uma perspectiva de negócios e, em seguida, converter este conhecimento em uma definição de problema de mineração de dados e um plano preliminar para alcançá-lo. Essas atividades incluem identificar os principais fatores que influenciam os objetivos do negócio, avaliar a situação inicial, determinar os objetivos do projeto de mineração de dados e elaborar um plano inicial.

### 2.3.2 Entendimento dos dados

Entendimento dos dados consiste em iniciar a coleta de dados e proceder com as atividades para se familiarizar com os dados, identificar problemas de qualidade dos dados, descobrir insights iniciais. Essa etapa envolve a coleta de dados iniciais, descrição de dados, exploração de dados, e verificação da qualidade dos dados.

### 2.3.3 Preparação dos dados

A preparação dos dados precisa construir o conjunto de dados final a partir dos dados iniciais brutos. As tarefas de preparação dos dados provavelmente serão realizadas várias vezes e não em uma ordem específica. As tarefas dessa fase abrangem a seleção de tabelas,

registros, e atributos, preparar para a limpeza de dados para lidar com dados ausentes ou errôneos e finalmente a construção de novos atributos para melhorar a análise.

### **2.3.4 Modelagem**

Na fase de modelagem, o foco é a seleção, aplicação e avaliação de técnicas de modelagem adequadas que atenda às necessidades do projeto de mineração de dados. Parte integrante desta fase é a escolha da técnica de modelagem, geração de conjunto de testes, construção do modelo, e avaliação do modelo.

### **2.3.5 Avaliação**

Para a fase de avaliação é avaliado o modelo ou modelos para garantir que eles atendam aos objetivos do negócio. É necessário revisar os passos executados para construir o modelo e ter certeza de que o modelo atende aos objetivos do negócio. Consequentemente, isso inclui avaliar os resultados, revisar o processo e determinar os próximos passos.

### **2.3.6 Implantação**

Para a implantação do modelo de mineração de dados, os resultados do projeto de mineração de dados são integrados aos sistemas de operação para efetivamente gerar valor. Nesse processo, é fundamental planejar a implantação, monitoramento e manutenção, produzir um relatório final e revisar o projeto.

## **2.4 Ferramentas utilizadas**

Para o tratamento e mineração dos microdados do ENADE foram utilizadas as seguintes ferramentas:

### **2.4.1 Python e Jupyter Notebook**

O trabalho utilizou a linguagem de programação Python (versão 3.7.5) [4], conhecida por sua versatilidade e ampla aplicação em análise de dados. O desenvolvimento dos *scripts* Python e a análise de dados foram realizados no ambiente integrado do Visual Studio Code, com a extensão Jupyter, que oferece uma integração perfeita com o Jupyter Notebook [13].

A escolha do Python como linguagem de programação para esta análise de dados neste trabalho se deve ao fato de ser uma linguagem de código aberto, o que a torna uma opção

gratuita. Além disso, Python oferece uma extensa variedade de bibliotecas especializadas, especialmente voltadas para análise de dados.

O Jupyter Notebook permite a execução de código em células individuais, o que facilita a experimentação e o teste de diferentes abordagens analíticas de forma iterativa e rápida. Isso proporcionou uma experiência de desenvolvimento mais eficiente e colaborativa, permitindo a criação de documentos interativos com código Python, visualizações de dados e documentação detalhada diretamente no Visual Studio Code [14].

## 2.4.2 Bibliotecas utilizadas

As seguintes bibliotecas desempenham papéis importantes na análise de dados em Python, permitindo a extração de informações valiosas, são parte essencial das ferramentas disponíveis para análise de dados em Python.

- **Pandas:** O Pandas [15] é uma biblioteca Python muito poderosa para manipulação e análise de dados estruturados. Fornece estruturas de dados flexíveis, como o DataFrame, que permite armazenar, limpar e analisar dados de maneira eficiente. Foi utilizada para fazer o mapeamento dos microdados.
- **Matplotlib e Seaborn:** São bibliotecas de visualização de dados. O Matplotlib [16] é uma biblioteca de gráficos 2D que permite criar uma variedade de gráficos. O Seaborn [17] é uma camada de alto nível construída sobre o Matplotlib que simplifica a criação de gráficos estatísticos mais atraentes e informativos. Foram utilizadas para gerar tabelas e gráficos para a análise dos dados.
- **Unidecode:** A biblioteca Unidecode [18] é usada para a normalização de *strings*, especialmente útil quando se lida com caracteres especiais ou acentuados como os dados dos cursos e faculdades. Converte esses caracteres para a forma ASCII mais próxima, facilitando a comparação e o processamento de *strings*, tornando o texto mais uniforme e fácil de lidar. Foi utilizada para padronizar os filtros dos dados tratados.
- **NumPy:** NumPy [19] é uma biblioteca que fornece suporte para arrays multidimensionais e funções matemáticas de alto desempenho, tornando possível realizar cálculos complexos e eficientes em grandes conjuntos de dados.

## 2.5 Impacto da LGPD

A Lei Geral de Proteção de Dados Pessoais (LGPD) [20], Lei nº 13.709/2018, representa um marco na legislação brasileira em termos de privacidade e proteção de dados. Esta



lei estabelece diretrizes para o tratamento de dados pessoais, incluindo os digitais, por pessoas físicas ou jurídicas, públicas ou privadas. Seu principal objetivo é proteger os direitos fundamentais de liberdade e privacidade.

A adequação dos microdados [21] do ENADE à LGPD foi realizada pela Diretoria de Avaliação da Educação Superior do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), ilustra o impacto direto da lei na análise de dados educacionais. Com o objetivo de preservar a privacidade dos estudantes, o INEP reestruturou a forma como os microdados do ENADE são disponibilizados. Essa nova estrutura organiza os dados em grupos de informações (arquivos específicos), com cada arquivo ordenado por variáveis distintas, dificultando na identificação de indivíduos. Embora ainda seja possível realizar estudos sobre o perfil dos cursos e seus resultados, a nova estrutura impede a agregação de informações no nível do estudante individual. Isso significa que os dados em um arquivo não podem ser correlacionados diretamente com os de outro arquivo no nível individual, pois a ordenação difere entre eles.

A mineração de dados foi feita analisando cada arquivo dos microdados e filtrando pelo curso analisado e faculdade analisada utilizando o código do curso como critério de agrupamento que será explicada de forma mais detalhada nos próximos capítulos.

# Capítulo 3

## Preparação dos dados

O ENADE do ano 2021 foi selecionado como foco deste trabalho devido à sua relevância como a avaliação mais recente que inclui dados detalhados sobre os cursos de Ciência da Computação, tanto em Licenciatura quanto em Bacharelado. Conforme abordado anteriormente na Fundamentação teórica, a adoção do modelo CRISP-DM envolve etapas específicas para a preparação dos dados:

- **Entendimento do negócio:** Essencial compreender os objetivos da análise dos dados do ENADE. Objetivos que já foram desenvolvidos na introdução do trabalho.
- **Entendimento dos dados:** Esta etapa envolve o *download* e a exploração inicial dos dados para compreender sua estrutura, conteúdo e o as limitações dos dados.
- **Preparação dos dados:** O objetivo é refinar os dados para que estejam prontos para análises subsequentes, assegurando que as informações sejam precisas e úteis.

### 3.1 Coleta de dados

A coleta de dados [22] é uma etapa fundamental em qualquer processo de análise de dados. Envolve a obtenção de informações de fontes relevantes para atender aos objetivos de um trabalho específico. Esta fase é importante para garantir a qualidade e a relevância dos dados para a pesquisa. Dados podem ser coletados de diversas maneiras, incluindo pesquisas, experimentos, bancos de dados existentes, ou através de dados abertos disponibilizados pelo governo.

### 3.1.1 Microdados do ENADE

Os microdados do ENADE [23] estão disponíveis para *download* no site do INEP, abrangendo as edições de 2004 até 2022. Após o *download* do arquivo comprimido do ano desejado (*microdados\_enade\_2021.zip*), os seguintes diretórios estarão disponíveis:

1. LEIA-ME: Contém a documentação dos microdados.
  - Manual do usuário\_Enade\_2021 (pdf).
  - Questionário do Estudante Enade 2021 (pdf).
  - Questionário Licenciaturas - Enade 2021 (pdf).
  - Questões complementares – pandemia (pdf).
  - Dicionário\_arquivos\_variáveis\_microdados\_Enade\_2021 (xlsx).
  - Dicionário\_arquivos\_variáveis\_microdados\_Enade\_2021 (ods).
2. DADOS: Contém os 43 arquivos de microdados referentes ao ano de 2021.
  - microdados2021\_arqX (txt), onde *X* representa o número específico do arquivo, começando de 1 ate 43.

Os microdados são um tipo de conjunto de dados caracterizado por uma estrutura organizada e planejada. Cada linha após o cabeçalho representa um conjunto de dados, com cada valor separado por ponto e vírgula, refletindo a informação correspondente a cada coluna. Contêm informações sobre o desempenho dos alunos nas provas, incluindo resultados das provas, questões das provas e contexto socioeconômico e educacional.

### 3.1.2 Cursos de Graduação do Brasil

O site indicadores sobre ensino superior - dados abertos - Ministério da Educação (MEC) [24] fornece indicadores governamentais relacionados ao ensino superior no Brasil. Após o *download* do arquivo *Cursos de Graduação do Brasil* estará disponível da seguinte forma:

- PDA\_Dados\_Cursos\_Graduacao\_Brasil (csv).

O arquivo “Cursos de Graduação do Brasil” apresentado no formato CSV, que oferece benefícios como fácil manipulação dos dados em diversas ferramentas de análise e programação. A estrutura inclui diversas colunas que fornecem informações detalhadas sobre as instituições e os cursos oferecidos, vai ser necessário para fazer o mapeamento dos nomes das instituições na etapa de tratamento dos dados.

## 3.2 Tratamento dos dados

A fase de tratamento de dados ou pré-processamento [8] [12] é essencial na mineração de dados, essa etapa prepara os dados para a aplicação eficaz de algoritmos de mineração. As principais atividades envolvidas no pré-processamento incluem:

- **Limpeza dos dados:** Esta etapa visa resolver problemas como registros incompletos, valores incorretos ou dados inconsistentes, para que não afetem negativamente os resultados dos algoritmos de mineração.
- **Integração dos dados:** Frequentemente, os dados são coletados de fontes diversas, como bancos de dados, arquivos de texto, planilhas. Isso torna necessária a integração desses dados em um repositório unificado e consistente. Nesse processo, é importante analisar os dados em busca de redundâncias, dependências entre variáveis e conflitos, como categorias distintas para os mesmos valores.
- **Redução dos dados:** Devido ao volume frequentemente elevado dos dados na mineração, em alguns casos é impraticável processar e analisar a totalidade dos dados. As técnicas de redução de dados visam transformar o grande volume de dados em um conjunto menor, mas que ainda assim mantenha a representatividade dos dados originais. Isso possibilita a execução mais eficiente dos algoritmos de mineração, preservando a qualidade dos resultados.
- **Transformação dos dados:** Certos algoritmos para a leitura dos dados, exigem dados em formatos específicos, como numéricos ou nominais. Portanto, pode ser necessário transformar valores numéricos em nominais ou vice-versa.

Essas etapas garantem que os dados estejam na forma mais adequada para análises e descobertas significativas através dos algoritmos de mineração de dados envolve várias etapas críticas, algumas das quais serão implementadas nos *scripts* das seções seguintes.

### 3.2.1 Tratamento dos vetores

Este *script* realiza uma análise do arquivo do ENADE que contém os dados das notas formação geral e componente específico, identificando as seguintes colunas específicas que tem o conteúdo dos vetores de notas [25].

1. **DS\_VT\_ACE\_OFG:** Vetor que representa os acertos da parte objetiva na formação geral.
2. **DS\_VT\_ACE\_OCE:** Vetor que representa os acertos da parte objetiva do componente específico.

3. **DS\_VT\_ESC\_OFG:** Vetor que representa a escolha de resposta da parte objetiva da formação geral.
4. **DS\_VT\_ESC\_OCE:** Vetor que representa a escolha de resposta da parte objetiva do componente específico.

Os vetores que tem conteúdo da formação geral são de tipo *string* e tem 8 caracteres enquanto os vetores que tem conteúdo do componente específico tem 27 caracteres. Após a verificação, ele extrai informações de cada vetor e divide ele em caracteres criando várias novas colunas com esses dados reformatados. Portanto, ao realizar essa transformação, os dados são estruturados de forma a permitir uma análise mais detalhada do desempenho individual dos alunos nas questões objetivas. Isso facilita a identificação de quais respostas dos alunos estão corretas ou incorretas, com base no gabarito final disponibilizado. O resultado é um arquivo CSV enriquecido com dados mais acessíveis para análises posteriores.

- **Entrada:** `microdados2021_arq3` (txt, com microdados do ENADE).
- **Saída:** `vetor_microdado_3` (csv, que contém dados do `microdados2021_arq3`, com dados reestruturados para análise).

### 3.2.2 Arquivo auxiliar para mapeamento

Este *script* processa o arquivo de microdados do ENADE que contém o código e caracterização do curso, adicionando informações claras sobre cursos e instituições. Essa informação não está disponível nos microdados e não era possível determinar os nomes das universidades correspondentes. No entanto, essa limitação foi superada com a introdução do arquivo auxiliar para fazer o mapeamento. O *script* cria uma coluna e faz o mapeamento inicialmente do Nome da Instituição de Educação Superior (**NOME\_IES**) pelo Código da IES (**CODIGO\_IES**). Se for comprovado que ainda faltam dados dos nomes, é feito o segundo mapeamento pelo Código do Curso (**CODIGO\_CURSO**). Finalmente faz um mapeamento para ter os nomes dos 30 cursos participantes pelo Código do Curso (**CO\_CURSO**). Facilitando a compreensão dos dados. Após mapear e adicionar esses dados, salva o resultado em um novo arquivo CSV.

- **Entradas:**
  - `PDA_Dados_Cursos_Graduacao_Brasil` (csv, com dados complementares.)
  - `microdados2021_arq1` (txt, com microdados do ENADE.)
- **Saída:** `microdados_aux` (csv, que contém dados do `microdados2021_arq1` adicionando as colunas com o Nome da IES e Nome do curso.)

### 3.2.3 Formatação do dados dos vetores

O *script* processa o arquivo de entrada e realiza mapeamentos e pelo Código do Curso (**CO\_CURSO**) e adiciona novas colunas com os dados do Código da IES, Nome da IES e Nome do Curso. As colunas com os dados Nome da IES e Nome do Curso passam pela função *unidecode* para padronizar o conteúdo das *strings*, que vai tirar todos os acentos e cedilhas, o que pode ser útil para evitar problemas de compatibilidade e simplificar a análise de dados. Isso garante que os algoritmos utilizados na análise de dados processem o texto de forma uniforme, minimizando a possibilidade de erros ou comportamentos inesperados.

Finalmente são classificadas as colunas do arquivo com o conteúdo das notas para ter apenas a informação mais importante, essas colunas tem informações do Nome e Código do curso; Nome e Código da IES; Dados dos vetores de formação geral e componente específico.

- **Entradas:**
  - PDA\_Dados\_Cursos\_Graduacao\_Brasil (csv, com dados complementares)
  - vetor\_microdado\_3 (csv, com microdados do ENADE melhor estruturados)
- **Saída:** form\_microdados\_arq3\_vetores\_filtrados (csv, com dados classificados)

### 3.2.4 Formatação dos dados com a percepção da prova

O *script* processa o arquivo de entrada, inicialmente faz o mapeamento das respostas com o conteúdo da percepção da prova, são 9 perguntas então tem 9 mapeamentos para cada pergunta depois faz o mesmo processo que o *script* anterior onde realiza os mapeamentos pelo Código do Curso (**CO\_CURSO**) e adiciona novas colunas com os dados do Código da IES, Nome da IES e Nome do Curso. As colunas com os dados Nome da IES e Nome do Curso passam pela função *unidecode* para padronizar o conteúdo das *strings* porque vários algoritmos podem apresentar erros ao lidar com caracteres acentuados e cedilhas.

Finalmente são classificadas as colunas do arquivo com o conteúdo do Nome e Código do curso; Nome e Código da IES; Respostas mapeadas da percepção da prova.

- **Entradas:**
  - PDA\_Dados\_Cursos\_Graduacao\_Brasil (csv, com dados complementares)
  - vetor\_microdado\_3 (csv, com microdados do ENADE melhor estruturados)
- **Saída:** form\_microdados\_arq3\_percepcao\_prova (csv, com dados classificados)

### 3.2.5 Tratamento dos questionários

Para processar os dados do questionário do estudante foi implementada a função (*processar\_arquivo*). Esse mapeamento ajuda para a interpretação dos dados de forma direta, sem a necessidade constante de consultar tabelas ou dicionários [25]. Isso simplifica a análise e torna os resultados mais acessíveis e compreensíveis. Esta função tem os seguintes parâmetros de entrada:

1. `arquivo_entrada`, diretório onde lê os dados do arquivo de entrada.
2. `arquivo_auxiliar`, utiliza ele para mapear o nome e código do curso.
3. `arquivo_saida`, diretório destino para os dados processados.
4. `mapeamento_adicional`, mapeamento específico que depende de cada questão.
5. `coluna_qe`, coluna a ser mapeada.

Na seção seguinte tem um exemplo prático do uso da função (*processar\_arquivo*), utilizando os dados de entrada para a primeira questão do Questionário do Estudante que é a seguinte; “QE1: Qual o seu estado civil?”.

```
mapeamento_qe_i01 = {
    "A": "Solteiro(a)",
    "B": "Casado(a)",
    "C": "Separado(a) judicialmente/divorciado(a)",
    "D": "Viúvo(a)",
    "E": "Outro"
}

processar_arquivo(
    'microdados_Enade_2021_LGPD/2.DADOS/microdados2021_arq7.txt',
    'aux_files/vetores/microdados_aux.csv',
    f'{dir_destino}/form_microdados_arq7.csv',
    mapeamento_qe_i01,
    'QE_I01'
)
```

O arquivo processado tem o formato de CSV, foi feito esse processo para o questionário do estudante com seu respectivo mapeamento específico. Todos os arquivos formatados são armazenados no diretório (*microdados\_tratados*), cada arquivo tem o nome com o seguinte formato:

- **form\_microdados\_arqY** (csv), onde “Y” representa o número específico do arquivo já formatado.



# Capítulo 4

## Análise dos dados

Seguindo o modelo CRISP-DM na seção 2.3, o desenvolvimento das seguintes etapas é fundamental para iniciar a análise de dados do ENADE. Este processo de análise é seguido com em três fases essenciais:

- **Modelagem:** O objetivo desta etapa é estabelecer uma base para a análise de dados. Inicialmente, formulamos os objetivos do trabalho, que incluem a identificação de fatores que influenciam o desempenho dos alunos.
- **Avaliação:** Nesta seção, abordamos a avaliação dos modelos criados na etapa de modelagem. A avaliação é fundamental para determinar o quão bem os modelos generalizam para dados não observados.
- **Implantação:** Esta etapa envolve a implantação prática do modelo escolhido em um contexto de uso real. Documentação adequada deve ser fornecida para que os usuários entendam como usar o modelo corretamente e interpretar seus resultados.

### 4.1 Análise geral dos dados

Iniciando a análise com os dados gerais de todos os cursos participantes do ENADE a Tabela 4.1 representa a participação dos alunos em cada curso. A tabela tem uma estrutura de quatro colunas principais:

- **Cursos participantes do ENADE:** Esta coluna lista os 30 cursos avaliados pelo ENADE, categorizados por grau acadêmico de Licenciatura ou Bacharelado.
- **Número total de alunos inscritos:** Reflete a quantidade total de estudantes inscritos para participar do ENADE em cada curso.
- **Número de alunos que realizaram a prova:** Indica o número efetivo de estudantes que compareceram para realizar o exame.

- **Percentual que realizou a prova:** Mostra o percentual dos inscritos que de fato realizaram o exame, oferecendo uma perspectiva sobre a taxa de adesão ao ENADE por curso.

Tabela 4.1: Tabela com os cursos participantes do ENADE

| Cursos participantes do ENADE                       | Número total de alunos inscritos | Número de alunos que realizaram a prova | Percentual que realizou a prova |
|---|----------------------------------|---|---------------------------------|
| Artes Visuais (Licenciatura)                        | 4892                             | 3481                                    | 71.16%                          |
| Ciencia Da Computacao (Bacharelado)                 | 12733                            | 9987                                    | 78.43%                          |
| Ciencia Da Computacao (Licenciatura)                | 2064                             | 1447                                    | 70.11%                          |
| Ciencias Biologicas (Bacharelado)                   | 8505                             | 6753                                    | 79.4%                           |
| Ciencias Biologicas (Licenciatura)                  | 18542                            | 14150                                   | 76.31%                          |
| Ciencias Sociais (Bacharelado)                      | 3309                             | 1967                                    | 59.44%                          |
| Ciencias Sociais (Licenciatura)                     | 4550                             | 2836                                    | 62.33%                          |
| Design  | 4732                             | 3846                                    | 81.28%                          |
| Educacao Fisica (Bacharelado)                       | 38367                            | 28570                                   | 74.47%                          |
| Educacao Fisica (Licenciatura)                      | 26688                            | 18432                                   | 69.06%                          |
| Filosofia (Bacharelado)                             | 2256                             | 1413                                    | 62.63%                          |
| Filosofia (Licenciatura)                            | 5724                             | 3434                                    | 59.99%                          |
| Fisica (Licenciatura)                               | 4788                             | 3559                                    | 74.33%                          |
| Geografia (Bacharelado)                             | 3287                             | 1982                                    | 60.3%                           |
| Geografia (Licenciatura)                            | 15395                            | 10027                                   | 65.13%                          |
| Historia (Bacharelado)                              | 2621                             | 1426                                    | 54.41%                          |
| Historia (Licenciatura)                             | 22371                            | 14973                                   | 66.93%                          |
| Letras-Ingles (Licenciatura)                        | 4431                             | 2883                                    | 65.06%                          |
| Letras-Portugues (Licenciatura)                     | 17775                            | 11879                                   | 66.83%                          |
| Letras-Portugues e Espanhol (Licenciatura)          | 2323                             | 1733                                    | 74.6%                           |
| Letras-Portugues e Ingles (Licenciatura)            | 8311                             | 6276                                    | 75.51%                          |
| Matematica (Licenciatura)                           | 19382                            | 13824                                   | 71.32%                          |
| Musica (Licenciatura)                               | 4119                             | 3052                                    | 74.1%                           |
| Pedagogia (Licenciatura)                            | 195137                           | 148303                                  | 76.0%                           |
| Quimica (Bacharelado)                               | 2729                             | 2185                                    | 80.07%                          |
| Quimica (Licenciatura)                              | 7601                             | 5765                                    | 75.85%                          |
| Sistemas de Informacao                              | 12789                            | 8579                                    | 67.08%                          |
| Tecnologia em Analise e Desenvolvimento de Sistemas | 24428                            | 16065                                   | 65.76%                          |
| Tecnologia em Gestao da Tecnologia da Informacao    | 6754                             | 4025                                    | 59.59%                          |
| Tecnologia em Redes de Computadores                 | 3263                             | 2047                                    | 62.73%                          |
| Total alunos ENADE                                  | 489866                           | 354899                                  | 72.45%                          |

#### 4.1.1 Heatmap dos dados do questionário do estudante

Um *heatmap* [26] é uma representação gráfica que sintetiza e apresenta conjuntos de dados complexos de forma intuitiva e imediatamente compreensível. Cada célula da matriz é colorida de forma proporcional ao valor que representa, facilitando a identificação de padrões, tendências e *outliers*. Além disso, os heatmaps são frequentemente usados para visualizar a correlação entre variáveis, densidade de informação, ou a variação de valores ao longo do tempo ou espaço.

Neste caso eles podem ser aplicados para visualizar uma variedade de aspectos relacionados ao desempenho e participação dos alunos e a concentração da maior parte dos alunos nos diferentes cursos utilizando as perguntas do questionário do estudante.

Para ilustrar a distribuição de gênero entre os participantes do Exame, a Figura 4.1 apresenta um *heatmap* onde as cores do mapa variam do azul claro a azul escuro, indicando uma escala crescente de respostas, onde o azul mais claro representa uma menor quantidade de respostas e o azul mais escuro denota uma maior quantidade. As linhas no eixo vertical listam os cursos participantes e no eixo horizontal, temos duas categorias: Feminino e Masculino, que indicam o sexo dos respondentes.

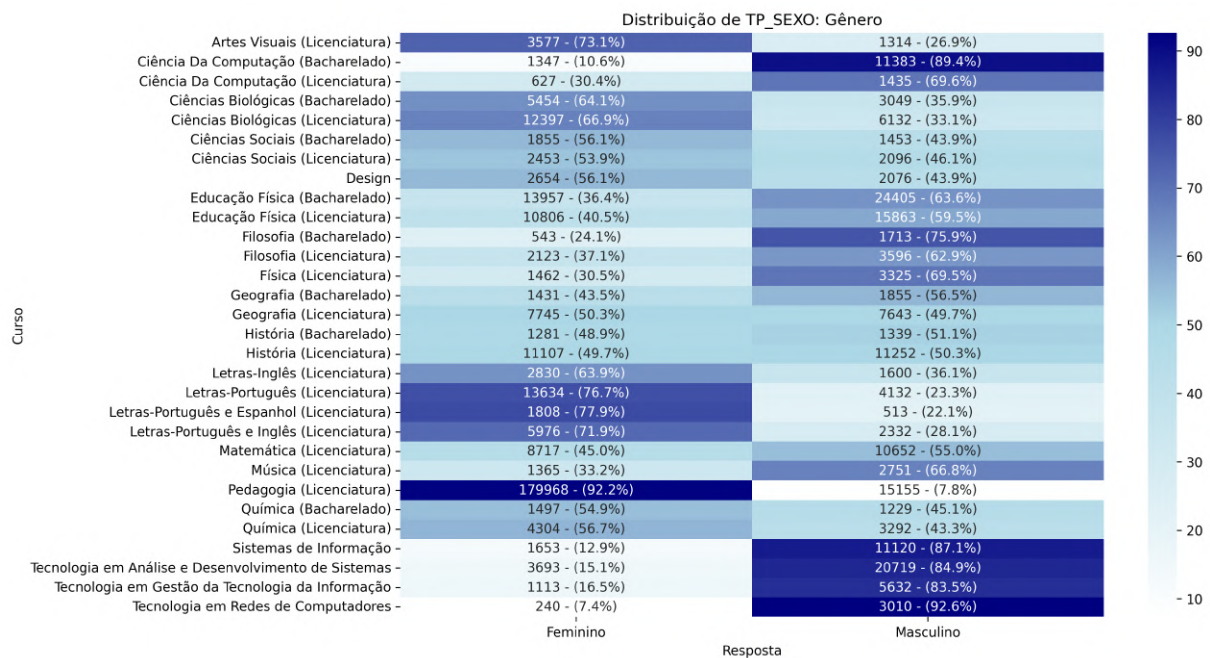


Figura 4.1: Heatmap do sexo dos participantes.

Um padrão notável que pode ser observado onde o curso de Pedagogia (Licenciatura) tem um número significativamente maior de respostas femininas, sugerindo uma predominância de mulheres neste campo. Em contraste, alguns cursos apresentam uma distribuição mais equilibrada entre os sexos, indicada por tons intermediários de azul.

A Figura 4.2 representa a idade dos alunos no dia da prova do ENADE, tendo uma grande concentração no grupo de 18 a 25 anos na maioria dos cursos.

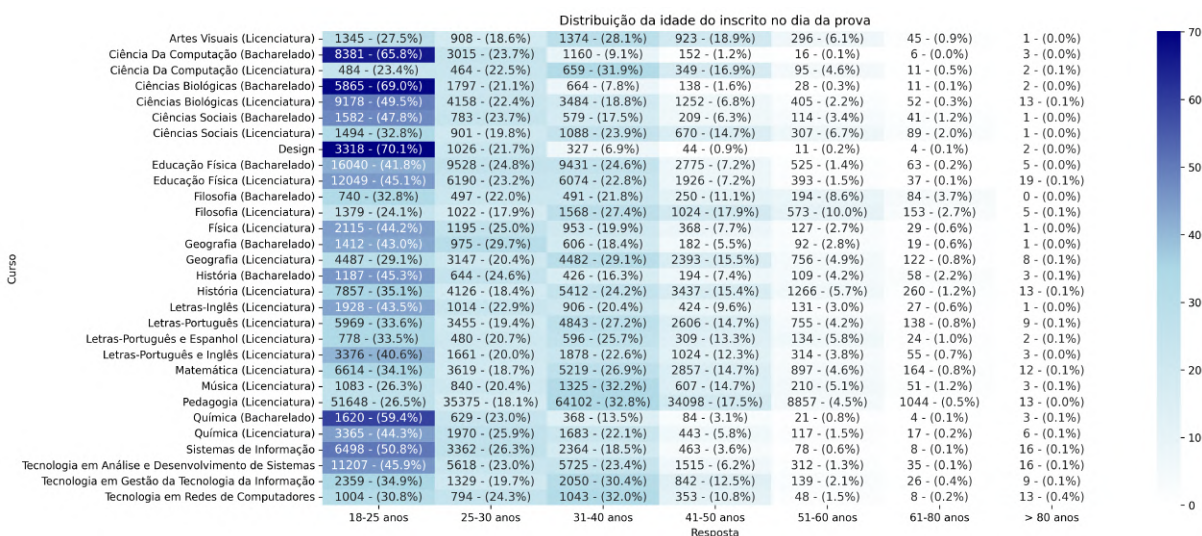


Figura 4.2: Heatmap da idade dos participantes.

Seguindo temos a Figura 4.3 que ilustra as respostas dos alunos para a questão 1 do questionário do estudante sendo “Qual o seu estado civil?”. As respostas estão divididas em seis categorias:

- Solteiro(a).
- Casado(a).
- Separado(a) judicialmente/divorciado(a).
- Viúvo(a).
- Outro.

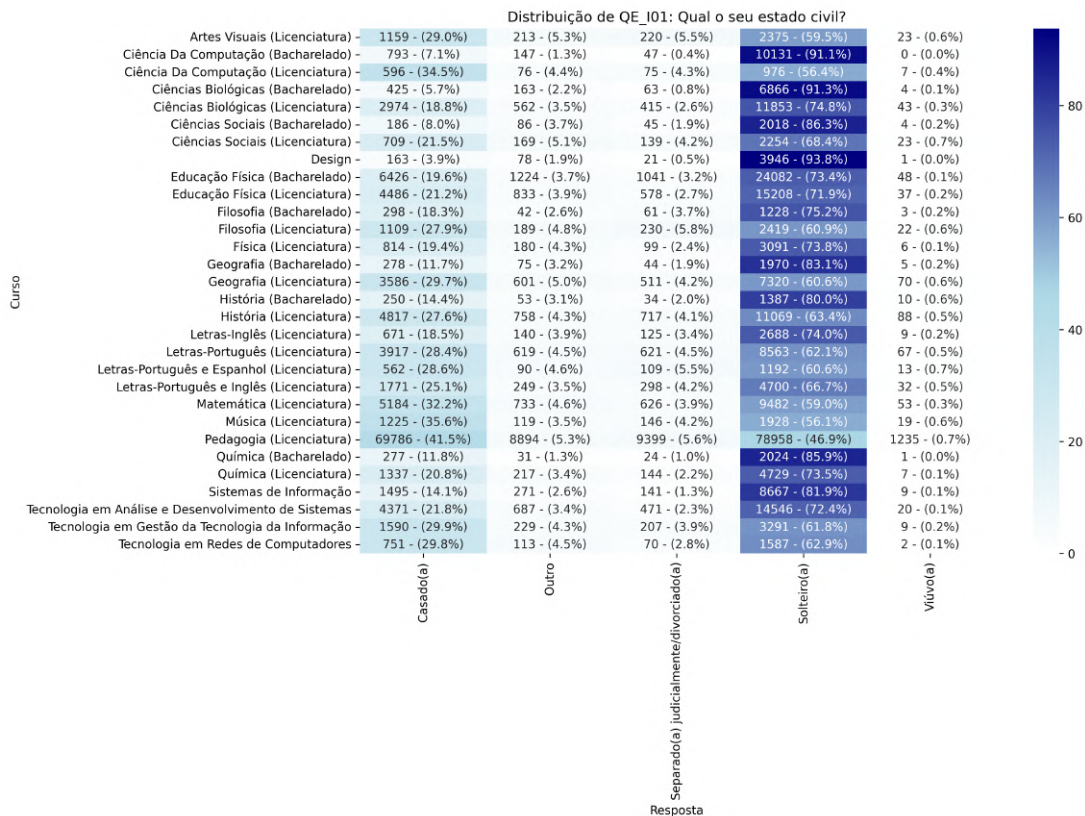


Figura 4.3: Heatmap do estado civil dos participantes.

Observa-se uma predominância da cor que representa alunos solteiros em quase todos os cursos, sugerindo que a maioria dos participantes tem esse estado civil. Entretanto, um destaque notável para a segunda resposta na maioria dos cursos que tem uma boa parte do grupo com participantes casados.

A Figura 4.4 detalha a autodeclaração de cor/raça dos alunos, a coleta de dados sobre raça ou etnia são fundamentais em muitos contextos sociais e educacionais, especialmente em países como o Brasil, onde tem uma diversidade étnico-racial muito abrangente, os dados são da questão 2 do questionário do estudante sendo “Qual é a sua cor ou raça?” e são apresentados nas seguintes categorias:

- Branca.
- Preta.
- Amarela.
- Parda.
- Indígena.
- Não quero declarar.



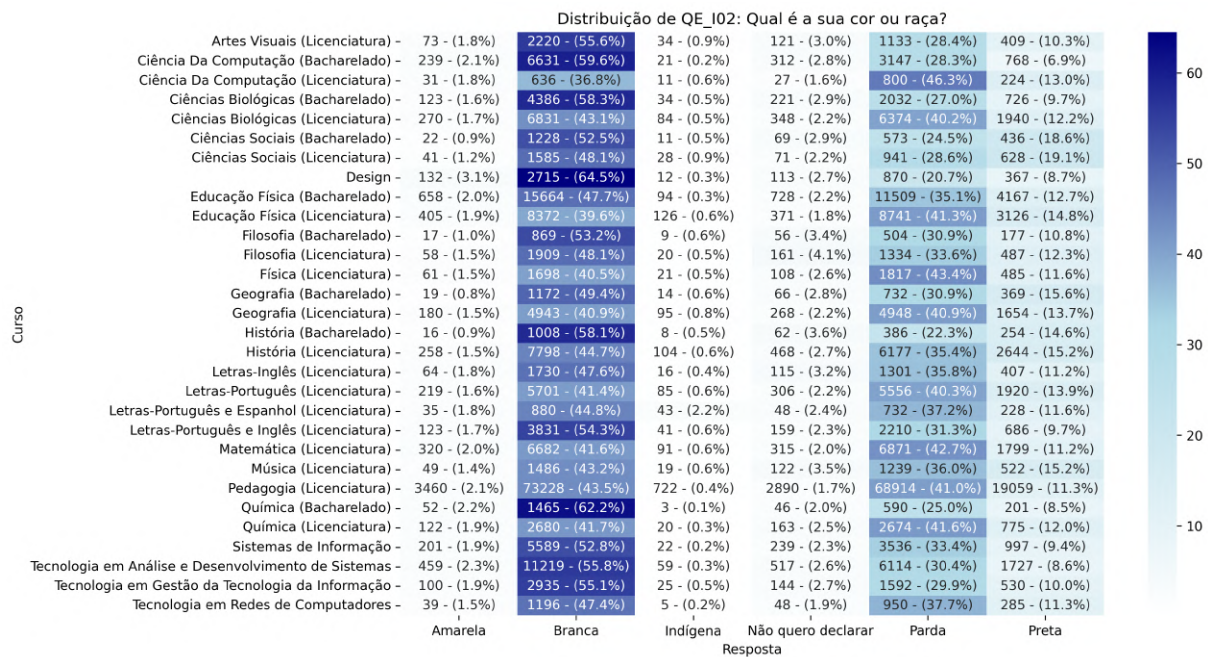


Figura 4.4: Heatmap da cor/raça dos participantes.

Tem uma predominância de estudantes que se declaram de cor branca, seguidos por aqueles que se declaram pardos, indicados pelas células em cor azul nestas categorias. Isto pode refletir as demografias da população universitária brasileira de uma forma mais geral. Este padrão de autodeclaração de cor/raça nos cursos pode oferecer um ponto de partida para discussões sobre equidade e inclusão dentro das instituições de ensino superior, e destaca a importância de políticas de acesso e suporte que promovam a diversidade no ambiente acadêmico.

A Figura 4.5 mapeia a quantidade de alunos por faixa de renda familiar mensal. A análise da distribuição de renda dos estudantes que participam do ENADE é crucial porque fornece *insights* sobre a diversidade socioeconômica da população estudantil do ensino superior. Este dado pode refletir as barreiras econômicas que influenciam quem tem a oportunidade de chegar ao ensino superior e quais são os cursos mais acessíveis para estudantes de diferentes estratos sociais. Os dados são da questão 8 do questionário do estudante sendo “Qual a renda total de sua família, incluindo seus rendimentos?”, em 2021 o valor do salário mínimo no Brasil era de R\$ 1.100,00, e os grupos de salario são apresentados nas seguintes categorias:

- Até 1,5 salário mínimo (até R\$ 1.650,00).
- De 1,5 a 3 salários mínimos (R\$ 1.650,01 a R\$ 3.300,00).
- De 3 a 4,5 salários mínimos (R\$ 3.300,01 a R\$ 4.950,00).

- De 4,5 a 6 salários mínimos (R\$ 4.950,01 a R\$ 6.600,00).
- De 6 a 10 salários mínimos (R\$ 6.600,01 a R\$ 11.000,00).
- De 10 a 30 salários mínimos (R\$ 11.000,01 a R\$ 33.000,00).
- Acima de 30 salários mínimos (mais de R\$ 33.000,00).

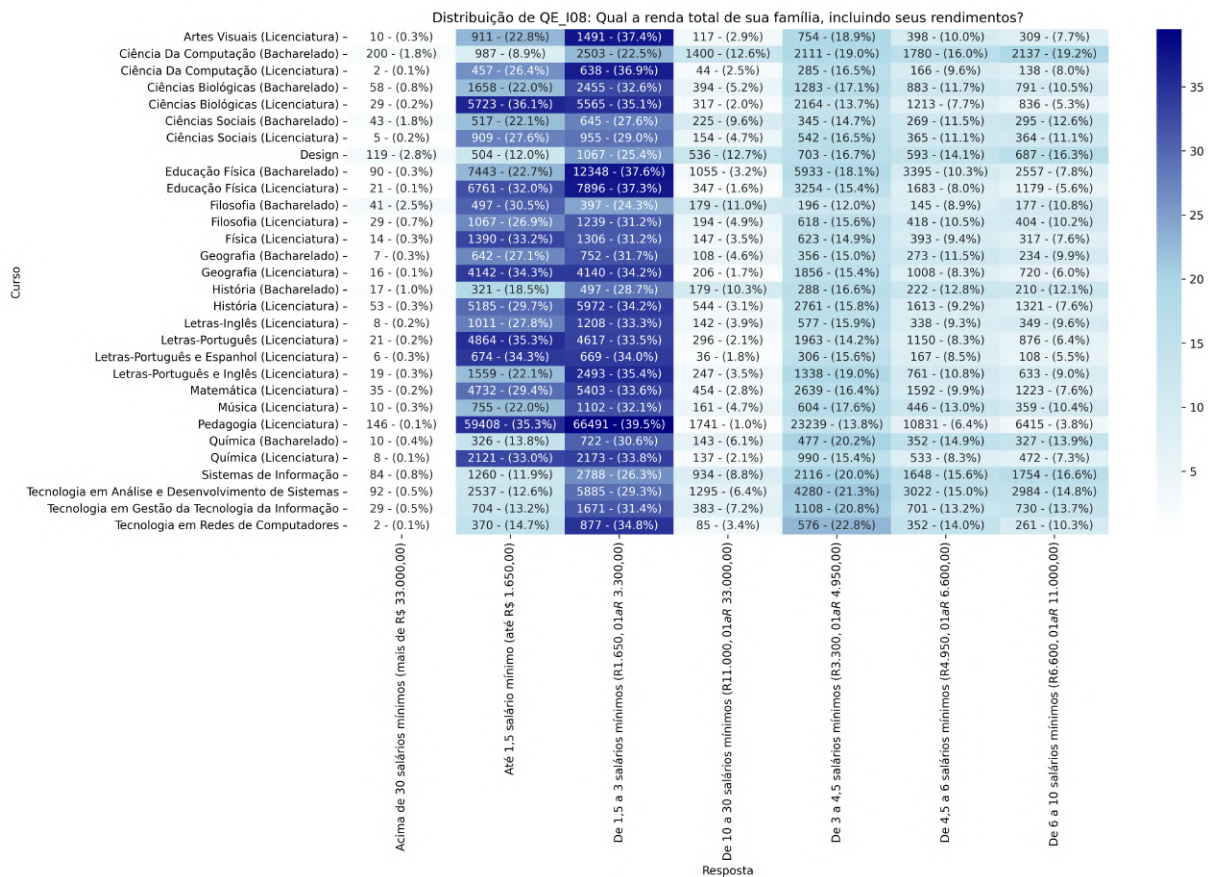


Figura 4.5: Heatmap da renda familiar mensal dos participantes.

Ao examinar os padrões entre os cursos de bacharelados e licenciaturas emergem algumas distinções notáveis. Cursos de bacharelado como Ciência da Computação e Sistemas de Informação mostram uma representatividade considerável na faixa de 6 a 10 salários mínimos (R\$ 6.600,01 a R\$ 11.000,00), possivelmente refletindo um acesso proporcional de estudantes de famílias com renda mais alta a essas áreas de estudo. A análise também revela que a categoria de 4,5 a 6 salários mínimos tende a ter uma distribuição mais uniforme, sugerindo que estudantes dessa faixa de renda estão relativamente bem distribuídos entre os diferentes cursos.

A Figura 4.6 tem a representação do estado onde concluíram o ensino médio. Compreender onde os alunos concluíram o ensino médio é importante porque pode indicar a

existência de fluxos migratórios de estudantes em busca de oportunidades de ensino superior, os dados são da questão 16 do questionário do estudante sendo “Em que unidade da Federação você concluiu o ensino médio?” e representam os 26 estados brasileiros e 1 uma opção de “Não se aplica” destinada para alunos que concluíram o ensino médio fora do Brasil.





Há uma concentração notável de estudantes que concluíram o ensino médio em estados como São Paulo, Minas Gerais e Rio de Janeiro. Isso pode sugerir uma maior centralização de instituições de ensino superior nestas áreas ou uma maior procura por esses estados devido a oportunidades educacionais percebidas. Cursos como Pedagogia (Licenciatura) em São Paulo e Sistemas de Informação em estados como Minas Gerais demonstram um alto volume de estudantes, potencialmente refletindo a demanda regional por formação nestes campos e a concentração de oportunidades educacionais.

Complementando o perfil educacional dos participantes do ENADE, a Figura 4.7 tem dados sobre o tipo de escola onde concluíram o ensino médio. Isso pode refletir nas diferenças de acesso à educação de qualidade, as desigualdades socioeconômicas, e pode impactar o desempenho acadêmico no ensino superior. Os dados são da questão 17 do questionário do estudante sendo “Em que tipo de escola você cursou o ensino médio?” e são apresentados nas seguintes categorias:

- Todo em escola pública.
- Todo em escola privada (particular).
- Todo no exterior.
- A maior parte em escola pública.
- A maior parte em escola privada (particular).
- Parte no Brasil e parte no exterior.

Há um padrão predominante de estudantes que cursaram o ensino médio em escolas públicas em quase todos os cursos listados. Este dado pode ser influenciado pela implementação de cotas para estudantes de escolas públicas. Observa-se também uma representatividade notável de alunos oriundos de escolas particulares em cursos como Ciência da Computação (Bacharelado) e Tecnologia em Análise e Desenvolvimento de Sistemas.

### 4.1.2 Análise das notas

O ENADE desempenha um papel importante na avaliação do ensino superior no Brasil, fornecendo dados essenciais sobre a qualidade dos cursos e a competência dos graduandos. A prova do ENADE compreende de dois componentes fundamentais:

- **Formação geral (FG):** Esta seção é formada por um conjunto de 10 questões, das quais 8 são de múltipla escolha e 2 são questões discursivas. O propósito desta seção é avaliar a capacidade do aluno de compreender e refletir sobre questões que vão

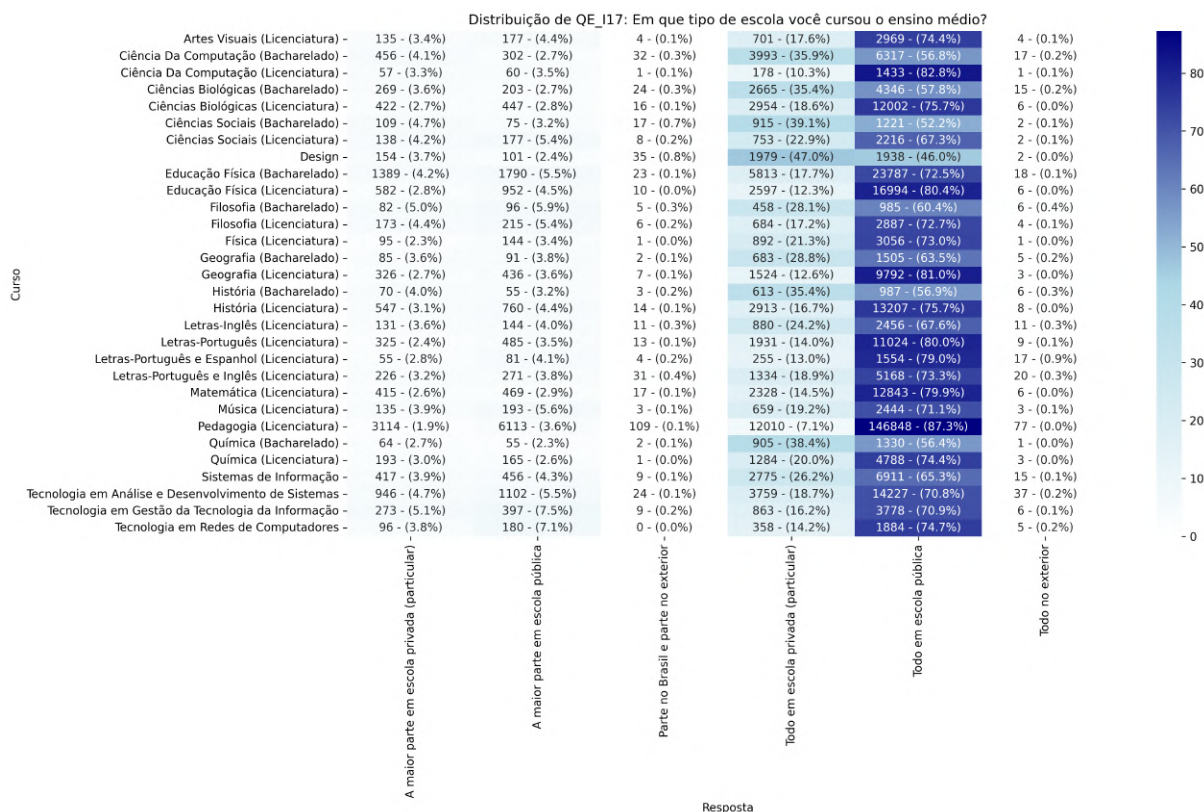


Figura 4.7: Heatmap do tipo de escola dos participantes.

além de sua área de formação profissional. Os alunos devem demonstrar um entendimento crítico de questões contemporâneas, habilidades de síntese contextualizada e a capacidade de expressar-se com clareza na correção gramatical.

- **Conhecimentos específicos (CE):** Composta por 30 questões — 27 de múltipla escolha e 3 discursivas, esta seção visa avaliar os conhecimentos e habilidades específicas relacionadas à área de estudo do curso do aluno. O exame explora diferentes níveis de complexidade e avalia a prontidão do aluno para o exercício profissional.

A avaliação dos graduandos nas questões de formação geral é crucial, pois reflete a capacidade de integrar conhecimentos e desenvolver uma visão crítica sobre temas relevantes no mundo atual. Espera-se que os estudantes demonstrem competências em diversas áreas que incluem, mas não se limitam à área de estudo, avalia a clareza na comunicação, a coerência e coesão argumentativa, bem como o uso adequado do vocabulário. Por outro lado, o componente de conhecimentos específicos foca na particularidade de cada área de avaliação do ENADE, abordando o domínio de conhecimentos e habilidades esperadas para os conhecimentos do curso específico. Uma análise mais detalhada do cálculo das notas pode ser encontrado na seção 2.1.1 da Fundamentação Teórica.

Para fazer a padronização dos dados utilizaremos medidas como a média e mediana [27], que medem a tendência central em estatísticas que ajudam a resumir um conjunto de dados com um único valor que representa o centro do conjunto, calculam a centralidade de maneiras diferentes e têm implicações distintas na análise de dados.

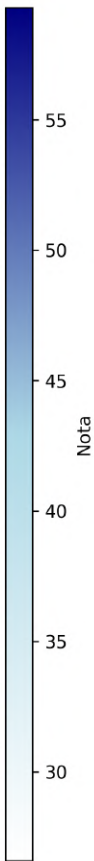
- **Média:** A média é o valor que se obtém ao somar todos os números de um conjunto de dados e depois dividir essa soma pelo número total de elementos no conjunto.
- **Mediana:** A mediana é o valor que divide um conjunto de dados ordenados em duas partes iguais. Para encontrar a mediana, os valores devem primeiro ser ordenados em sequência crescente ou decrescente. Se o número de observações for ímpar, a mediana corresponderá ao valor central. Se for par, será a média dos dois valores centrais.

A principal vantagem da mediana sobre a média é sua resistência a valores extremos, que garante que a medida de tendência central seja mais representativa do conjunto de dados em situações onde a média poderia ser distorcida por valores atípicos.

A Tabela 4.2 mostra as medianas para cada curso das notas na prova geral ( $NT_{FG}$ ), nas questões específicas do curso ( $NT_{CE}$ ), e a nota geral ( $NT_{GER}$ ), respectivamente.

Tabela 4.2: Tabela das medianas das notas dos participantes do ENADE

| Cursos participantes do ENADE                       | Número de alunos que realizaram a prova | NT_FG | NT_CE | NT_GER |
|---|---|-------|-------|--------|
| Artes Visuais (Licenciatura)                        | 3481                                    | 33.30 | 48.40 | 44.60  |
| Ciencia Da Computacao (Bacharelado)                 | 9987                                    | 37.50 | 35.80 | 35.70  |
| Ciencia Da Computacao (Licenciatura)                | 1447                                    | 32.30 | 28.30 | 29.40  |
| Ciencias Biologicas (Bacharelado)                   | 6753                                    | 42.00 | 46.50 | 45.50  |
| Ciencias Biologicas (Licenciatura)                  | 14150                                   | 37.50 | 44.50 | 42.70  |
| Ciencias Sociais (Bacharelado)                      | 1967                                    | 41.70 | 43.10 | 42.80  |
| Ciencias Sociais (Licenciatura)                     | 2836                                    | 40.40 | 47.80 | 45.90  |
| Design  | 3846                                    | 45.00 | 50.80 | 49.00  |
| Educacao Fisica (Bacharelado)                       | 28570                                   | 28.10 | 38.30 | 35.70  |
| Educacao Fisica (Licenciatura)                      | 18432                                   | 29.20 | 35.80 | 34.00  |
| Filosofia (Bacharelado)                             | 1413                                    | 37.50 | 49.00 | 46.50  |
| Filosofia (Licenciatura)                            | 3434                                    | 38.55 | 53.30 | 49.65  |
| Fisica (Licenciatura)                               | 3559                                    | 37.50 | 38.00 | 38.00  |
| Geografia (Bacharelado)                             | 1982                                    | 42.00 | 43.70 | 43.40  |
| Geografia (Licenciatura)                            | 10027                                   | 34.20 | 42.80 | 40.50  |
| Historia (Bacharelado)                              | 1426                                    | 45.00 | 59.30 | 55.80  |
| Historia (Licenciatura)                             | 14973                                   | 36.70 | 42.50 | 41.30  |
| Letras-Ingles (Licenciatura)                        | 2883                                    | 38.80 | 41.20 | 40.70  |
| Letras-Portugues (Licenciatura)                     | 11879                                   | 34.30 | 44.10 | 41.40  |
| Letras-Portugues e Espanhol (Licenciatura)          | 1733                                    | 35.70 | 40.20 | 38.90  |
| Letras-Portugues e Ingles (Licenciatura)            | 6276                                    | 39.20 | 40.70 | 40.30  |
| Matematica (Licenciatura)                           | 13824                                   | 33.50 | 39.50 | 38.10  |
| Musica (Licenciatura)                               | 3052                                    | 35.30 | 43.50 | 41.80  |
| Pedagogia (Licenciatura)                            | 148303                                  | 28.00 | 37.00 | 34.60  |
| Quimica (Bacharelado)                               | 2185                                    | 38.80 | 32.30 | 34.10  |
| Quimica (Licenciatura)                              | 5765                                    | 37.50 | 38.00 | 37.30  |
| Sistemas de Informacao                              | 8579                                    | 34.80 | 39.20 | 38.30  |
| Tecnologia em Analise e Desenvolvimento de Sistemas | 16065                                   | 32.00 | 28.30 | 28.40  |
| Tecnologia em Gestao da Tecnologia da Informacao    | 4025                                    | 30.00 | 37.00 | 35.20  |
| Tecnologia em Redes de Computadores                 | 2047                                    | 28.10 | 26.90 | 26.60  |



Os cursos como Ciências Biológicas (Bacharelado) e História (Bacharelado) apresentam medianas relativamente altas no componente específico do curso (NT\_CE), um sólido desempenho dos estudantes nas áreas focais de seus estudos. Por outro lado, cursos como Educação Física (Bacharelado) e Pedagogia (Licenciatura) exibem medianas mais baixas na formação geral, sugerindo que essas áreas podem demandar atenção adicional para potencialmente elevar o nível de desempenho dos estudantes.

É notável que a mediana da nota geral (NT\_GER) tende a seguir as tendências das medianas dos componentes específicos dos cursos. Isso sugere que o desempenho dos estudantes nas disciplinas específicas pode ter um impacto significativo na nota geral.

## 4.2 Análise das notas do curso Ciência da Computação

Na continuidade da análise de desempenho, o foco recairá especificamente sobre os cursos de Ciência da Computação, tanto bacharelado quanto licenciatura, com o objetivo de examinar o rendimento particular desses programas.

As medidas de primeiro quartil, mediana e terceiro quartil são estatísticas de posição que dividem um conjunto de dados ordenados em partes iguais e são fundamentais para entender a distribuição dos dados das notas.

- **Primeiro Quartil (25º Quartil):** Representa o valor abaixo do qual se encontram 25% dos dados quando ordenados do menor para o maior. É um ponto de corte que separa os 25% dos valores mais baixos do restante dos dados. Em outras palavras, é a mediana da primeira metade dos dados.
- **Terceiro Quartil (75º Quartil):** Representa o valor acima do qual se encontram 25% dos dados. Ele separa os 25% dos valores mais altos do resto do conjunto de dados. O terceiro quartil é a mediana da segunda metade dos dados.

### 4.2.1 Nota bruta na formação geral

A Tabela 4.3 fornece um olhar detalhado sobre as notas brutas na formação geral (*NT\_FG*), na primeira coluna ENADE avalia a nota para todos os cursos participantes, depois para os cursos de Ciência da Computação, apresenta medidas de posição, como o primeiro quartil (25 Quartil), a mediana e o terceiro quartil (75 Quartil), além do número total de alunos que realizaram a prova e o percentual de participação.

Tabela 4.3: Notas da formação geral

|                            | ENADE  | Ciencia Da Computacao (Licenciatura) | Ciencia Da Computacao (Bacharelado) |
|----------------------------|--------|--------------------------------------|-------------------------------------|
| 25 Quartil                 | 21.60  | 23.25                                | 26.90                               |
| Mediana                    | 31.30  | 32.30                                | 37.50                               |
| 75 Quartil                 | 42.80  | 41.65                                | 48.30                               |
| Número de alunos           | 354899 | 1447                                 | 9987                                |
| Percentual que fez a prova | 72.45% | 70.11%                               | 78.43%                              |

Os alunos de Ciência da Computação (Bacharelado) têm uma mediana de 37.50, que é superior à mediana geral do ENADE, indicando um desempenho acima da mediana nacional. O mesmo se observa para os quartis, onde tanto o 25 quartil quanto o 75 quartil dos alunos de Ciência da Computação (Bacharelado) estão acima dos valores gerais do ENADE, sugerindo que os alunos deste curso têm um desempenho geralmente mais alto na formação geral do que a média dos alunos de todos os cursos. O número de alunos que

realizaram a prova e o percentual que efetivamente compareceu ao exame também são relevantes, pois refletem o engajamento e a representatividade das notas em relação ao total de inscritos. Com 78.43% dos inscritos no curso Ciência da Computação bacharelado e 70.11% na licenciatura fazendo a prova, é possível inferir uma maior adesão dos alunos do bacharelado ao ENADE, o que pode estar associado a diversos fatores institucionais ou motivacionais.

### 4.2.2 Nota bruta nos conhecimentos específicos

A Tabela 4.4 concentra as notas brutas dos conhecimentos específicos, para todos os cursos do ENADE e dos cursos de Ciência da Computação.

Tabela 4.4: Notas dos conhecimentos específicos.

|                            | ENADE  | Ciencia Da Computacao (Licenciatura) | Ciencia Da Computacao (Bacharelado) |
|----------------------------|--------|--------------------------------------|-------------------------------------|
| 25 Quartil                 | 27.80  | 20.80                                | 24.90                               |
| Mediana                    | 38.30  | 28.30                                | 35.80                               |
| 75 Quartil                 | 50.90  | 38.70                                | 46.50                               |
| Número de alunos           | 354899 | 1447                                 | 9987                                |
| Percentual que fez a prova | 72.45% | 70.11%                               | 78.43%                              |

A mediana das notas dos conhecimentos específicos para o curso de Ciência da Computação (Bacharelado) é 35.80, enquanto para a licenciatura é 28.30. Além disso, os quartis para o bacharelado também são mais altos.

Os dados gerais do ENADE para todos os cursos mostram que o 25 Quartil é 27.80, a mediana é 38.30 e o 75 Quartil é 50.90. Quando comparamos esses números com os dos cursos de Ciência da Computação, vemos que a mediana para o bacharelado está um pouco abaixo da mediana geral do ENADE, enquanto a mediana para a licenciatura está significativamente abaixo. Isto pode indicar diferenças no nível de rigor ou foco entre os tipos de curso, ou refletir variações nas bases de conhecimento dos alunos ao ingressarem no ensino superior ou provas com complexidades diferentes.

### 4.2.3 Nota bruta da prova

Continuando com a avaliação dos resultados, a Tabela 4.5 concentra-se a na nota bruta total, abrangendo tanto a formação geral quanto a específica, dos estudantes dos cursos de Ciência da Computação.

A partir dos dados da Tabela 4.5, é possível inferir que o desempenho dos alunos no bacharelado em Ciência da Computação é comparável ao desempenho médio nacional em termos de mediana. No entanto, a variação de desempenho, tanto para mais quanto para menos, é ligeiramente menor do que a mediana geral, como sugerido pelos quartis. Já

Tabela 4.5: Nota bruta da prova.

|                            | ENADE  | Ciencia Da Computacao (Licenciatura) | Ciencia Da Computacao (Bacharelado) |
|----------------------------|--------|--------------------------------------|-------------------------------------|
| 25 Quartil                 | 27.20  | 23.00                                | 26.60                               |
| Mediana                    | 36.60  | 29.40                                | 35.70                               |
| 75 Quartil                 | 47.70  | 37.60                                | 46.10                               |
| Número de alunos           | 354899 | 1447                                 | 9987                                |
| Percentual que fez a prova | 72.45% | 70.11%                               | 78.43%                              |

a licenciatura em Ciência da Computação apresenta um desempenho abaixo da mediana nacional, o que pode levantar questões sobre a eficácia do currículo, a preparação dos alunos ou outros fatores pedagógicos que impactam os resultados do ENADE ou provas com dificuldades diferentes.

#### 4.2.4 Boxplot das notas do curso Ciência da Computação

O *boxplot* [28] também conhecido como diagrama de caixa, é uma representação gráfica estatística que mostra a distribuição de um conjunto de dados. Ele é projetado para fornecer uma visão rápida da localização, dispersão e simetria dos dados, bem como para destacar *outliers*. Aqui estão os componentes principais de um *boxplot* e sua interpretação:

- **Mediana**
- **Quartis**
  - Primeiro quartil (Q1/25º quartil)
  - Terceiro quartil (Q3/75º quartil)
- **Intervalo interquartil (IIQ):** É a diferença entre o terceiro e o primeiro quartil (Q3 e Q1). O IIQ representa a variação média dos dados e é uma medida da dispersão.
- **Bigodes (*whiskers*):** Estendem-se do Q1 até o menor valor dentro de 1,5 IIQ abaixo do Q1 e do Q3 até o maior valor dentro de 1,5 IIQ acima do Q3. Eles fornecem uma indicação visual da variabilidade dos dados.
- ***Outliers*:** São pontos que caem fora dos bigodes. Geralmente, são considerados valores atípicos e são marcados individualmente, muitas vezes com pontos ou asteriscos.

#### Boxplot da nota bruta na formação geral

A Figura 4.8 ilustra um *boxplot* comparativo das notas brutas na formação geral dos alunos do ENADE. O gráfico exibe três categorias distintas: a performance de todos os alunos



do ENADE, seguida pelos alunos dos cursos de Ciência da Computação de Bacharelado e Licenciatura. As provas em branco não foram consideradas na elaboração do boxplot.

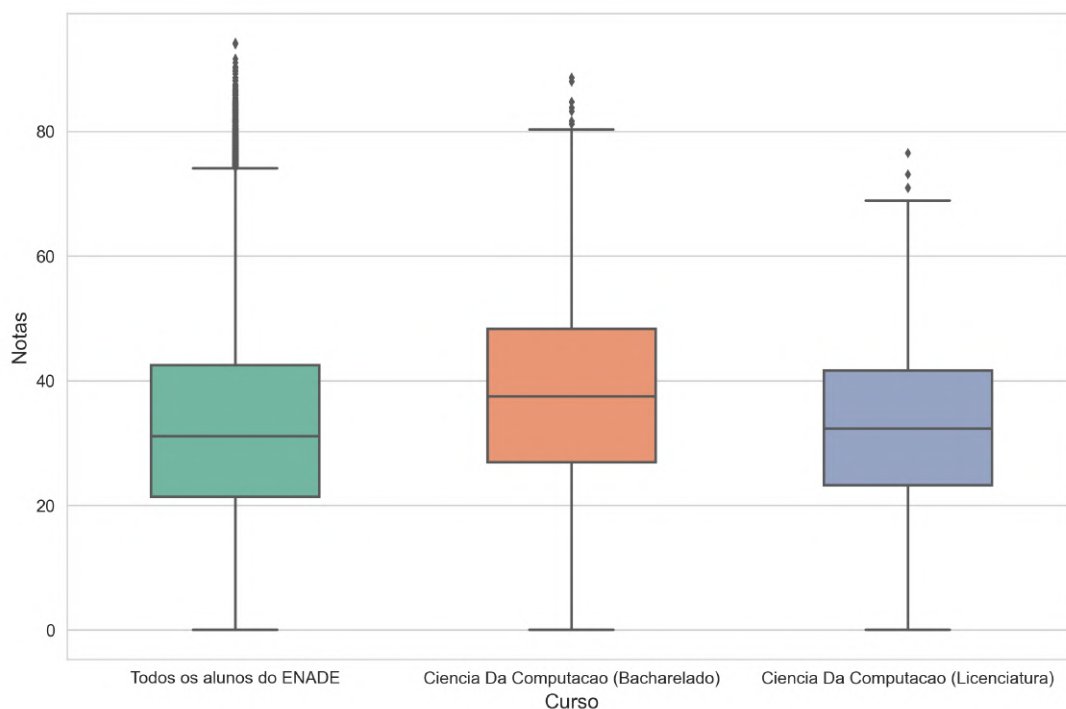


Figura 4.8: Boxplot de comparação de NT\_FG.

Os alunos de Ciência da Computação, tanto de bacharelado quanto de licenciatura, têm medianas de nota mais altas e uma distribuição de notas mais concentrada em comparação com a população geral do ENADE. O grupo de Ciência da Computação (Licenciatura) demonstra uma consistência ligeiramente maior nas notas do que o grupo de Bacharelado, conforme indicado pela caixa mais estreita do *boxplot*. Existem alunos com desempenho excepcionalmente alto em todas as categorias, como mostram os *outliers*.

### ***Boxplot da nota bruta nos conhecimentos específicos***

A Figura 4.8 compara as notas brutas dos conhecimentos específicos entre três grupos; Todos os alunos do ENADE, Ciência da Computação de Bacharelado e Licenciatura.

Em termos de formação específica, os alunos participantes do ENADE parecem ter uma gama mais ampla de desempenho do que os grupos focados em Ciência da Computação, com mais *outliers* no alto desempenho. Os alunos de Ciência da Computação de ambos os cursos (bacharelado e licenciatura) têm uma distribuição de notas mais concentrada em torno da mediana do que o grupo geral do ENADE. A quantidade de *outliers* sugere

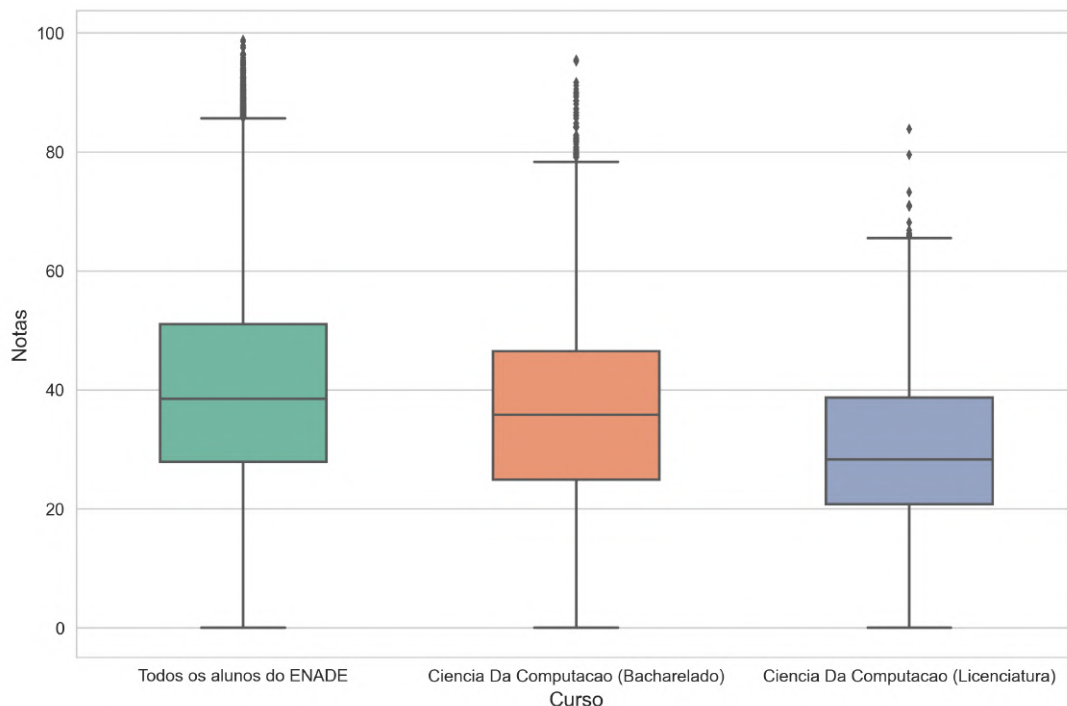


Figura 4.9: Boxplot de comparação de NT\_CE.

que, enquanto a maioria dos alunos tem um desempenho consistente, há uma minoria significativa com desempenho excepcionalmente alto.

### Boxplot da nota bruta da prova

A Figura 4.10 compara as notas brutas da prova do ENADE para três grupos, Todos os alunos do ENADE, Ciência da Computação (Bacharelado) e Ciência da Computação (Licenciatura).

Os alunos de Ciência da Computação, têm uma distribuição de notas ligeiramente mais estreita em comparação com a totalidade dos alunos do ENADE, sugerindo um desempenho mais homogêneo.

Os *outliers* em todos os grupos indicam a presença de alunos com desempenho notavelmente superior ao da maioria.

O perfil de distribuição das notas é distintamente mais compacto para os alunos de Ciência da Computação (Licenciatura), indicando uma concentração de notas em uma faixa mais estreita.

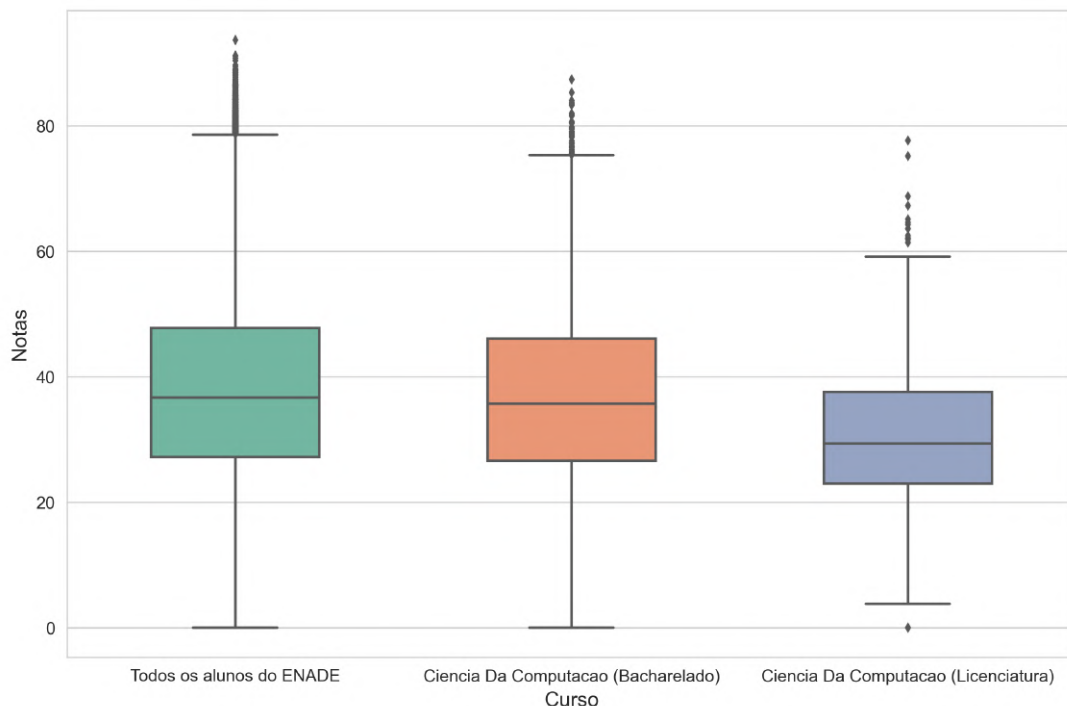


Figura 4.10: Boxplot de comparação de NT\_GER.

### 4.3 Análise dos vetores de notas

Para escolher as faculdades que vão ser analisadas foram selecionadas as 20 faculdades pela coluna “Conceito ENADE (Contínuo)” [29], onde o Conceito ENADE é um indicador de qualidade que avalia os cursos de graduação com base nos resultados alcançados pelos estudantes no Exame Nacional de Desempenho dos Estudantes (ENADE).

Este indicador é calculado e divulgado para cada curso de graduação que foi avaliado, sendo identificado pelo código de curso do Sistema e-MEC, utilizado pelas Instituições de Educação Superior (IES) para a inscrição dos estudantes elegíveis no Exame e para o enquadramento dos cursos em uma das áreas de avaliação do ENADE. A amostra total de instituições de ensino superior que oferecem o curso e participaram do ENADE 2021 é a seguinte:

- Ciência da Computação (Bacharelado): 72 instituições.
- Ciência da Computação (Licenciatura): 13 instituições.

A Tabela 4.6 foi obtida do site do INEP [30], classificando as 20 maiores notas do Conceito ENADE em ordem decrescente, as quais estão destacadas em verde, que tenham o

curso de “Ciência da Computação (Bacharelado)” para a análise <sup>1</sup>. Adicionalmente, os códigos das Instituições de Educação Superior (IES) relevantes são realçados em amarelo para facilitar a avaliação. Esta abordagem permite uma fácil expansão da lista e a utilização dos códigos de IES proporciona um método eficiente para o mapeamento e identificação para os *scripts* de Python que vão gerar as tabelas e gráficos.

Tabela 4.6: Faculdades selecionadas.

| Código da Área | Área de Avaliação     | Grau Acadêmico | Código da IES | Nome da IES*   | Sigla da IES* | Organização Acadêmica | Conceito Enade (Contínuo) |
|----------------|-----------------------|----------------|---------------|--|---------------|-----------------------|---------------------------|
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 575           | UNIVERSIDADE FEDERAL DE MINAS GERAIS                   | UFMG          | Universidade          | 5                         |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 570           | UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE            | UFRN          | Universidade          | 4.535                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 8             | UNIVERSIDADE FEDERAL DE VIÇOSA                         | UFV           | Universidade          | 4.379                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 576           | UNIVERSIDADE FEDERAL DE JUIZ DE FORA                   | UFJF          | Universidade          | 4.349                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 581           | UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL              | UFRGS         | Universidade          | 4.338                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 571           | UNIVERSIDADE FEDERAL DO PARANÁ                         | UFPR          | Universidade          | 4.175                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 577           | UNIVERSIDADE FEDERAL DE ALAGOAS                        | UFAL          | Universidade          | 4.143                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 573           | UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO                 | UFES          | Universidade          | 4.119                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 585           | UNIVERSIDADE FEDERAL DE SANTA CATARINA                 | UFSC          | Universidade          | 4.088                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 4925          | FUNDAÇÃO UNIVERSIDADE FEDERAL DO ABC                   | UFABC         | Universidade          | 4.064                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 54            | UNIVERSIDADE ESTADUAL DE CAMPINAS                      | UNICAMP       | Universidade          | 4.047                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 591           | UNIVERSIDADE FEDERAL DE SÃO PAULO                      | UNIFESP       | Universidade          | 3.995                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 7             | UNIVERSIDADE FEDERAL DE SÃO CARLOS                     | UFSCAR        | Universidade          | 3.989                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 56            | UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO | UNESP         | Universidade          | 3.989                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 572           | UNIVERSIDADE FEDERAL FLUMINENSE                        | UFF           | Universidade          | 3.971                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 586           | UNIVERSIDADE FEDERAL DO RIO DE JANEIRO                 | UFRJ          | Universidade          | 3.826                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 582           | UNIVERSIDADE FEDERAL DE SANTA MARIA                    | UFSM          | Universidade          | 3.819                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 271           | UNIVERSIDADE DO OESTE PAULISTA                         | UNOESTE       | Universidade          | 3.818                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 2             | UNIVERSIDADE DE BRASÍLIA                               | UNB           | Universidade          | 3.816                     |
| 4004           | CIÊNCIA DA COMPUTAÇÃO | Bacharelado    | 57            | UNIVERSIDADE ESTADUAL DE MARINGÁ                       | UEM           | Universidade          | 3.798                     |

Para a análise, foram selecionadas as mesmas 20 instituições de ensino superior para ambos os Ciência da Computação Bacharelado e Licenciatura. Essa escolha visa evidenciar a discrepância entre a demanda de alunos e a oferta das universidades, bem como possibilita uma análise da qualidade de ensino ao comparar as mesmas 20 instituições, sendo que entre elas apenas 5 oferecem cursos tanto para o bacharelado quanto para a licenciatura.

### 4.3.1 Análise de número total de alunos

Foi feita uma análise do número total de alunos inscritos e número de alunos que realizaram a prova.

Para ser classificado como participante que completou a prova, o aluno deve ter alcançado uma nota igual ou superior a zero. A UnB vai ser destacada na lista para melhor visualização. Os cursos analisados serão:

- Ciência da Computação (Bacharelado)
- Ciência da Computação (Licenciatura)

<sup>1</sup>As instituições CENTRO UNIVERSITÁRIO JORGE AMADO, FACULDADE CESAR e UNIVERSIDADE DO VALE DO RIO DOS SINOS estavam entre as 20 melhores classificadas, mas não foram selecionadas para a análise devido ao número insuficiente de participantes, o que comprometeria a relevância estatística dos resultados.

## Ciência da Computação (Bacharelado)

A Tabela 4.7 lista as 20 instituições de ensino superior selecionadas anteriormente que ofereçam o curso de Ciência da Computação (Bacharelado).

Tabela 4.7: Faculdades selecionadas para Ciência da Computação (Bacharelado).

| Nome da Instituição                                    | Número total de alunos inscritos | Número de alunos que realizaram a prova |
|--|----------------------------------|---|
| UNIVERSIDADE ESTADUAL PAULISTA JULIO DE MESQUITA FILHO | 194                              | 176                                     |
| FUNDAÇÃO UNIVERSIDADE FEDERAL DO ABC                   | 185                              | 94                                      |
| UNIVERSIDADE FEDERAL FLUMINENSE                        | 169                              | 138                                     |
| UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL              | 154                              | 120                                     |
| UNIVERSIDADE FEDERAL DE SÃO CARLOS                     | 148                              | 133                                     |
| UNIVERSIDADE FEDERAL DO RIO DE JANEIRO                 | 135                              | 122                                     |
| UNIVERSIDADE FEDERAL DO PARANÁ                         | 113                              | 72                                      |
| UNIVERSIDADE DE BRASÍLIA                               | 112                              | 85                                      |
| UNIVERSIDADE FEDERAL DE ALAGOAS                        | 111                              | 98                                      |
| UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO                 | 109                              | 80                                      |
| UNIVERSIDADE FEDERAL DE SANTA CATARINA                 | 99                               | 79                                      |
| UNIVERSIDADE FEDERAL DE JUIZ DE FORA                   | 79                               | 56                                      |
| UNIVERSIDADE FEDERAL DE VICOSA                         | 66                               | 62                                      |
| UNIVERSIDADE FEDERAL DE MINAS GERAIS                   | 52                               | 48                                      |
| UNIVERSIDADE FEDERAL DE SÃO PAULO                      | 50                               | 40                                      |
| UNIVERSIDADE FEDERAL DE SANTA MARIA                    | 47                               | 42                                      |
| UNIVERSIDADE ESTADUAL DE MARINGÁ                       | 46                               | 41                                      |
| UNIVERSIDADE ESTADUAL DE CAMPINAS                      | 43                               | 34                                      |
| UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE            | 32                               | 23                                      |
| UNIVERSIDADE DO OESTE PAULISTA                         | 25                               | 18                                      |

Observa-se uma variação considerável no número de alunos inscritos e que realizaram a prova entre as instituições. A instituição com o maior número de inscritos é a Universidade Estadual Paulista Júlio de Mesquita Filho com 194 inscritos, dos quais 176 realizaram a prova. Em contraste, a Universidade do Oeste Paulista tem o menor número de inscritos, com 25, e 18 alunos realizaram a prova. A maioria das instituições tem uma alta taxa de comparecimento para a prova, o que pode ser indicativo de um forte engajamento dos alunos com o processo de avaliação do ENADE.

O ENADE contou com 1969 participantes das 20 instituições de ensino superior selecionadas no curso de Ciência da Computação - Bacharelado. Destes, 1561 realizaram a prova. Assim, o percentual da amostra foi de 15,45%. Isto garante uma confiança no resultado acima de 15,64% [31].

## Ciência da Computação (Licenciatura)

A Tabela 4.8 filtra as mesmas 20 faculdades anteriores, mas apenas 5 faculdades oferecem o curso de Ciência da Computação (Licenciatura), esse dado pode sugerir que é menos

comum ou tem menos demanda em comparação ao bacharelado, refletindo em menos alunos inscritos e participantes do ENADE neste nível de formação.

Tabela 4.8: Faculdades selecionadas para Ciência da Computação (Licenciatura).

| Nome da Instituição                  | Número total de alunos inscritos | Número de alunos que realizaram a prova |
|--------------------------------------|----------------------------------|---|
| UNIVERSIDADE DE BRASÍLIA             | 60                               | 39                                      |
| UNIVERSIDADE FEDERAL DE JUIZ DE FORA | 36                               | 26                                      |
| UNIVERSIDADE FEDERAL DO PARANÁ       | 14                               | 12                                      |
| UNIVERSIDADE FEDERAL FLUMINENSE      | 14                               | 10                                      |
| UNIVERSIDADE FEDERAL DE SANTA MARIA  | 4                                | 3                                       |

A UnB tem o maior número de alunos inscritos (60) e que realizaram a prova (39) entre as instituições listadas para o curso de licenciatura em Ciência da Computação. Isto é significativo, especialmente se considerarmos que nem todas as universidades oferecem esse curso, o que pode explicar a lista mais curta em comparação com o bacharelado. O ENADE contou com 128 alunos das 20 instituições de ensino superior selecionadas no curso de Ciência da Computação - Bacharelado. Destes, 90 alunos realizaram a prova. Assim, o percentual da amostra foi de 1,01%. Isto garante uma confiança no resultado acima de 0,90%.

### 4.3.2 Análise da parte objetiva na formação geral

O componente de formação geral (FG) da parte objetiva é formada por um conjunto de 8 questões que são de múltipla escolha. O vetor de notas **DS\_VT\_ACE\_OFG** representa os acertos da parte objetiva na formação geral, ele é de tipo *string* e tem 8 caracteres. Cada caractere representa o acerto ou erro da questão da parte objetiva, o vetor foi dividido na etapa de preparação de dados e segue o seguinte formato “forma\_geral\_q\_x”, onde *X* vai representar o número da questão objetiva. Foi feita a análise da taxa de acertos por questão da parte objetiva na formação geral do ENADE para o curso de Ciência da Computação (Bacharelado) e Ciência da Computação (Licenciatura) para as 20 faculdades selecionadas anteriormente.

#### Ciência da Computação (Bacharelado)

A Tabela 4.9 e Tabela 4.10 apresentam a taxa de acertos por questão da parte objetiva na formação geral para o curso de Ciência da Computação (Bacharelado) para as 20 faculdades.

Tabela 4.9: Percentual de acertos nas questões objetivas (FG) 1 a 4.

| Posição | Sigla da IES | forma_geral_q_1 | Posição | Sigla da IES | forma_geral_q_2 | Posição | Sigla da IES | forma_geral_q_3 | Posição | Sigla da IES | forma_geral_q_4 |
|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|
| 1       | UFV          | 72.58%          | 1       | UNIFESP      | 42.50%          | 1       | UFF          | 77.54%          | 1       | UFMG         | 60.42%          |
| 2       | UFMG         | 70.83%          | 2       | UFES         | 38.75%          | 2       | UNICAMP      | 76.47%          | 2       | UFES         | 60.00%          |
| 3       | UEM          | 70.73%          | 3       | UFPR         | 37.50%          | 3       | UFRGS        | 75.83%          | 3       | UFRJ         | 59.84%          |
| 4       | UFF          | 70.29%          | 4       | <b>UNB</b>   | <b>36.47%</b>   | 4       | UNIFESP      | 75.00%          | 4       | UFJF         | 58.93%          |
| 5       | UNIFESP      | 67.50%          | 5       | UFABC        | 35.11%          | 5       | UFJF         | 75.00%          | 5       | <b>UNB</b>   | <b>58.82%</b>   |
| 6       | UFABC        | 65.96%          | 6       | UFMG         | 33.33%          | 6       | UFV          | 74.19%          | 6       | UFSCAR       | 56.39%          |
| 7       | UFSC         | 65.82%          | 7       | UFSC         | 31.65%          | 7       | UFRN         | 73.91%          | 7       | UFRGS        | 55.00%          |
| 8       | UFRN         | 65.22%          | 8       | UFF          | 31.16%          | 8       | UFSM         | 73.81%          | 8       | UFSM         | 54.76%          |
| 9       | UNICAMP      | 64.71%          | 9       | UFSM         | 30.95%          | 9       | UFMG         | 72.92%          | 9       | UNICAMP      | 52.94%          |
| 10      | UNESP        | 64.20%          | 10      | UFRGS        | 30.83%          | 10      | UFSCAR       | 72.18%          | 10      | UFF          | 52.90%          |
| 11      | UFSCAR       | 63.91%          | 11      | UFSCAR       | 28.57%          | 11      | <b>UNB</b>   | <b>71.76%</b>   | 11      | UFPR         | 52.78%          |
| 12      | UFES         | 62.50%          | 12      | UFJF         | 28.57%          | 12      | UFABC        | 71.28%          | 12      | UNOESTE      | 50.00%          |
| 13      | UFSM         | 61.90%          | 13      | UFRJ         | 27.87%          | 13      | UEM          | 70.73%          | 13      | UNIFESP      | 50.00%          |
| 14      | UFPR         | 61.11%          | 14      | UFV          | 27.42%          | 14      | UFRJ         | 69.67%          | 14      | UFABC        | 50.00%          |
| 15      | <b>UNB</b>   | <b>60.00%</b>   | 15      | UNESP        | 27.27%          | 15      | UFSC         | 68.35%          | 15      | UEM          | 46.34%          |
| 16      | UFRJ         | 59.84%          | 16      | UEM          | 26.83%          | 16      | UFES         | 67.50%          | 16      | UFAL         | 44.90%          |
| 17      | UFJF         | 58.93%          | 17      | UNICAMP      | 26.47%          | 17      | UNOESTE      | 66.67%          | 17      | UNESP        | 43.75%          |
| 18      | UFRGS        | 58.33%          | 18      | UFRN         | 26.09%          | 18      | UFPR         | 65.28%          | 18      | UFRN         | 43.48%          |
| 19      | UFAL         | 51.02%          | 19      | UFAL         | 23.47%          | 19      | UFAL         | 63.27%          | 19      | UFV          | 41.94%          |
| 20      | UNOESTE      | 38.89%          | 20      | UNOESTE      | 22.22%          | 20      | UNESP        | 60.80%          | 20      | UFSC         | 41.77%          |

Tabela 4.10: Percentual de acertos nas questões objetivas (FG) 5 a 8.

| Posição | Sigla da IES | forma_geral_q_5 | Posição | Sigla da IES | forma_geral_q_6 | Posição | Sigla da IES | forma_geral_q_7 | Posição | Sigla da IES | forma_geral_q_8 |
|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|
| 1       | UFMG         | 72.92%          | 1       | UNICAMP      | 88.24%          | 1       | UFJF         | 89.29%          | 1       | UFMG         | 35.42%          |
| 2       | <b>UNB</b>   | <b>71.76%</b>   | 2       | UFSCAR       | 83.46%          | 2       | UNICAMP      | 88.24%          | 2       | UFRGS        | 35.00%          |
| 3       | UFABC        | 68.09%          | 3       | UFMG         | 83.33%          | 3       | UFSM         | 88.10%          | 3       | UFF          | 31.88%          |
| 4       | UFRN         | 65.22%          | 4       | UFABC        | 80.85%          | 4       | UFSC         | 84.81%          | 4       | <b>UNB</b>   | <b>31.76%</b>   |
| 5       | UNIFESP      | 65.00%          | 5       | UFRGS        | 77.50%          | 5       | UFABC        | 84.04%          | 5       | UNOESTE      | 27.78%          |
| 6       | UFV          | 64.52%          | 6       | UNIFESP      | 77.50%          | 6       | UNOESTE      | 83.33%          | 6       | UFRJ         | 27.05%          |
| 7       | UFSM         | 64.29%          | 7       | UFSC         | 75.95%          | 7       | UFPR         | 83.33%          | 7       | UNICAMP      | 26.47%          |
| 8       | UFRGS        | 64.17%          | 8       | UNESP        | 75.00%          | 8       | UFV          | 82.26%          | 8       | UFV          | 25.81%          |
| 9       | UEM          | 63.41%          | 9       | UFRN         | 73.91%          | 9       | UFMG         | 81.25%          | 9       | UFSC         | 25.32%          |
| 10      | UFES         | 62.50%          | 10      | UFV          | 72.58%          | 10      | UFRJ         | 79.51%          | 10      | UEM          | 24.39%          |
| 11      | UFSC         | 62.03%          | 11      | UFAL         | 72.45%          | 11      | UFRGS        | 79.17%          | 11      | UFAL         | 23.47%          |
| 12      | UFF          | 60.87%          | 12      | UFPR         | 72.22%          | 12      | UFRN         | 78.26%          | 12      | UNIFESP      | 22.50%          |
| 13      | UFPR         | 59.72%          | 13      | UFJF         | 71.43%          | 13      | UNIFESP      | 77.50%          | 13      | UFES         | 22.50%          |
| 14      | UFAL         | 57.14%          | 14      | UFF          | 70.29%          | 14      | UFF          | 76.81%          | 14      | UFABC        | 22.34%          |
| 15      | UNICAMP      | 55.88%          | 15      | UFES         | 68.75%          | 15      | UNESP        | 76.70%          | 15      | UFPR         | 22.22%          |
| 16      | UFRJ         | 54.92%          | 16      | <b>UNB</b>   | <b>68.24%</b>   | 16      | <b>UNB</b>   | <b>71.76%</b>   | 16      | UFSCAR       | 20.30%          |
| 17      | UNESP        | 52.84%          | 17      | UNOESTE      | 66.67%          | 17      | UFES         | 71.25%          | 17      | UNESP        | 19.89%          |
| 18      | UFSCAR       | 51.13%          | 18      | UFRJ         | 64.75%          | 18      | UEM          | 70.73%          | 18      | UFJF         | 17.86%          |
| 19      | UFJF         | 39.29%          | 19      | UFSM         | 54.76%          | 19      | UFSCAR       | 70.68%          | 19      | UFSM         | 14.29%          |
| 20      | UNOESTE      | 38.89%          | 20      | UEM          | 53.66%          | 20      | UFAL         | 69.39%          | 20      | UFRN         | 8.70%           |

A Questão 7 tem consistentemente a maior taxa de acertos entre as instituições em ambas as tabelas. Isso pode indicar que o conteúdo dessa questão é melhor dominado pelos alunos ou que a questão foi mais clara e direta. Por outro lado, a Questão 8 destaca-se por ter as menores taxas de acertos.

A informação da Tabela 4.11 é um resumo consolidado dos dados de desempenho das instituições nas questões. Destaca o melhor e o pior desempenho por questão, indicando quais instituições tiveram as maiores e menores taxas de acerto.

Tabela 4.11: Comparação de desempenho (FG) questões objetivas 1 a 8.

| Questão         | Melhor Desempenho                    | Pior Desempenho  |
|-----------------|--------------------------------------|--|
| forma_geral_q_1 | UNIVERSIDADE FEDERAL DE VICOSA       | UNIVERSIDADE DO OESTE PAULISTA                         |
| forma_geral_q_2 | UNIVERSIDADE FEDERAL DE SAO PAULO    | UNIVERSIDADE DO OESTE PAULISTA                         |
| forma_geral_q_3 | UNIVERSIDADE FEDERAL FLUMINENSE      | UNIVERSIDADE ESTADUAL PAULISTA JULIO DE MESQUITA FILHO |
| forma_geral_q_4 | UNIVERSIDADE FEDERAL DE MINAS GERAIS | UNIVERSIDADE FEDERAL DE SANTA CATARINA                 |
| forma_geral_q_5 | UNIVERSIDADE FEDERAL DE MINAS GERAIS | UNIVERSIDADE DO OESTE PAULISTA                         |
| forma_geral_q_6 | UNIVERSIDADE ESTADUAL DE CAMPINAS    | UNIVERSIDADE ESTADUAL DE MARINGA                       |
| forma_geral_q_7 | UNIVERSIDADE FEDERAL DE JUIZ DE FORA | UNIVERSIDADE FEDERAL DE ALAGOAS                        |
| forma_geral_q_8 | UNIVERSIDADE FEDERAL DE MINAS GERAIS | UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE            |

A tabela mostra que a Universidade Federal de Minas Gerais teve o melhor desempenho nas questões 4, 5 e 8, o que pode indicar uma forte preparação ou alinhamento do currículo desta universidade com o conteúdo abordado nessas questões específicas. Por outra parte a Universidade do Oeste Paulista aparece como a instituição com o pior desempenho nas questões 1, 2 e 5, o que pode indicar áreas onde os alunos dessa universidade podem necessitar de um suporte adicional. A UnB não apareceu na tabela como melhor e o pior desempenho o que pode dizer que os alunos tiveram um desempenho médio nas questões da parte objetiva na formação geral.

Para uma visão mais aprofundada a Tabela 4.12 apresenta as respostas dadas pelos alunos da UnB do curso Ciência da Computação (Bacharelado). A tabela está dividida em colunas das questões de 1 a 8 e linhas que representam as opções de resposta marcadas (A, B, C, D, E, opção em branco e várias onde o aluno marcou várias opções).  $Z$  representa que a questão foi excluída devido à anulação e  $X$  que a questão foi excluída devido ao coeficiente ponto-bisserial  $< 0,20$ . Ponto bisserial [32] é uma medida de correlação que avalia quão bem uma questão individual consegue distinguir entre alunos de alto e baixo desempenho no exame. Um coeficiente ponto-bisserial inferior a 0,20 geralmente sugere que a questão não é eficaz em discriminar entre alunos com diferentes níveis de habilidade ou conhecimento. Cada linha tem destacada a opção mais marcada com a cor azul, a última coluna indica o gabarito correto para cada questão.



Tabela 4.12: Respostas da UnB para as questões objetivas (FG) Bacharelado.

| Questões                    | Opção A | Opção B | Opção C | Opção D | Opção E | Opção em branco | Várias | Gabarito |
|-----------------------------|---------|---------|---------|---------|---------|-----------------|--------|----------|
| forma_geral_alternativa_q_1 | 2       | 15      | 15      | 1       | 51      | 0               | 1      | E        |
| forma_geral_alternativa_q_2 | 24      | 23      | 31      | 6       | 1       | 0               | 0      | C        |
| forma_geral_alternativa_q_3 | 10      | 61      | 0       | 4       | 10      | 0               | 0      | B        |
| forma_geral_alternativa_q_4 | 25      | 50      | 3       | 4       | 3       | 0               | 0      | B        |
| forma_geral_alternativa_q_5 | 61      | 1       | 21      | 1       | 1       | 0               | 0      | A        |
| forma_geral_alternativa_q_6 | 58      | 7       | 9       | 9       | 2       | 0               | 0      | A        |
| forma_geral_alternativa_q_7 | 14      | 3       | 61      | 3       | 4       | 0               | 0      | C        |
| forma_geral_alternativa_q_8 | 17      | 12      | 6       | 27      | 23      | 0               | 0      | D        |

Os alunos da UnB tiveram um bom desempenho em geral, com a maioria das respostas coincidindo com o gabarito. Nas questões 2 e 8, muitos alunos não marcaram a opção correta.

### Ciência da Computação (Licenciatura)

A Tabela 4.13 e Tabela 4.14 apresentam a taxa de acertos por questão da parte objetiva na formação geral para o curso de Ciência da Computação (Licenciatura) para as 5 faculdades que ofertavam o curso das 20 faculdades selecionadas.

Tabela 4.13: Percentual de acertos nas questões objetivas 1 a 4.

| Posição | Sigla da IES | forma_geral_q_1 | Posição | Sigla da IES | forma_geral_q_2 | Posição | Sigla da IES | forma_geral_q_3 | Posição | Sigla da IES | forma_geral_q_4 |
|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|
| 1       | UFPR         | 66.67%          | 1       | UNB          | 41.03%          | 1       | UFPR         | 83.33%          | 1       | UFSM         | 100.00%         |
| 2       | UFF          | 50.00%          | 2       | UFPR         | 33.33%          | 2       | UFF          | 70.00%          | 2       | UNB          | 38.46%          |
| 3       | UFJF         | 50.00%          | 3       | UFF          | 30.00%          | 3       | UNB          | 66.67%          | 3       | UFPR         | 33.33%          |
| 4       | UNB          | 48.72%          | 4       | UFJF         | 26.92%          | 4       | UFSM         | 66.67%          | 4       | UFJF         | 26.92%          |
| 5       | UFSM         | 33.33%          | 5       | UFSM         | 0.00%           | 5       | UFJF         | 65.38%          | 5       | UFF          | 20.00%          |

Tabela 4.14: Percentual de acertos nas questões objetivas 5 a 8.

| Posição | Sigla da IES | forma_geral_q_5 | Posição | Sigla da IES | forma_geral_q_6 | Posição | Sigla da IES | forma_geral_q_7 | Posição | Sigla da IES | forma_geral_q_8 |
|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|
| 1       | UFSM         | 100.00%         | 1       | UFPR         | 75.00%          | 1       | UFPR         | 66.67%          | 1       | UFSM         | 66.67%          |
| 2       | UNB          | 82.05%          | 2       | UFF          | 60.00%          | 2       | UNB          | 64.10%          | 2       | UNB          | 30.77%          |
| 3       | UFJF         | 65.38%          | 3       | UFJF         | 57.69%          | 3       | UFJF         | 61.54%          | 3       | UFPR         | 25.00%          |
| 4       | UFF          | 60.00%          | 4       | UNB          | 56.41%          | 4       | UFF          | 50.00%          | 4       | UFJF         | 11.54%          |
| 5       | UFPR         | 58.33%          | 5       | UFSM         | 33.33%          | 5       | UFSM         | 33.33%          | 5       | UFF          | 10.00%          |

A Questão 5 parece ser aquela com a maior taxa de acertos entre as instituições, enquanto a Questão 8 é a que tem consistentemente a menor taxa de acertos entre todas as instituições<sup>2</sup>.

<sup>2</sup>A Universidade Federal de Santa Maria (UFSM) teve apenas 3 participantes

A informação da Tabela 4.15 é um resumo consolidado dos dados de desempenho das instituições nas questões. Destaca o melhor e o pior desempenho por questão, indicando quais instituições tiveram as maiores e menores taxas de acerto.

Tabela 4.15: Comparação de desempenho (FG) questões objetivas 1 a 8.

| Questão         | Melhor Desempenho                   | Pior Desempenho                      |
|-----------------|-------------------------------------|--------------------------------------|
| forma_geral_q_1 | UNIVERSIDADE FEDERAL DO PARANA      | UNIVERSIDADE FEDERAL DE SANTA MARIA  |
| forma_geral_q_2 | UNIVERSIDADE DE BRASÍLIA            | UNIVERSIDADE FEDERAL DE SANTA MARIA  |
| forma_geral_q_3 | UNIVERSIDADE FEDERAL DO PARANA      | UNIVERSIDADE FEDERAL DE JUIZ DE FORA |
| forma_geral_q_4 | UNIVERSIDADE FEDERAL DE SANTA MARIA | UNIVERSIDADE FEDERAL FLUMINENSE      |
| forma_geral_q_5 | UNIVERSIDADE FEDERAL DE SANTA MARIA | UNIVERSIDADE FEDERAL DO PARANA       |
| forma_geral_q_6 | UNIVERSIDADE FEDERAL DO PARANA      | UNIVERSIDADE FEDERAL DE SANTA MARIA  |
| forma_geral_q_7 | UNIVERSIDADE FEDERAL DO PARANA      | UNIVERSIDADE FEDERAL DE SANTA MARIA  |
| forma_geral_q_8 | UNIVERSIDADE FEDERAL DE SANTA MARIA | UNIVERSIDADE FEDERAL FLUMINENSE      |

A UnB tem um ponto forte na Questão 2, mas não aparece como tendo o pior desempenho em nenhuma questão, o que sugere um desempenho geralmente consistente.

A Tabela 4.16 representa as respostas dadas pelos alunos da UnB do curso Ciência da Computação (Licenciatura). Cada linha tem destacada a opção mais marcada com a cor azul, a última coluna indica o gabarito correto para cada questão.

Tabela 4.16: Respostas da UnB para as questões objetivas (FG) Licenciatura.

| Questões                    | Opção A | Opção B | Opção C | Opção D | Opção E | Opção em branco | Várias | Gabarito |
|-----------------------------|---------|---------|---------|---------|---------|-----------------|--------|----------|
| forma_geral_alternativa_q_1 | 2       | 8       | 6       | 4       | 19      | 0               | 0      | E        |
| forma_geral_alternativa_q_2 | 15      | 6       | 16      | 1       | 1       | 0               | 0      | C        |
| forma_geral_alternativa_q_3 | 1       | 26      | 0       | 3       | 9       | 0               | 0      | B        |
| forma_geral_alternativa_q_4 | 23      | 15      | 0       | 0       | 1       | 0               | 0      | B        |
| forma_geral_alternativa_q_5 | 32      | 0       | 6       | 0       | 1       | 0               | 0      | A        |
| forma_geral_alternativa_q_6 | 22      | 8       | 3       | 0       | 6       | 0               | 0      | A        |
| forma_geral_alternativa_q_7 | 7       | 1       | 25      | 4       | 2       | 0               | 0      | C        |
| forma_geral_alternativa_q_8 | 6       | 5       | 7       | 12      | 9       | 0               | 0      | D        |

Podemos notar que os alunos se saíram bem na maioria das questões, especialmente nas questões 1, 5, 6 onde a maioria escolheu a resposta correta. Os alunos demonstraram uma tendência a selecionar a uma resposta para esta questão.

### 4.3.3 Análise da parte objetiva de conhecimentos específicos

O componente de conhecimentos específicos (CE) da parte objetiva é formada por um conjunto de 27 questões que são de múltipla escolha.

O vetor de notas **DS\_VT\_ACE\_OCE** representa os acertos da parte objetiva de conhecimentos específicos, ele é de tipo *string* e tem 27 caracteres. Cada caracter representa a opção marcada pelo aluno na questão da parte objetiva, o vetor foi dividido na etapa de preparação de dados e segue o seguinte formato “forma\_espec\_q\_x”, onde X vai representar o número da questão objetiva. Foi feita a análise da taxa de acertos por questão da parte objetiva na formação geral do ENADE para o curso de Ciência da Computação (Bacharelado) e Ciência da Computação (Licenciatura) para as 20 faculdades selecionadas anteriormente.

## Ciência da Computação (Bacharelado)

As Tabelas 4.17 a 4.20 apresentam a taxa de acertos por questão da parte objetiva do Componente Específico para o curso de Ciência da Computação (Bacharelado) para as 20 faculdades. As seguintes questões foram anuladas:

- **Questões (21, 25)** - Anulada pela comissão.
- **Questões (4, 5, 9, 13, 17, 24)** - Anulada pelo índice de discriminação (correlação pontobisserial < 0,20).

Tabela 4.17: Percentual de acertos nas questões (CE) 1, 2, 3, 6, 7.

| Posição | Sigla da IES | forma_espec_q_1 | Posição | Sigla da IES | forma_espec_q_2 | Posição | Sigla da IES | forma_espec_q_3 | Posição | Sigla da IES | forma_espec_q_6 | Posição | Sigla da IES | forma_espec_q_7 |
|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|
| 1       | UFPR         | 52.78%          | 1       | UFRN         | 91.30%          | 1       | UNIFESP      | 97.50%          | 1       | UFPR         | 50.00%          | 1       | UNOESTE      | 77.78%          |
| 2       | UFSCAR       | 51.88%          | 2       | UFRGS        | 75.83%          | 2       | UFRN         | 95.65%          | 2       | <b>UNB</b>   | <b>42.35%</b>   | 2       | UFABC        | 73.40%          |
| 3       | UFJF         | 51.79%          | 3       | UEM          | 75.61%          | 3       | UFMS         | 92.86%          | 3       | UFSC         | 39.24%          | 3       | UFSCAR       | 69.17%          |
| 4       | UNIFESP      | 50.00%          | 4       | <b>UNB</b>   | <b>67.06%</b>   | 4       | UFMG         | 91.67%          | 4       | UFMG         | 37.50%          | 4       | UFRGS        | 69.17%          |
| 5       | UFSC         | 46.84%          | 5       | UNOESTE      | 66.67%          | 5       | UEM          | 90.24%          | 5       | UFES         | 37.50%          | 5       | UFES         | 65.00%          |
| 6       | UEM          | 46.34%          | 6       | UFSCAR       | 64.66%          | 6       | UFRGS        | 89.17%          | 6       | UFV          | 37.10%          | 6       | <b>UNB</b>   | <b>64.71%</b>   |
| 7       | UFES         | 46.25%          | 7       | UFMG         | 64.58%          | 7       | UNOESTE      | 88.89%          | 7       | UF RJ        | 36.89%          | 7       | UFSC         | 64.56%          |
| 8       | UFMS         | 45.24%          | 8       | UFV          | 62.90%          | 8       | UFPR         | 88.89%          | 8       | UFRGS        | 36.67%          | 8       | UFF          | 63.77%          |
| 9       | UNOESTE      | 44.44%          | 9       | UF RJ        | 61.48%          | 9       | UNICAMP      | 88.24%          | 9       | UFAL         | 35.71%          | 9       | UFMG         | 60.42%          |
| 10      | UFRGS        | 44.17%          | 10      | UFABC        | 59.57%          | 10      | UFABC        | 87.23%          | 10      | UFMS         | 35.71%          | 10      | UNESP        | 60.23%          |
| 11      | UFAL         | 42.86%          | 11      | UFPR         | 58.33%          | 11      | UFV          | 87.10%          | 11      | UFJF         | 33.93%          | 11      | UFJF         | 58.93%          |
| 12      | UFABC        | 42.55%          | 12      | UFMS         | 57.14%          | 12      | UF RJ        | 85.25%          | 12      | UEM          | 31.71%          | 12      | UFPR         | 58.33%          |
| 13      | UNICAMP      | 41.18%          | 13      | UFSC         | 56.96%          | 13      | UFJF         | 83.93%          | 13      | UNESP        | 31.25%          | 13      | UFRN         | 56.52%          |
| 14      | UNESP        | 40.34%          | 14      | UNICAMP      | 55.88%          | 14      | UFSCAR       | 81.95%          | 14      | UFF          | 30.43%          | 14      | UEM          | 56.10%          |
| 15      | UFV          | 40.32%          | 15      | UFES         | 53.75%          | 15      | <b>UNB</b>   | <b>80.00%</b>   | 15      | UNOESTE      | 27.78%          | 15      | UF RJ        | 53.28%          |
| 16      | UF RJ        | 39.34%          | 16      | UFF          | 52.90%          | 16      | UFF          | 79.71%          | 16      | UFRN         | 26.09%          | 16      | UFV          | 53.23%          |
| 17      | UFRN         | 39.13%          | 17      | UNESP        | 52.84%          | 17      | UNESP        | 79.55%          | 17      | UNIFESP      | 25.00%          | 17      | UNIFESP      | 52.50%          |
| 18      | <b>UNB</b>   | <b>32.94%</b>   | 18      | UFJF         | 50.00%          | 18      | UFSC         | 74.68%          | 18      | UFSCAR       | 24.81%          | 18      | UFMS         | 52.39%          |
| 19      | UFF          | 32.61%          | 19      | UNIFESP      | 50.00%          | 19      | UFES         | 72.50%          | 19      | UNICAMP      | 23.53%          | 19      | UNICAMP      | 47.06%          |
| 20      | UFMG         | 27.08%          | 20      | UFAL         | 38.78%          | 20      | UFAL         | 68.37%          | 20      | UFABC        | 21.28%          | 20      | UFAL         | 38.78%          |

Tabela 4.18: Percentual de acertos nas questões (CE) 8, 10, 11, 12, 14.

| Posição | Sigla da IES | forma_espec_q_8 | Posição | Sigla da IES | forma_espec_q_10 | Posição | Sigla da IES | forma_espec_q_11 | Posição | Sigla da IES | forma_espec_q_12 | Posição | Sigla da IES | forma_espec_q_14 |
|---------|--------------|-----------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|
| 1       | UFRN         | 86.96%          | 1       | UNOESTE      | 66.67%           | 1       | UFRN         | 60.87%           | 1       | UFRN         | 95.65%           | 1       | UNOESTE      | 88.89%           |
| 2       | UNIFESP      | 82.50%          | 2       | UNIFESP      | 60.00%           | 2       | UFRGS        | 57.50%           | 2       | UFMG         | 87.50%           | 2       | UFSC         | 78.48%           |
| 3       | UNICAMP      | 82.35%          | 3       | UNICAMP      | 58.82%           | 3       | UNICAMP      | 55.88%           | 3       | UEM          | 78.05%           | 3       | UNIFESP      | 75.00%           |
| 4       | UFF          | 80.43%          | 4       | UFSCAR       | 57.89%           | 4       | UFV          | 54.84%           | 4       | UFV          | 77.42%           | 4       | UFF          | 73.19%           |
| 5       | UFABC        | 78.72%          | 5       | UFMS         | 57.14%           | 5       | UFMG         | 54.17%           | 5       | UNIFESP      | 75.00%           | 5       | UFMG         | 72.92%           |
| 6       | UFMS         | 78.57%          | 6       | UFMG         | 56.25%           | 6       | UFSCAR       | 52.63%           | 6       | UFPR         | 75.00%           | 6       | UFV          | 72.58%           |
| 7       | UNOESTE      | 77.78%          | 7       | UFSC         | 55.70%           | 7       | UEM          | 48.78%           | 7       | UFABC        | 74.47%           | 7       | UFJF         | 71.43%           |
| 8       | UF RJ        | 77.05%          | 8       | UFRGS        | 53.33%           | 8       | UFPR         | 48.61%           | 8       | UFRGS        | 74.17%           | 8       | UFRN         | 69.57%           |
| 9       | UNESP        | 76.70%          | 9       | UFRN         | 52.17%           | 9       | <b>UNB</b>   | <b>47.06%</b>    | 9       | UNICAMP      | 70.59%           | 9       | UFSCAR       | 69.17%           |
| 10      | UFSCAR       | 76.69%          | 10      | UF RJ        | 50.00%           | 10      | UFABC        | 45.74%           | 10      | UFF          | 68.84%           | 10      | UNICAMP      | 67.65%           |
| 11      | UFMG         | 75.00%          | 11      | UFABC        | 48.94%           | 11      | UF RJ        | 45.08%           | 11      | UFSC         | 68.35%           | 11      | UFES         | 67.50%           |
| 12      | UFJF         | 75.00%          | 12      | UNESP        | 48.86%           | 12      | UFJF         | 44.64%           | 12      | UFJF         | 67.86%           | 12      | UF RJ        | 66.39%           |
| 13      | UFSC         | 74.68%          | 13      | UFV          | 48.39%           | 13      | UNESP        | 42.05%           | 13      | UFSCAR       | 66.92%           | 13      | UEM          | 65.85%           |
| 14      | UEM          | 73.17%          | 14      | <b>UNB</b>   | <b>47.06%</b>    | 14      | UFF          | 39.13%           | 14      | UFMS         | 66.67%           | 14      | UFMS         | 64.29%           |
| 15      | UFV          | 72.58%          | 15      | UFPR         | 44.44%           | 15      | UFMS         | 38.10%           | 15      | <b>UNB</b>   | <b>65.88%</b>    | 15      | UFRGS        | 63.33%           |
| 16      | <b>UNB</b>   | <b>71.76%</b>   | 16      | UFAL         | 42.86%           | 16      | UNIFESP      | 37.50%           | 16      | UFES         | 65.00%           | 16      | <b>UNB</b>   | <b>61.18%</b>    |
| 17      | UFPR         | 69.44%          | 17      | UFF          | 40.58%           | 17      | UFAL         | 36.73%           | 17      | UNOESTE      | 61.11%           | 17      | UFPR         | 58.33%           |
| 18      | UFES         | 67.50%          | 18      | UFJF         | 39.29%           | 18      | UFES         | 33.75%           | 18      | UF RJ        | 59.02%           | 18      | UFAL         | 57.14%           |
| 19      | UFRGS        | 66.67%          | 19      | UEM          | 39.02%           | 19      | UNESP        | 32.91%           | 19      | UNESP        | 55.11%           | 19      | UNESP        | 56.82%           |
| 20      | UFAL         | 63.27%          | 20      | UFES         | 32.50%           | 20      | UNOESTE      | 22.22%           | 20      | UFAL         | 42.86%           | 20      | UFABC        | 52.13%           |

Tabela 4.19: Percentual de acertos nas questões (CE) 15, 16, 18, 19, 20.

| Posição | Sigla da IES | forma_espec_q_15 | Posição | Sigla da IES | forma_espec_q_16 | Posição | Sigla da IES | forma_espec_q_18 | Posição | Sigla da IES | forma_espec_q_19 | Posição | Sigla da IES | forma_espec_q_20 |
|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|
| 1       | UNOESTE      | 94.44%           | 1       | UFRGS        | 63.33%           | 1       | UFRJ         | 24.59%           | 1       | UNIFESP      | 87.50%           | 1       | UFPR         | 63.89%           |
| 2       | UFRN         | 60.87%           | 2       | UFMG         | 58.33%           | 2       | UFSC         | 22.78%           | 2       | UFMG         | 85.42%           | 2       | UNICAMP      | 61.76%           |
| 3       | UFMG         | 60.42%           | 3       | UFPR         | 56.94%           | 3       | UFV          | 22.58%           | 3       | UFRN         | 78.26%           | 3       | UFMS         | 57.14%           |
| 4       | UFMS         | 59.52%           | 4       | UFSC         | 54.43%           | 4       | UFAL         | 22.45%           | 4       | UFABC        | 77.66%           | 4       | UFJF         | 57.14%           |
| 5       | UNICAMP      | 58.82%           | 5       | UNB          | 48.24%           | 5       | UNB          | 22.35%           | 5       | UFV          | 77.42%           | 5       | UFRGS        | 56.67%           |
| 6       | UNIFESP      | 57.50%           | 6       | UFMS         | 47.62%           | 6       | UNOESTE      | 22.22%           | 6       | UNICAMP      | 76.47%           | 6       | UFV          | 56.45%           |
| 7       | UFV          | 54.84%           | 7       | UNICAMP      | 47.06%           | 7       | UFRGS        | 21.67%           | 7       | UFSCAR       | 75.19%           | 7       | UFMG         | 56.25%           |
| 8       | UFSCAR       | 54.14%           | 8       | UFABC        | 45.74%           | 8       | UNIFESP      | 20.00%           | 8       | UFPR         | 75.00%           | 8       | UEM          | 56.10%           |
| 9       | UFES         | 53.75%           | 9       | UNESP        | 43.75%           | 9       | UFMG         | 18.75%           | 9       | UFRJ         | 74.59%           | 9       | UFSC         | 54.43%           |
| 10      | UFJF         | 53.57%           | 10      | UFF          | 39.86%           | 10      | UFPR         | 18.06%           | 10      | UFRGS        | 74.17%           | 10      | UNIFESP      | 52.50%           |
| 11      | UFABC        | 51.06%           | 11      | UFRJ         | 38.52%           | 11      | UFRN         | 17.39%           | 11      | UFF          | 73.91%           | 11      | UFRN         | 52.17%           |
| 12      | UFPR         | 50.00%           | 12      | UEM          | 36.59%           | 12      | UFES         | 16.25%           | 12      | UFJF         | 73.21%           | 12      | UNB          | 50.59%           |
| 13      | UNESP        | 47.16%           | 13      | UFSCAR       | 34.59%           | 13      | UFF          | 15.22%           | 13      | UFMS         | 71.43%           | 13      | UFF          | 48.55%           |
| 14      | UFAL         | 45.92%           | 14      | UNOESTE      | 33.33%           | 14      | UFABC        | 14.89%           | 14      | UFSC         | 68.35%           | 14      | UFRJ         | 45.08%           |
| 15      | UFRJ         | 45.90%           | 15      | UFJF         | 30.36%           | 15      | UNICAMP      | 14.71%           | 15      | UNB          | 64.71%           | 15      | UNOESTE      | 44.44%           |
| 16      | UFSC         | 44.30%           | 16      | UNIFESP      | 30.00%           | 16      | UNESP        | 14.20%           | 16      | UFES         | 63.75%           | 16      | UFABC        | 42.55%           |
| 17      | UEM          | 43.90%           | 17      | UFES         | 30.00%           | 17      | UFSCAR       | 12.78%           | 17      | UNESP        | 61.93%           | 17      | UNESP        | 42.03%           |
| 18      | UFRGS        | 43.33%           | 18      | UFV          | 29.03%           | 18      | UFJF         | 12.50%           | 18      | UNOESTE      | 61.13%           | 18      | UFES         | 41.23%           |
| 19      | UFF          | 42.75%           | 19      | UFAL         | 27.55%           | 19      | UFMS         | 11.90%           | 19      | UEM          | 60.98%           | 19      | UFSCAR       | 40.60%           |
| 20      | UNB          | 36.47%           | 20      | UFRN         | 26.09%           | 20      | UEM          | 9.76%            | 20      | UFAL         | 56.12%           | 20      | UFAL         | 36.73%           |

Tabela 4.20: Percentual de acertos nas questões (CE) 22, 23, 26, 27.

| Posição | Sigla da IES | forma_espec_q_22 | Posição | Sigla da IES | forma_espec_q_23 | Posição | Sigla da IES | forma_espec_q_26 | Posição | Sigla da IES | forma_espec_q_27 |
|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|
| 1       | UFPR         | 75.00%           | 1       | UFRGS        | 53.33%           | 1       | UFMG         | 62.50%           | 1       | UFMS         | 85.71%           |
| 2       | UFSC         | 72.15%           | 2       | UFJF         | 51.79%           | 2       | UFV          | 38.71%           | 2       | UFRN         | 82.61%           |
| 3       | UFMG         | 70.83%           | 3       | UEM          | 51.22%           | 3       | UFRGS        | 38.33%           | 3       | UNIFESP      | 82.50%           |
| 4       | UFV          | 67.74%           | 4       | UFSC         | 48.10%           | 4       | UFRN         | 34.78%           | 4       | UFMG         | 81.25%           |
| 5       | UNOESTE      | 66.67%           | 5       | UFMS         | 45.24%           | 5       | UFJF         | 33.93%           | 5       | UFV          | 79.03%           |
| 6       | UFRGS        | 65.00%           | 6       | UNB          | 42.35%           | 6       | UFSCAR       | 33.83%           | 6       | UFABC        | 77.66%           |
| 7       | UNICAMP      | 64.71%           | 7       | UFPR         | 40.28%           | 7       | UNICAMP      | 32.35%           | 7       | UFJF         | 75.00%           |
| 8       | UFMS         | 64.29%           | 8       | UFSCAR       | 39.85%           | 8       | UNB          | 30.59%           | 8       | UFPR         | 73.61%           |
| 9       | UFABC        | 62.77%           | 9       | UNESP        | 39.77%           | 9       | UFPR         | 30.56%           | 9       | UFRGS        | 73.33%           |
| 10      | UNB          | 61.18%           | 10      | UFMG         | 39.58%           | 10      | UFABC        | 29.79%           | 10      | UNOESTE      | 72.22%           |
| 11      | UEM          | 60.98%           | 11      | UNOESTE      | 38.89%           | 11      | UFSC         | 29.11%           | 11      | UFSCAR       | 72.18%           |
| 12      | UFRN         | 60.87%           | 12      | UFRJ         | 37.70%           | 12      | UFF          | 28.26%           | 12      | UNB          | 71.76%           |
| 13      | UFRJ         | 59.02%           | 13      | UNIFESP      | 37.50%           | 13      | UNOESTE      | 27.78%           | 13      | UFRJ         | 68.85%           |
| 14      | UNIFESP      | 57.50%           | 14      | UFF          | 36.96%           | 14      | UFAL         | 26.53%           | 14      | UFAL         | 66.33%           |
| 15      | UNESP        | 54.55%           | 15      | UFES         | 35.00%           | 15      | UEM          | 24.39%           | 15      | UEM          | 65.85%           |
| 16      | UFF          | 52.90%           | 16      | UFABC        | 34.04%           | 16      | UNESP        | 23.86%           | 16      | UFSC         | 65.82%           |
| 17      | UFES         | 52.50%           | 17      | UFAL         | 31.63%           | 17      | UFES         | 23.75%           | 17      | UFES         | 65.00%           |
| 18      | UFSCAR       | 49.62%           | 18      | UFRN         | 30.43%           | 18      | UNIFESP      | 22.50%           | 18      | UFF          | 64.49%           |
| 19      | UFJF         | 46.43%           | 19      | UFV          | 22.58%           | 19      | UFRJ         | 21.31%           | 19      | UNESP        | 59.66%           |
| 20      | UFAL         | 42.86%           | 20      | UNICAMP      | 20.59%           | 20      | UFMS         | 19.05%           | 20      | UNICAMP      | 58.82%           |

As questões com melhor desempenho são as questões 3, 8, 27 e as com pior desempenho as questões 18, 26, o resto apresenta um padrão de resposta médio.

A informação da Tabela 4.21 é um resumo dos dados de desempenho das instituições nas questões. Destaca o melhor e o pior desempenho por questão, indicando quais instituições tiveram as maiores e menores taxas de acerto.

Tabela 4.21: Comparação de desempenho (CE) questões objetivas 1 a 27.

| Questão          | Melhor Desempenho                           | Pior Desempenho                             |
|------------------|---|---|
| forma_espec_q_1  | UNIVERSIDADE FEDERAL DO PARANA              | UNIVERSIDADE FEDERAL DE MINAS GERAIS        |
| forma_espec_q_2  | UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE | UNIVERSIDADE FEDERAL DE ALAGOAS             |
| forma_espec_q_3  | UNIVERSIDADE FEDERAL DE SAO PAULO           | UNIVERSIDADE FEDERAL DE ALAGOAS             |
| forma_espec_q_4  | QUESTÃO ANULADA                             | QUESTÃO ANULADA                             |
| forma_espec_q_5  | QUESTÃO ANULADA                             | QUESTÃO ANULADA                             |
| forma_espec_q_6  | UNIVERSIDADE FEDERAL DO PARANA              | FUNDACAO UNIVERSIDADE FEDERAL DO ABC        |
| forma_espec_q_7  | UNIVERSIDADE DO OESTE PAULISTA              | UNIVERSIDADE FEDERAL DE ALAGOAS             |
| forma_espec_q_8  | UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE | UNIVERSIDADE FEDERAL DE ALAGOAS             |
| forma_espec_q_9  | QUESTÃO ANULADA                             | QUESTÃO ANULADA                             |
| forma_espec_q_10 | UNIVERSIDADE DO OESTE PAULISTA              | UNIVERSIDADE FEDERAL DO ESPIRITO SANTO      |
| forma_espec_q_11 | UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE | UNIVERSIDADE DO OESTE PAULISTA              |
| forma_espec_q_12 | UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE | UNIVERSIDADE FEDERAL DE ALAGOAS             |
| forma_espec_q_13 | QUESTÃO ANULADA                             | QUESTÃO ANULADA                             |
| forma_espec_q_14 | UNIVERSIDADE DO OESTE PAULISTA              | FUNDACAO UNIVERSIDADE FEDERAL DO ABC        |
| forma_espec_q_15 | UNIVERSIDADE DO OESTE PAULISTA              | UNIVERSIDADE DE BRASÍLIA                    |
| forma_espec_q_16 | UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL   | UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE |
| forma_espec_q_17 | QUESTÃO ANULADA                             | QUESTÃO ANULADA                             |
| forma_espec_q_18 | UNIVERSIDADE FEDERAL DO RIO DE JANEIRO      | UNIVERSIDADE ESTADUAL DE MARINGA            |
| forma_espec_q_19 | UNIVERSIDADE FEDERAL DE SAO PAULO           | UNIVERSIDADE FEDERAL DE ALAGOAS             |
| forma_espec_q_20 | UNIVERSIDADE FEDERAL DO PARANA              | UNIVERSIDADE FEDERAL DE ALAGOAS             |
| forma_espec_q_21 | QUESTÃO ANULADA                             | QUESTÃO ANULADA                             |
| forma_espec_q_22 | UNIVERSIDADE FEDERAL DO PARANA              | UNIVERSIDADE FEDERAL DE ALAGOAS             |
| forma_espec_q_23 | UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL   | UNIVERSIDADE ESTADUAL DE CAMPINAS           |
| forma_espec_q_24 | QUESTÃO ANULADA                             | QUESTÃO ANULADA                             |
| forma_espec_q_25 | QUESTÃO ANULADA                             | QUESTÃO ANULADA                             |
| forma_espec_q_26 | UNIVERSIDADE FEDERAL DE MINAS GERAIS        | UNIVERSIDADE FEDERAL DE SANTA MARIA         |
| forma_espec_q_27 | UNIVERSIDADE FEDERAL DE SANTA MARIA         | UNIVERSIDADE ESTADUAL DE CAMPINAS           |

A Universidade Federal do Paraná se destaca com o melhor desempenho em várias questões, o que pode indicar um forte alinhamento do seu currículo com o conteúdo da prova, enquanto a Universidade Federal de Alagoas figura várias vezes na coluna de pior desempenho.

Olhando a Tabela 4.17 a questão<sup>3</sup> que os alunos de Ciência da Computação (Bacharelado) da UnB tiveram um melhor desempenho foi a questão 11 (questão número três do Componente Específico) que registrou uma taxa de acerto de 80%.

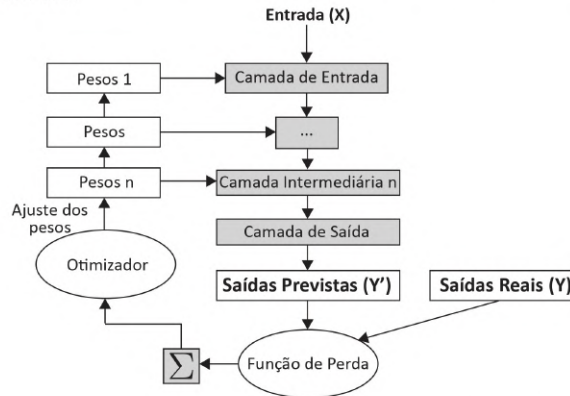
<sup>3</sup>As provas e gabaritos de cada curso podem ser consultados no site do INEP [33]. No caderno de questões, as questões de formação geral estão numeradas de 1 a 8, enquanto as questões de Componente Específico vão da 9 até a 35, sendo a questão 9 a primeira e a 35 a vigésima sétima deste grupo.



### QUESTÃO 11

Uma equipe de cientistas da computação de uma determinada empresa de animação foi designada para desenvolver um sistema capaz de varrer a web no intuito de detectar sites que possam estar usando imagens de seus personagens de animação sem o devido consentimento. Portanto, o sistema deverá receber imagens como entrada, classificá-las entre imagens da empresa e imagens não produzidas pela empresa.

A figura abaixo esboça uma arquitetura de rede neural profunda e o processo de treinamento que os cientistas pretendem usar.



CHOLLET, F. *Deep Learning with Python*. New York: Manning Publications, 2017 (adaptado).

Após uma tentativa, notaram-se duas dificuldades: 1) o tempo de treinamento da rede estava muito longo e 2) a acurácia da rede treinada não estava no patamar aceito pela empresa.

Diante deste contexto, avalie as afirmações a seguir.

- I. Aumentar o número de camadas é uma alternativa que pode levar a uma melhora na acurácia, além de diminuir o tempo de treinamento da rede.
- II. Fazer uso de redes convolucionais é uma alternativa que pode levar a uma melhora na acurácia, no entanto, pode exigir uso de máquinas com maior poder de processamento.
- III. Aumentar o número de unidades de processamento (neurônios) nas camadas pode levar a uma piora na acurácia, além de diminuir o tempo de treinamento da rede.
- IV. Aumentar o número de amostras de treinamento é uma alternativa que pode levar a uma melhora na acurácia, apesar de aumentar o tempo de treinamento da rede.
- V. Fazer uso de redes recorrentes é uma alternativa que pode levar a uma melhora na acurácia, no entanto, pode exigir uso de máquinas com maior poder de processamento.

É correto apenas o que se afirma em

- A I e IV.
- B I e V.
- C II e III.
- D II e IV.
- E III e V.

Figura 4.11: Questão 11 para Ciência da Computação (Bacharelado).

Por outro lado, a UnB teve seu desempenho mais baixo na questão 26, (Questão número dezoito do Componente Específico) com uma taxa de acerto de 22.35%.

#### QUESTÃO 26

As interfaces adaptativas realizam as adaptações e personalizações de forma automática e dinâmica durante o processo de navegação, com base na aprendizagem da navegação e da interação do usuário. Técnicas de interfaces adaptativas podem ser utilizadas para adaptar interfaces às preferências do usuário, à sua capacidade cognitiva e ao seu estilo de navegação, tornando as interações mais naturais e atrativas. Essas técnicas de adaptação podem ser empregadas tanto para a reorganização dos objetos no ambiente, como para alterar a forma de apresentar informações. As adaptações podem ocorrer em diferentes níveis ou de diferentes formas: adaptação de conteúdo, adaptação da navegação e adaptação da apresentação do conteúdo. Cada um desses níveis de adaptação possui métodos e técnicas de adaptação próprios.

NIENOW, A. L. Interfaces adaptativas no comércio eletrônico como facilitadoras da inclusão digital de idosos. *Revista Tecnologia e Tendências*, v. 9, n. 2, p. 116-136, 2017 (adaptado).

Considerando a construção de interfaces adaptativas na interação homem-computador, avalie as afirmações a seguir quanto aos níveis e técnicas de adaptação.

- I. No nível de conteúdo, podem-se considerar as técnicas de fragmentos de texto, fragmentos condicionais, páginas variantes e abordagem baseada em *frames*.
- II. No nível de navegação, podem-se considerar as técnicas de *layouts* de página e guias de estilos.
- III. No nível de apresentação, podem-se considerar as técnicas de orientação direta, anotação de *links*, apresentação e ocultação e ordenação de *links*.

É correto o que se afirma em

- A I, apenas.
- B III, apenas.
- C I e II, apenas.
- D II e III, apenas.
- E I, II e III.

Figura 4.12: Questão 26 para Ciência da Computação (Bacharelado).

Para uma visão mais aprofundada, as Tabela 4.22 representa as respostas dadas pelos alunos da UnB. A tabela está dividida em colunas das questões de 1 a 27 e linhas que representam as opções de resposta marcadas (A, B, C, D, E, opção em branco e várias onde o aluno marcou várias opções). Z representa que a questão excluída devido à anulação e X questão excluída devido ao coeficiente ponto-bisserial  $< 0,20$ . A última coluna indica o gabarito correto para cada questão. Cada linha tem destacada a opção mais marcada com a cor azul, a última coluna indica o gabarito correto para cada questão.

Tabela 4.22: Respostas da UnB para as questões objetivas (CE) Bacharelado.

| Questões                     | Opção A | Opção B | Opção C | Opção D | Opção E | Opção em branco | Várias | Gabarito |
|------------------------------|---------|---------|---------|---------|---------|-----------------|--------|----------|
| forma_espec_alternativa_q_1  | 2       | 5       | 11      | 39      | 28      | 0               | 0      | E        |
| forma_espec_alternativa_q_2  | 22      | 57      | 4       | 1       | 1       | 0               | 0      | B        |
| forma_espec_alternativa_q_3  | 5       | 3       | 6       | 68      | 2       | 1               | 0      | D        |
| forma_espec_alternativa_q_4  | 0       | 17      | 2       | 46      | 19      | 1               | 0      | X        |
| forma_espec_alternativa_q_5  | 11      | 5       | 7       | 41      | 20      | 1               | 0      | X        |
| forma_espec_alternativa_q_6  | 31      | 11      | 36      | 4       | 2       | 1               | 0      | C        |
| forma_espec_alternativa_q_7  | 4       | 1       | 55      | 6       | 19      | 0               | 0      | C        |
| forma_espec_alternativa_q_8  | 4       | 61      | 3       | 0       | 16      | 0               | 1      | B        |
| forma_espec_alternativa_q_9  | 27      | 11      | 16      | 21      | 10      | 0               | 0      | X        |
| forma_espec_alternativa_q_10 | 9       | 15      | 9       | 10      | 40      | 2               | 0      | E        |
| forma_espec_alternativa_q_11 | 17      | 10      | 8       | 40      | 9       | 1               | 0      | D        |
| forma_espec_alternativa_q_12 | 56      | 10      | 10      | 4       | 4       | 1               | 0      | A        |
| forma_espec_alternativa_q_13 | 3       | 13      | 24      | 30      | 8       | 6               | 1      | X        |
| forma_espec_alternativa_q_14 | 52      | 6       | 23      | 0       | 2       | 2               | 0      | A        |
| forma_espec_alternativa_q_15 | 2       | 23      | 31      | 12      | 16      | 1               | 0      | C        |
| forma_espec_alternativa_q_16 | 13      | 3       | 17      | 10      | 41      | 1               | 0      | E        |
| forma_espec_alternativa_q_17 | 15      | 22      | 11      | 28      | 8       | 1               | 0      | X        |
| forma_espec_alternativa_q_18 | 19      | 7       | 16      | 8       | 33      | 2               | 0      | A        |
| forma_espec_alternativa_q_19 | 3       | 11      | 7       | 55      | 5       | 4               | 0      | D        |
| forma_espec_alternativa_q_20 | 12      | 43      | 10      | 11      | 6       | 3               | 0      | B        |
| forma_espec_alternativa_q_21 | 6       | 9       | 55      | 5       | 6       | 4               | 0      | Z        |
| forma_espec_alternativa_q_22 | 22      | 3       | 4       | 4       | 52      | 0               | 0      | E        |
| forma_espec_alternativa_q_23 | 4       | 15      | 17      | 8       | 36      | 5               | 0      | E        |
| forma_espec_alternativa_q_24 | 12      | 12      | 10      | 41      | 8       | 2               | 0      | X        |
| forma_espec_alternativa_q_25 | 10      | 5       | 5       | 55      | 7       | 3               | 0      | Z        |
| forma_espec_alternativa_q_26 | 11      | 16      | 26      | 21      | 10      | 1               | 0      | C        |
| forma_espec_alternativa_q_27 | 8       | 7       | 3       | 61      | 4       | 2               | 0      | D        |

### Ciência da Computação (Licenciatura)

As Tabelas 4.23 a 4.26 apresentam a taxa de acertos por questão da parte objetiva de conhecimentos específicos para o curso de Ciência da Computação (Licenciatura) para as 5 faculdades que ofertavam o curso. As seguintes questões foram anuladas:

- **Questões (4, 9, 13, 17, 18)** - Anuladas pelo índice de discriminação (correlação pontobisserial  $< 0,20$ ).



Tabela 4.23: Percentual de acertos nas questões 1 a 6.

| Posição | Sigla da IES | forma_espec_q_1 | Posição | Sigla da IES | forma_espec_q_2 | Posição | Sigla da IES | forma_espec_q_3 | Posição | Sigla da IES | forma_espec_q_5 | Posição | Sigla da IES | forma_espec_q_6 |
|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|-----------------|
| 1       | UFSM         | 66.67%          | 1       | UNB          | 53.85%          | 1       | UNB          | 74.36%          | 1       | UFF          | 50.00%          | 1       | UFF          | 50.00%          |
| 2       | UFPR         | 41.67%          | 2       | UFPR         | 41.67%          | 2       | UFPR         | 66.67%          | 2       | UNB          | 41.03%          | 2       | UNB          | 33.33%          |
| 3       | UFJF         | 34.62%          | 3       | UFJF         | 34.62%          | 3       | UFJF         | 42.31%          | 3       | UFPR         | 33.33%          | 3       | UFJF         | 30.77%          |
| 4       | UNB          | 30.77%          | 4       | UFSM         | 33.33%          | 4       | UFSM         | 33.33%          | 4       | UFSM         | 33.33%          | 4       | UFPR         | 8.33%           |
| 5       | UFF          | 20.00%          | 5       | UFF          | 10.00%          | 5       | UFF          | 20.00%          | 5       | UFJF         | 23.08%          | 5       | UFSM         | 0.00%           |

Tabela 4.24: Percentual de acertos nas questões 7 a 12.

| Posição | Sigla da IES | forma_espec_q_7 | Posição | Sigla da IES | forma_espec_q_8 | Posição | Sigla da IES | forma_espec_q_10 | Posição | Sigla da IES | forma_espec_q_11 | Posição | Sigla da IES | forma_espec_q_12 |
|---------|--------------|-----------------|---------|--------------|-----------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|
| 1       | UFSM         | 66.67%          | 1       | UFPR         | 75.00%          | 1       | UFPR         | 58.33%           | 1       | UFF          | 50.00%           | 1       | UFSM         | 33.33%           |
| 2       | UNB          | 64.10%          | 2       | UNB          | 71.79%          | 2       | UNB          | 48.72%           | 2       | UNB          | 41.03%           | 2       | UNB          | 30.77%           |
| 3       | UFJF         | 53.85%          | 3       | UFSM         | 66.67%          | 3       | UFSM         | 33.33%           | 3       | UFSM         | 33.33%           | 3       | UFF          | 10.00%           |
| 4       | UFPR         | 50.00%          | 4       | UFF          | 60.00%          | 4       | UFF          | 30.00%           | 4       | UFJF         | 30.77%           | 4       | UFPR         | 8.33%            |
| 5       | UFF          | 40.00%          | 5       | UFJF         | 30.77%          | 5       | UFJF         | 26.92%           | 5       | UFPR         | 25.00%           | 5       | UFJF         | 7.69%            |

Tabela 4.25: Percentual de acertos nas questões 14 a 21.

| Posição | Sigla da IES | forma_espec_q_14 | Posição | Sigla da IES | forma_espec_q_15 | Posição | Sigla da IES | forma_espec_q_16 | Posição | Sigla da IES | forma_espec_q_19 | Posição | Sigla da IES | forma_espec_q_20 | Posição | Sigla da IES | forma_espec_q_21 |
|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|
| 1       | UFF          | 70.00%           | 1       | UFPR         | 91.67%           | 1       | UNB          | 41.03%           | 1       | UFSM         | 66.67%           | 1       | UFSM         | 100.00%          | 1       | UFJF         | 65.38%           |
| 2       | UFSM         | 66.67%           | 2       | UFJF         | 53.85%           | 2       | UFF          | 20.00%           | 2       | UNB          | 46.15%           | 2       | UFPR         | 83.33%           | 2       | UFF          | 60.00%           |
| 3       | UNB          | 33.33%           | 3       | UFSM         | 33.33%           | 3       | UFPR         | 8.33%            | 3       | UFJF         | 46.15%           | 3       | UFJF         | 73.08%           | 3       | UNB          | 46.15%           |
| 4       | UFPR         | 25.00%           | 4       | UNB          | 30.77%           | 4       | UFJF         | 7.69%            | 4       | UFPR         | 41.67%           | 4       | UNB          | 51.28%           | 4       | UFPR         | 41.67%           |
| 5       | UFJF         | 19.23%           | 5       | UFF          | 20.00%           | 5       | UFSM         | 0.00%            | 5       | UFF          | 10.00%           | 5       | UFF          | 30.00%           | 5       | UFSM         | 33.33%           |

Tabela 4.26: Percentual de acertos nas questões 22 a 27.

| Posição | Sigla da IES | forma_espec_q_22 | Posição | Sigla da IES | forma_espec_q_23 | Posição | Sigla da IES | forma_espec_q_24 | Posição | Sigla da IES | forma_espec_q_25 | Posição | Sigla da IES | forma_espec_q_26 | Posição | Sigla da IES | forma_espec_q_27 |
|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|---------|--------------|------------------|
| 1       | UFF          | 60.00%           | 1       | UFSM         | 66.67%           | 1       | UFSM         | 100.00%          | 1       | UFSM         | 66.67%           | 1       | UFPR         | 75.00%           | 1       | UNB          | 66.67%           |
| 2       | UNB          | 51.28%           | 2       | UNB          | 48.72%           | 2       | UFPR         | 66.67%           | 2       | UFJF         | 50.00%           | 2       | UFSM         | 66.67%           | 2       | UFSM         | 66.67%           |
| 3       | UFPR         | 50.00%           | 3       | UFF          | 30.00%           | 3       | UNB          | 53.85%           | 3       | UNB          | 46.15%           | 3       | UNB          | 64.10%           | 3       | UFJF         | 57.69%           |
| 4       | UFSM         | 33.33%           | 4       | UFPR         | 25.00%           | 4       | UFF          | 50.00%           | 4       | UFF          | 30.00%           | 4       | UFJF         | 50.00%           | 4       | UFF          | 50.00%           |
| 5       | UFJF         | 30.77%           | 5       | UFJF         | 23.08%           | 5       | UFJF         | 42.31%           | 5       | UFPR         | 25.00%           | 5       | UFF          | 20.00%           | 5       | UFPR         | 25.00%           |

A Tabela 4.27 é um resumo dos dados de desempenho das instituições nas questões. Destaca o melhor e o pior desempenho por questão, indicando quais instituições tiveram as maiores e menores taxas de acerto.

Tabela 4.27: Comparação de desempenho (CE) questões objetivas 1 a 27 - Licenciatura.

| Questão          | Melhor Desempenho                    | Pior Desempenho                      |
|------------------|--------------------------------------|--------------------------------------|
| forma_espec_q_1  | UNIVERSIDADE FEDERAL DE SANTA MARIA  | UNIVERSIDADE FEDERAL FLUMINENSE      |
| forma_espec_q_2  | UNIVERSIDADE DE BRASÍLIA             | UNIVERSIDADE FEDERAL FLUMINENSE      |
| forma_espec_q_3  | UNIVERSIDADE DE BRASÍLIA             | UNIVERSIDADE FEDERAL FLUMINENSE      |
| forma_espec_q_4  | QUESTÃO ANULADA                      | QUESTÃO ANULADA                      |
| forma_espec_q_5  | UNIVERSIDADE FEDERAL FLUMINENSE      | UNIVERSIDADE FEDERAL DE JUIZ DE FORA |
| forma_espec_q_6  | UNIVERSIDADE FEDERAL FLUMINENSE      | UNIVERSIDADE FEDERAL DE SANTA MARIA  |
| forma_espec_q_7  | UNIVERSIDADE FEDERAL DE SANTA MARIA  | UNIVERSIDADE FEDERAL FLUMINENSE      |
| forma_espec_q_8  | UNIVERSIDADE FEDERAL DO PARANA       | UNIVERSIDADE FEDERAL DE JUIZ DE FORA |
| forma_espec_q_9  | QUESTÃO ANULADA                      | QUESTÃO ANULADA                      |
| forma_espec_q_10 | UNIVERSIDADE FEDERAL DO PARANA       | UNIVERSIDADE FEDERAL DE JUIZ DE FORA |
| forma_espec_q_11 | UNIVERSIDADE FEDERAL FLUMINENSE      | UNIVERSIDADE FEDERAL DO PARANA       |
| forma_espec_q_12 | UNIVERSIDADE FEDERAL DE SANTA MARIA  | UNIVERSIDADE FEDERAL DE JUIZ DE FORA |
| forma_espec_q_13 | QUESTÃO ANULADA                      | QUESTÃO ANULADA                      |
| forma_espec_q_14 | UNIVERSIDADE FEDERAL FLUMINENSE      | UNIVERSIDADE FEDERAL DE JUIZ DE FORA |
| forma_espec_q_15 | UNIVERSIDADE FEDERAL DO PARANA       | UNIVERSIDADE FEDERAL FLUMINENSE      |
| forma_espec_q_16 | UNIVERSIDADE DE BRASÍLIA             | UNIVERSIDADE FEDERAL DE SANTA MARIA  |
| forma_espec_q_17 | QUESTÃO ANULADA                      | QUESTÃO ANULADA                      |
| forma_espec_q_18 | QUESTÃO ANULADA                      | QUESTÃO ANULADA                      |
| forma_espec_q_19 | UNIVERSIDADE FEDERAL DE SANTA MARIA  | UNIVERSIDADE FEDERAL FLUMINENSE      |
| forma_espec_q_20 | UNIVERSIDADE FEDERAL DE SANTA MARIA  | UNIVERSIDADE FEDERAL FLUMINENSE      |
| forma_espec_q_21 | UNIVERSIDADE FEDERAL DE JUIZ DE FORA | UNIVERSIDADE FEDERAL DE SANTA MARIA  |
| forma_espec_q_22 | UNIVERSIDADE FEDERAL FLUMINENSE      | UNIVERSIDADE FEDERAL DE JUIZ DE FORA |
| forma_espec_q_23 | UNIVERSIDADE FEDERAL DE SANTA MARIA  | UNIVERSIDADE FEDERAL DE JUIZ DE FORA |
| forma_espec_q_24 | UNIVERSIDADE FEDERAL DE SANTA MARIA  | UNIVERSIDADE FEDERAL DE JUIZ DE FORA |
| forma_espec_q_25 | UNIVERSIDADE FEDERAL DE SANTA MARIA  | UNIVERSIDADE FEDERAL DO PARANA       |
| forma_espec_q_26 | UNIVERSIDADE FEDERAL DO PARANA       | UNIVERSIDADE FEDERAL FLUMINENSE      |
| forma_espec_q_27 | UNIVERSIDADE DE BRASÍLIA             | UNIVERSIDADE FEDERAL DO PARANA       |

A UnB mostra um desempenho forte em questões 2, 3, 16, 27 indicando áreas em que seus alunos podem ter tido uma preparação particularmente eficaz, enquanto a Universidade Federal Fluminense aparece em um maior número de vezes na coluna de pior desempenho.

Olhando os dados podemos notar que os alunos de Ciência da Computação (Licenciatura) da UnB tiveram um melhor desempenho foi a questão 11 (Questão número três do Componente Específico) que registrou uma taxa de acerto de 74% que a Figura 4.11 é a mesma questão que a de bacharelado.

Por outro lado, teve seu desempenho mais baixo em três questões 9, 15 e 18 (questão número 1, 12 e 15 do componente específico) com uma taxa de acerto para as três de 30,77%. Na Figura 4.13 é mostrada apenas a questão 9.

### QUESTÃO 09

Quando um computador é multiprogramado, ele geralmente tem múltiplos processos ou *threads* que competem pela CPU ao mesmo tempo. Essa situação ocorre sempre que dois ou mais processos estão simultaneamente no estado pronto. Se somente uma CPU se encontrar disponível, deverá ser feita uma escolha de qual processo executar em seguida. A parte do sistema operacional que faz a escolha é chamada de **escalador**, e o algoritmo que ele usa é o **algoritmo de escalonamento**.

TANENBAUM, A. S. **Sistemas Operacionais Modernos**. 3. ed., São Paulo: Pearson, 2010 (adaptado).

Considerando que em ambientes diferentes são necessários algoritmos diferentes de escalonamento, garantindo assim que seja maximizado o uso de seus recursos, assinale a opção que apresenta um algoritmo de escalonamento seguido do tipo de ambiente no qual deva ser implementado.

- A** Primeiro a chegar, último a sair (*first in, last out* - FILO); propício para sistemas de tempo real.
- B** Escalonamento por taxas monotônicas (*rate monotonic scheduling* - RMS); propício para sistemas em lote.
- C** Tarefa mais curta primeiro; propício para sistemas interativos.
- D** Escalonamento por chave circular (*round-robin*); propício para sistemas de tempo real.
- E** Escalonamento por prioridades; propício para sistemas interativos.

Área livre

Figura 4.13: Questão 9 para Ciência da Computação (Licenciatura).

As Tabela 4.28 representam as respostas dadas pelos alunos da UnB do curso Ciência da Computação (Licenciatura). *Z* representa que a questão excluída devido à anulação e *X* questão excluída devido ao coeficiente ponto-bisserial  $< 0,20$ . Cada linha tem destacada a opção mais marcada com a cor azul, a última coluna indica o gabarito correto para cada questão.

Tabela 4.28: Respostas da UnB para as questões objetivas (CE) Licenciatura.

| Questões                     | Opção A | Opção B | Opção C | Opção D | Opção E | Opção em branco | Várias | Gabarito |
|------------------------------|---------|---------|---------|---------|---------|-----------------|--------|----------|
| forma_espec_alternativa_q_1  | 4       | 2       | 5       | 16      | 12      | 0               | 0      | E        |
| forma_espec_alternativa_q_2  | 14      | 21      | 3       | 1       | 0       | 0               | 0      | B        |
| forma_espec_alternativa_q_3  | 1       | 4       | 3       | 29      | 2       | 0               | 0      | D        |
| forma_espec_alternativa_q_4  | 2       | 5       | 2       | 22      | 8       | 0               | 0      | X        |
| forma_espec_alternativa_q_5  | 7       | 1       | 2       | 13      | 16      | 0               | 0      | E        |
| forma_espec_alternativa_q_6  | 15      | 11      | 13      | 0       | 0       | 0               | 0      | C        |
| forma_espec_alternativa_q_7  | 0       | 0       | 25      | 4       | 10      | 0               | 0      | C        |
| forma_espec_alternativa_q_8  | 3       | 28      | 1       | 2       | 5       | 0               | 0      | B        |
| forma_espec_alternativa_q_9  | 11      | 2       | 7       | 10      | 7       | 2               | 0      | X        |
| forma_espec_alternativa_q_10 | 3       | 3       | 5       | 8       | 19      | 1               | 0      | E        |
| forma_espec_alternativa_q_11 | 12      | 1       | 6       | 16      | 4       | 0               | 0      | D        |
| forma_espec_alternativa_q_12 | 12      | 5       | 12      | 5       | 5       | 0               | 0      | A        |
| forma_espec_alternativa_q_13 | 1       | 6       | 9       | 18      | 4       | 1               | 0      | X        |
| forma_espec_alternativa_q_14 | 13      | 11      | 13      | 0       | 2       | 0               | 0      | A        |
| forma_espec_alternativa_q_15 | 4       | 7       | 12      | 10      | 4       | 2               | 0      | C        |
| forma_espec_alternativa_q_16 | 2       | 4       | 10      | 7       | 16      | 0               | 0      | E        |
| forma_espec_alternativa_q_17 | 7       | 4       | 5       | 12      | 11      | 0               | 0      | X        |
| forma_espec_alternativa_q_18 | 6       | 2       | 4       | 8       | 19      | 0               | 0      | X        |
| forma_espec_alternativa_q_19 | 8       | 10      | 3       | 18      | 0       | 0               | 0      | D        |
| forma_espec_alternativa_q_20 | 2       | 4       | 6       | 20      | 7       | 0               | 0      | D        |
| forma_espec_alternativa_q_21 | 18      | 6       | 3       | 5       | 7       | 0               | 0      | A        |
| forma_espec_alternativa_q_22 | 20      | 5       | 7       | 3       | 4       | 0               | 0      | A        |
| forma_espec_alternativa_q_23 | 1       | 2       | 5       | 19      | 12      | 0               | 0      | D        |
| forma_espec_alternativa_q_24 | 21      | 10      | 4       | 3       | 1       | 0               | 0      | A        |
| forma_espec_alternativa_q_25 | 9       | 0       | 18      | 12      | 0       | 0               | 0      | C        |
| forma_espec_alternativa_q_26 | 0       | 5       | 25      | 9       | 0       | 0               | 0      | C        |
| forma_espec_alternativa_q_27 | 1       | 1       | 5       | 6       | 26      | 0               | 0      | E        |

#### 4.3.4 Comparação do desempenho dos alunos da UnB com outras universidades

Esta seção apresenta uma análise do desempenho dos alunos da UnB dos cursos Ciência da Computação Bacharelado e Computação Licenciatura em comparação com estudantes de outras universidades que foram selecionadas na Tabela 4.6.

## Ciência da Computação (Bacharelado)

A Tabela 4.29 mostra uma tabela representando o desempenho dos alunos da UnB em relação a outras universidades, em oito questões objetivas da formação geral. As questões são identificadas como “forma\_geral\_q\_x”, onde  $X$  representa o número da questão. Ao lado de cada questão, há uma porcentagem que indica o desempenho relativo da UnB, onde 100% significaria que a UnB foi a universidade com o melhor desempenho naquela questão e 0% indicaria o pior desempenho.

Tabela 4.29: Comparação do desempenho (FG) da UnB - Ciência da Computação - Bacharelado com outras instituições.

| Número da Questão | Desempenho da UnB em relação a outras universidades |
|-------------------|---|
| forma_geral_q_1   | 30.00 %   |
| forma_geral_q_2   | 85.00 %   |
| forma_geral_q_3   | 50.00 %   |
| forma_geral_q_4   | 80.00 %   |
| forma_geral_q_5   | 95.00 %   |
| forma_geral_q_6   | 25.00 %   |
| forma_geral_q_7   | 25.00 %   |
| forma_geral_q_8   | 85.00 %   |
| Média             | 59.38 %   |

A tabela mostra que a UnB teve desempenho excepcionalmente alto em algumas questões específicas, com taxas de 85% e 95% nas questões 2 e 5, respectivamente. No entanto, nas questões 1, 3, 6 e 7, o desempenho é significativamente menor, com apenas 25% a 50%. Isso indica que nessas áreas a UnB pode estar atrás de outras instituições, o que pode ser um sinal para revisar a abordagem de ensino.

A Tabela 4.30 exibe o desempenho percentual da UnB para as questões do Componente Específico identificadas como “forma\_espec\_q\_x” onde  $X$  representa o número da questão. As seguintes questões foram anuladas e não foram colocadas na tabela:

- **Questões (21, 25)** - Anulada pela comissão.
- **Questões (4, 5, 9, 13, 17, 24)** - Anulada pelo índice de discriminação (correlação pontobisserial  $< 0,20$ ).

Tabela 4.30: Comparação do desempenho (CE) da UnB - Ciência da Computação - Bacharelado com outras instituições.

| <b>Número da Questão</b> | <b>Desempenho da UnB em relação a outras universidades</b> |
|--------------------------|--|
| forma_espec_q_1          | 15.00 %  |
| forma_espec_q_2          | 85.00 %  |
| forma_espec_q_3          | 30.00 %  |
| forma_espec_q_6          | 95.00 %  |
| forma_espec_q_7          | 75.00 %  |
| forma_espec_q_8          | 25.00 %  |
| forma_espec_q_10         | 35.00 %  |
| forma_espec_q_11         | 60.00 %  |
| forma_espec_q_12         | 30.00 %  |
| forma_espec_q_14         | 25.00 %  |
| forma_espec_q_15         | 5.00 %   |
| forma_espec_q_16         | 80.00 %  |
| forma_espec_q_18         | 80.00 %  |
| forma_espec_q_19         | 30.00 %  |
| forma_espec_q_20         | 45.00 %  |
| forma_espec_q_22         | 55.00 %  |
| forma_espec_q_23         | 75.00 %  |
| forma_espec_q_26         | 65.00 %  |
| forma_espec_q_27         | 45.00 %  |
| Média                    | 50.26 %  |

Questões 6 e 16 apresentam desempenho notavelmente alto (95% e 80%, respectivamente), indicando que os alunos da UnB têm um excelente entendimento desses tópicos específicos em comparação que outras instituições. Por outro lado as questões 1, 14 e 15

têm um desempenho relativamente baixo (15%, 25% e 5%), o que sugere que os alunos da UnB podem estar encontrando dificuldades nesses temas ou que as outras universidades estão particularmente fortes nestas áreas.

### Computação (Licenciatura)

A Tabela 4.31 representa o desempenho dos alunos da UnB em relação a outras universidades, em oito questões objetivas da formação geral. As questões são identificadas como “forma\_geral\_q\_x”, onde  $X$  representa o número da questão. Ao lado de cada questão, há uma porcentagem que indica o desempenho relativo da UnB, onde 100% significaria que a UnB foi a universidade com o melhor desempenho naquela questão e 0% indicaria o pior desempenho.

Tabela 4.31: Comparação do desempenho (FG) da UnB - Computação - Licenciatura com outras instituições.

| Número da Questão | Desempenho da UnB em relação a outras universidades |
|-------------------|---|
| forma_geral_q_1   | 40.00 %   |
| forma_geral_q_2   | 100.00 %  |
| forma_geral_q_3   | 60.00 %   |
| forma_geral_q_4   | 80.00 %   |
| forma_geral_q_5   | 80.00 %   |
| forma_geral_q_6   | 40.00 %   |
| forma_geral_q_7   | 80.00 %   |
| forma_geral_q_8   | 80.00 %   |
| Média             | 70.00 %   |

A UnB destaca-se na questão 2, com um desempenho perfeito de 100% e de forma geral os dados refletem positivamente, com maior parte das questões acima de 60%, indicando que os alunos geralmente se saem bem quando comparados a outras instituições.

A Tabela 4.32 exibe o desempenho percentual da UnB para as questões do componente específico identificadas como “forma\_espec\_q\_x” onde  $X$  representa o número da questão. As seguintes questões foram anuladas e não foram colocadas na tabela:

- **Questões (4, 9, 13, 17, 18)**- Anuladas pelo índice de discriminação (correlação pontobisserial  $< 0,20$ ) (Carácter 9).



Tabela 4.32: Comparação do desempenho (CE) da UnB - Computação - Licenciatura com outras instituições.

| <b>Número da Questão</b> | <b>Desempenho da UnB em relação a outras universidades</b> |
|--------------------------|--|
| forma_espec_q_1          | 40.00 %  |
| forma_espec_q_2          | 100.00 %   |
| forma_espec_q_3          | 100.00 %   |
| forma_espec_q_5          | 80.00 %  |
| forma_espec_q_6          | 80.00 %  |
| forma_espec_q_7          | 80.00 %  |
| forma_espec_q_8          | 80.00 %  |
| forma_espec_q_10         | 80.00 %  |
| forma_espec_q_11         | 80.00 %  |
| forma_espec_q_12         | 80.00 %  |
| forma_espec_q_14         | 60.00 %  |
| forma_espec_q_15         | 40.00 %  |
| forma_espec_q_16         | 100.00 %   |
| forma_espec_q_19         | 80.00 %  |
| forma_espec_q_20         | 40.00 %  |
| forma_espec_q_21         | 60.00 %  |
| forma_espec_q_22         | 80.00 %  |
| forma_espec_q_23         | 80.00 %  |
| forma_espec_q_24         | 60.00 %  |
| forma_espec_q_25         | 60.00 %  |
| forma_espec_q_26         | 60.00 %  |
| forma_espec_q_27         | 100.00 %   |
| Média                    | 73.64 %  |

Questões com desempenho de 80% ou mais são numerosas, sugerindo que, em geral, os alunos da UnB estão bem alinhados ou superam outras instituições nessas áreas. Isso



sugere que os alunos da UnB têm um entendimento superior nessas áreas específicas.

A Tabela 4.33 representa as questões dos conhecimentos específicos e o conteúdo avaliado, essa tabela pode ajudar para saber quais áreas precisam de um maior apoio ao estudante.

Tabela 4.33: Disciplinas da UnB.

| Disciplina                                | Ciência da Computação (Bacharelado) | Computação (Licenciatura)           |
|---|-------------------------------------|-------------------------------------|
|   | Questões                            | Questões                            |
| SISTEMAS OPERACIONAIS                     | 9, 14, 17 e 35                      | 9, 14 e 17                          |
| ESTRUTURAS DE DADOS                       | 10, 23 e 34                         | 10 e 23                             |
| INTRODUCAO A INTELIGENCIA ARTIFICIAL      | 11 e 18                             | 11 e 18                             |
| BANCOS DE DADOS                           | 12 e 22                             | 12 e 22                             |
| ENGENHARIA DE SOFTWARE                    | 13, 15, 19 e 25                     | 13, 15, 19 e 25                     |
| CIRCUITOS LÓGICOS                         | 16                                  | 16                                  |
| ALGORITMOS E PROGRAMAÇÃO DE COMPUTADORES  | 20 e 33                             | 20                                  |
| REDES DE COMPUTADORES                     | 21                                  | 21                                  |
| SEGURANÇA COMPUTACIONAL                   | 24                                  | 24                                  |
| INTERAÇÃO HUMANO COMPUTADOR               | 26                                  | 26                                  |
| PROBABILIDADE E ESTATÍSTICA               | 27                                  | -                                   |
| ORGANIZAÇÃO E ARQUITETURA DE COMPUTADORES | 28                                  | -                                   |
| INTRODUCAO AO PROCESSAMENTO DE IMAGENS    | 29                                  | -                                   |
| COMPILADORES                              | 30                                  | -                                   |
| AUTÔMATOS E COMPUTABILIDADE               | 31                                  | -                                   |
| PROJETO E ANÁLISE DE ALGORITMOS           | 32                                  | -                                   |
| FORMAÇÃO DOCENTE EM COMPUTAÇÃO            | -                                   | 27, 28, 29, 30, 31, 32, 33, 34 e 35 |

#### 4.4 Análise do questionário do estudante

O questionário do estudante é aplicado em conjunto com a prova e é destinado a coletar informações relevantes sobre o perfil dos participantes. Os dados obtidos através deste questionário são usados para contextualizar os resultados das provas e fornecer uma visão mais ampla sobre a educação superior no Brasil, é fundamental para fornecer um contexto

mais amplo sobre a qualidade do ensino superior no Brasil. A UnB será a universidade específica para analisar as informações do questionário.

#### 4.4.1 Análise de gênero nos cursos de Computação da UnB

A Figura 4.14 mostra um gráfico de barras representando a distribuição de gênero dos alunos da Universidade de Brasília que participaram do ENADE. O gráfico compara dois cursos, Ciência Da Computação Bacharelado e Computação Licenciatura.

O gráfico reflete um maior número de alunos do sexo masculino nos cursos de Ciência da Computação tanto em nível de bacharelado quanto de licenciatura na UnB.

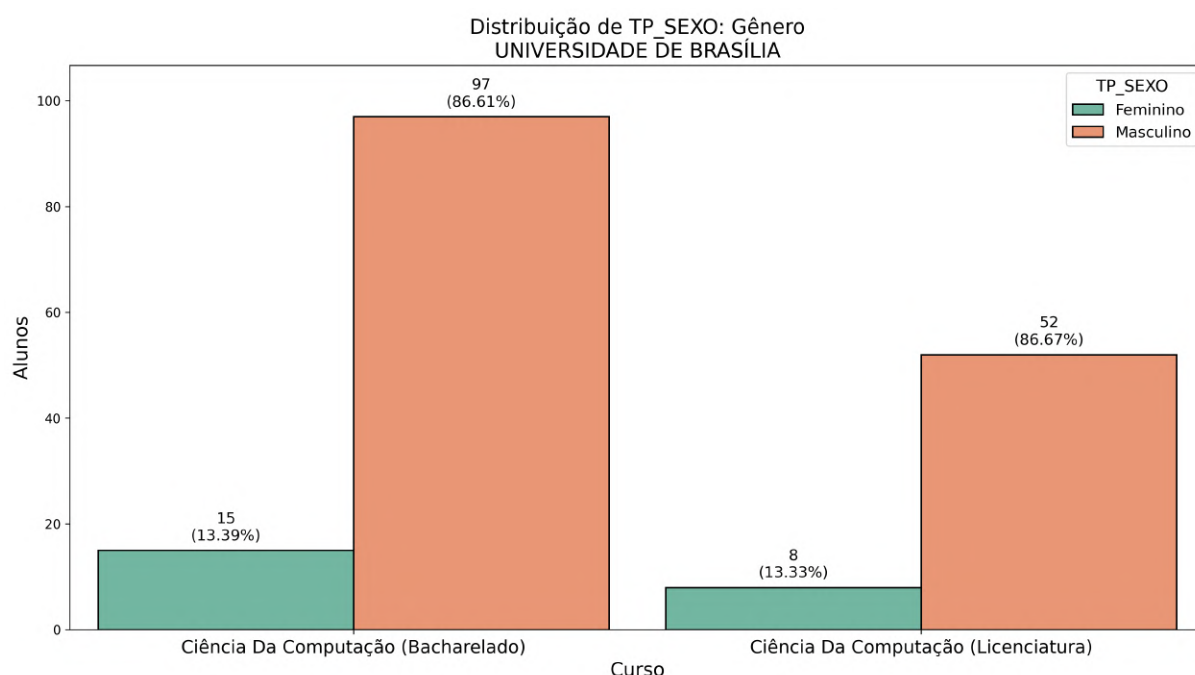


Figura 4.14: Gráfico do sexo dos alunos de Ciência Da Computação.

#### 4.4.2 Análise do estado civil nos cursos de Computação da UnB

A Figura 4.15 representa a distribuição dos estados civis dos alunos da Universidade de Brasília, os dados do estado civil dos estudantes podem ser importantes porque podem influenciar questões como a necessidade de serviços de apoio aos estudantes, a disponibilidade para participar de atividades extracurriculares ou estágios.

A maioria dos alunos de ambos os cursos, identifica-se como solteiro(a), o que é comum em populações universitárias compostas principalmente por jovens adultos. Também pode indicar que os alunos tendem a não ter compromissos matrimoniais, o que pode refletir

maior flexibilidade em termos de compromissos pessoais e disponibilidade de tempo para se dedicarem aos estudos e atividades relacionadas ao curso.

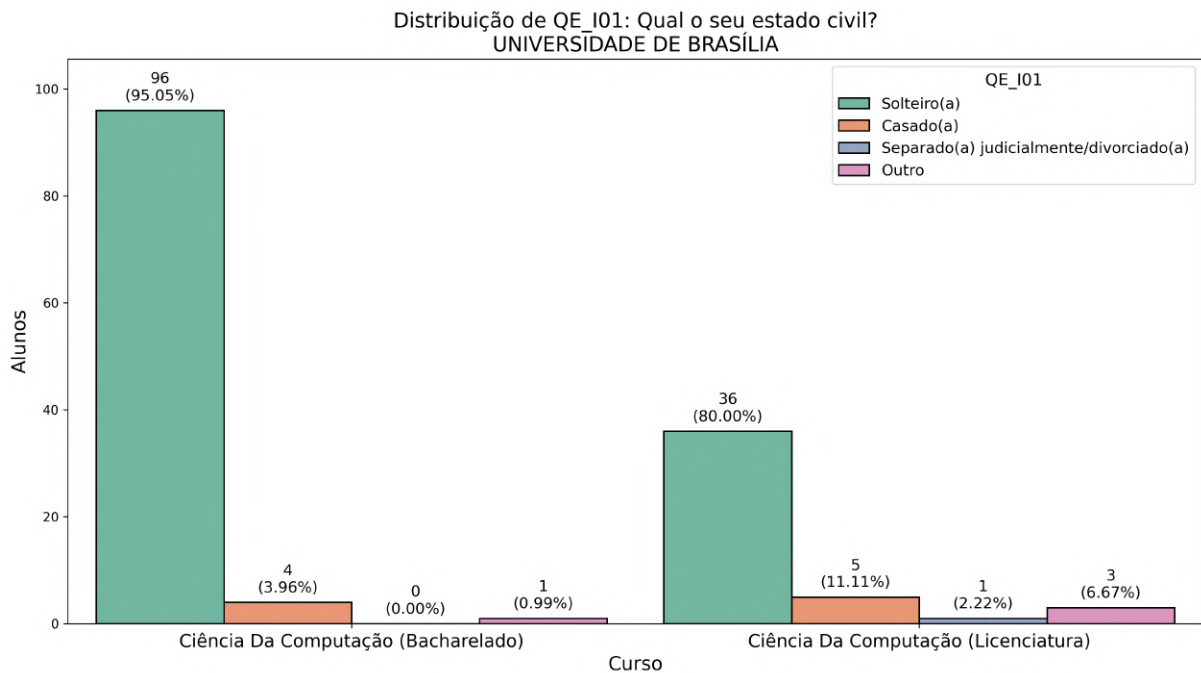


Figura 4.15: Gráfico do estado civil dos alunos de Ciência Da Computação.

#### 4.4.3 Análise da cor/raça nos cursos de Computação da UnB

A Figura 4.16 é um gráfico de barras que mostra a distribuição dos alunos por cor ou raça, os dados sobre a cor ou raça dos alunos são importantes para avaliar a diversidade dentro da instituição e podem ajudar a universidade a entender se suas políticas de inclusão estão alcançando uma ampla representatividade de grupos raciais e étnicos.

A UnB possui uma população estudantil racialmente diversa em ambos os cursos analisados. Há variações entre eles, o que pode ser reflexo de diferenças nas políticas de ingresso, na atratividade dos cursos ou em fatores socioeconômicos dos candidatos. O gráfico indica que tanto no curso de Bacharelado e Licenciatura, a maioria dos alunos se identifica como da cor ou raça branca e o segundo grupo tem maior representatividade é da cor ou raça parda.

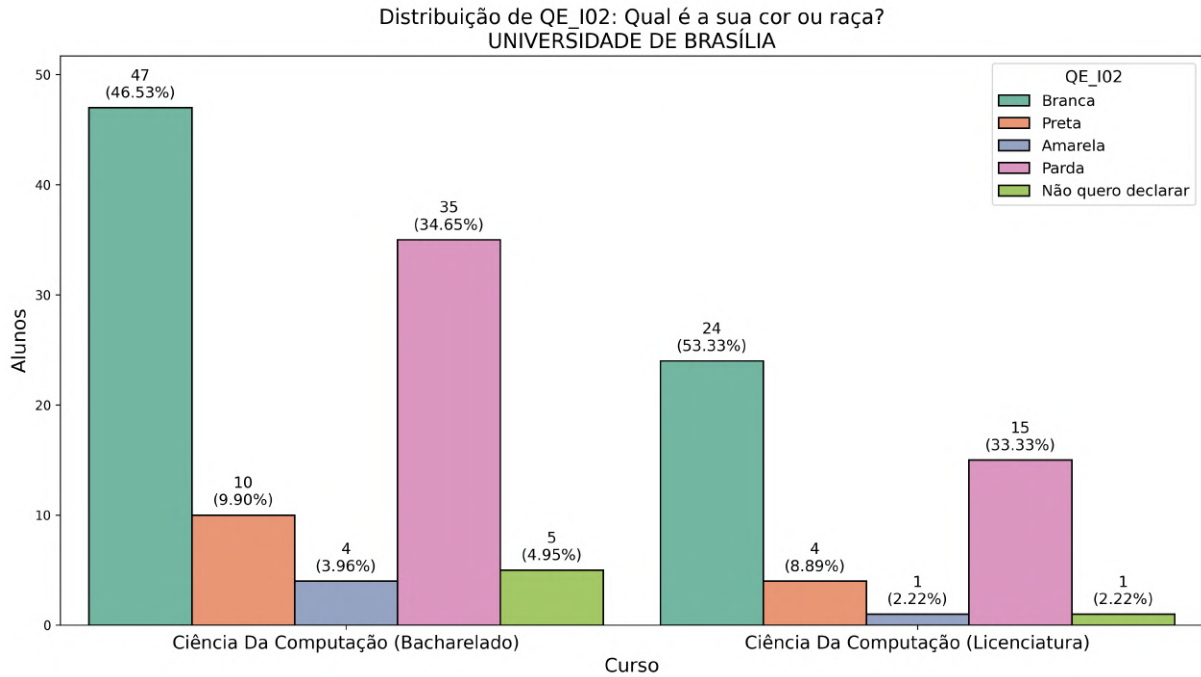


Figura 4.16: Gráfico da cor/raça dos alunos de Ciência Da Computação.

#### 4.4.4 Análise da nacionalidade nos cursos de Computação da UnB

A Figura 4.17 representa a distribuição de nacionalidades dos alunos. A informação é importante para compreender a demografia dos alunos em termos de nacionalidade dentro dos cursos de Ciência da Computação.

A totalidade dos alunos ser da mesma nacionalidade pode refletir um perfil de estudante uniforme em termos de origem nacional. Isso também pode apontar para oportunidades de diversificar o corpo discente, introduzindo programas de intercâmbio ou incentivos para estudantes estrangeiros.

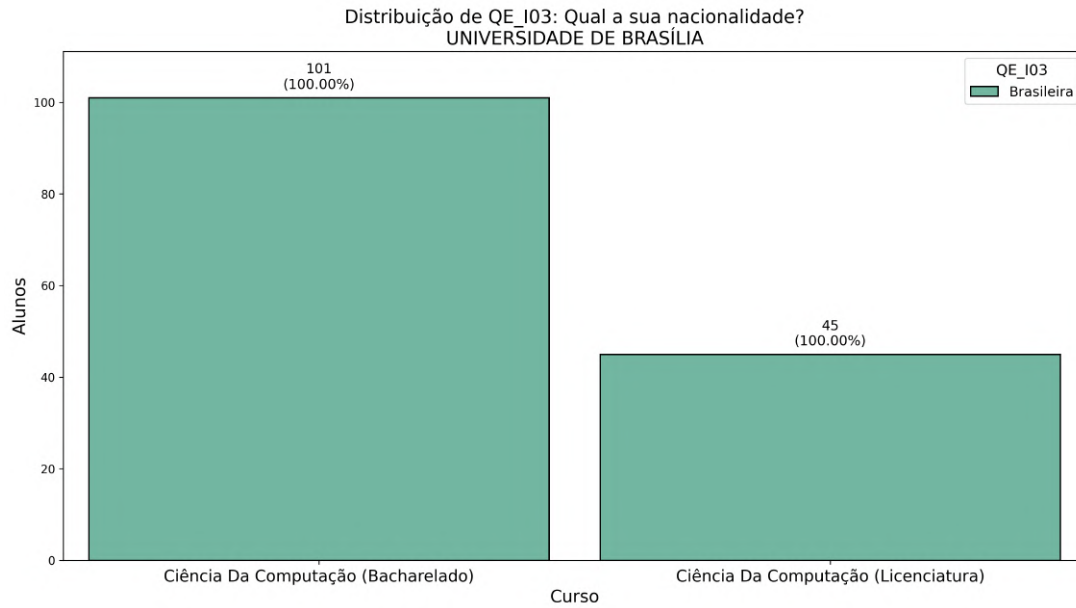


Figura 4.17: Gráfico da nacionalidade dos alunos de Ciência Da Computação.

#### 4.4.5 Análise do ambiente familiar nos cursos de Computação da UnB

A Figura 4.18 e Figura 4.19 representam o nível de escolaridade dos pais. A maioria dos alunos provém de contextos com ao menos nível superior. Isto pode contribuir para uma predisposição positiva em relação ao aprendizado e ao sucesso acadêmico, além de potencialmente influenciar as aspirações profissionais dos alunos. A representação significativa de níveis mais altos de escolaridade entre as mães e pais aponta para uma possível correlação para a escolha do curso superior pelos alunos.

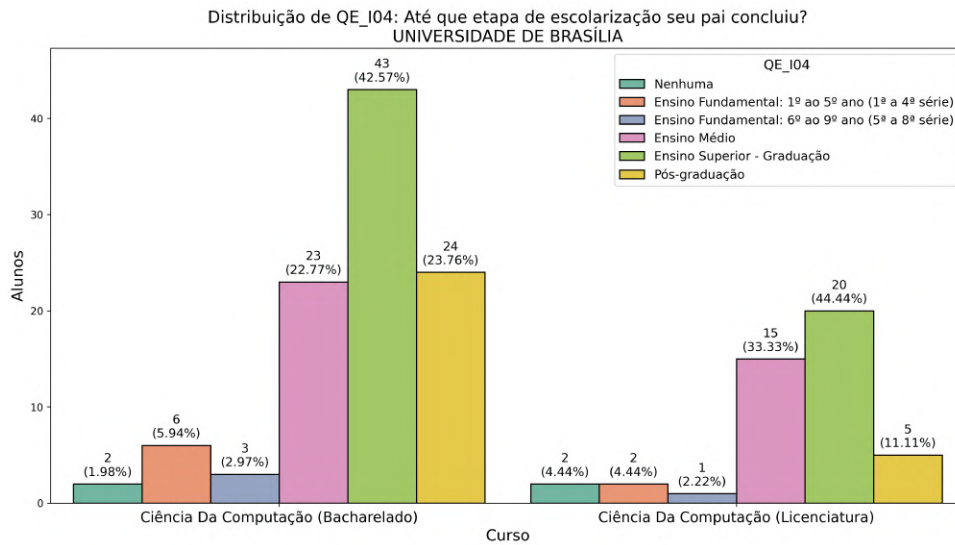


Figura 4.18: Gráfico da escolaridade do pai dos alunos de Ciência da Computação.

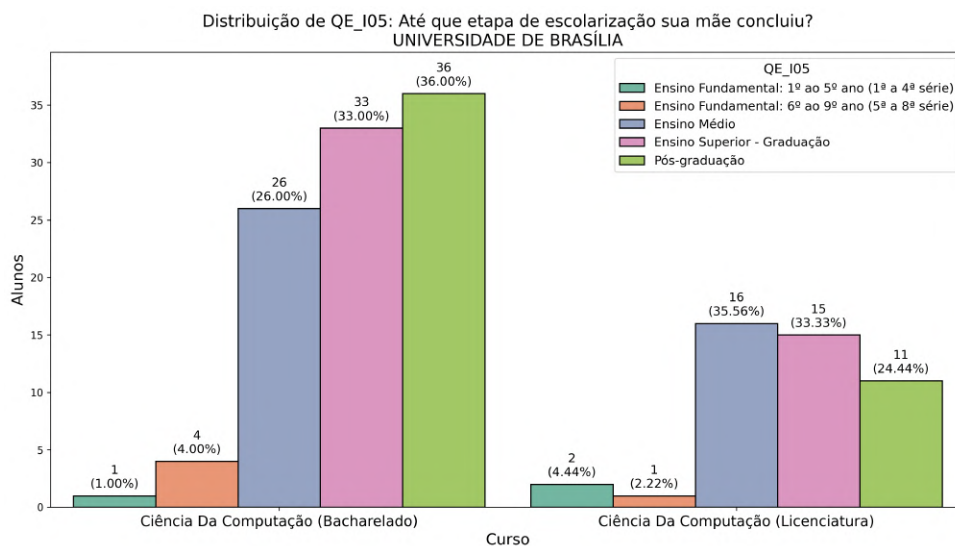


Figura 4.19: Gráfico da escolaridade da mãe dos alunos de Ciência da Computação.

A Figura 4.20 representa as condições de moradia dos alunos, é um fator importante que pode influenciar o desempenho acadêmico e o bem-estar geral. Estar em um ambiente estável e seguro pode contribuir para um melhor foco nos estudos.

Os alunos apresentam um perfil que combina apoio familiar com responsabilidades pessoais. A condição de moradia, junto com o contexto educacional dos pais, sugere que muitos desses alunos podem ter um ambiente propício ao sucesso acadêmico, embora também possam enfrentar desafios individuais relacionados à independência e gestão do tempo.

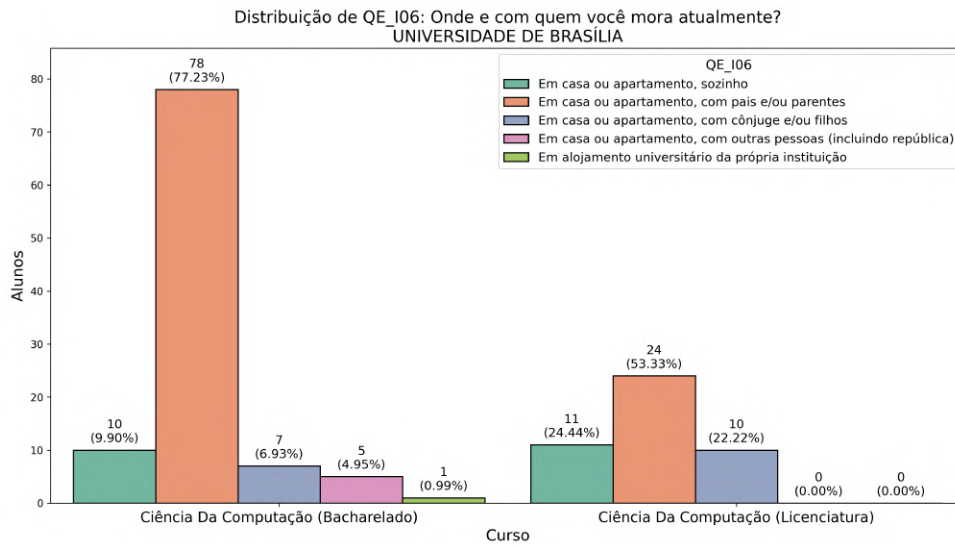


Figura 4.20: Gráfico da moradia dos alunos de Ciência da Computação.

A Figura 4.21 representa a quantidade de membros da família que vivem com os alunos. Esse dado pode influenciar a rotina de estudos dos alunos, viver com muitos membros da família pode fornecer um sistema de apoio, mas também pode criar um ambiente mais barulhento ou distrair, o que pode impactar o estudo e o tempo de lazer.

No curso de Bacharelado, a maioria dos alunos vive com dois membros da família, enquanto no curso de Licenciatura, a distribuição é mais variada, com uma quantidade notável de estudantes vivendo com nenhuma ou apenas uma pessoa, o que pode indicar famílias menores. A ausência de alunos vivendo com sete ou mais membros da família em ambos os cursos sugere que famílias muito grandes não são comuns entre essa população estudantil.

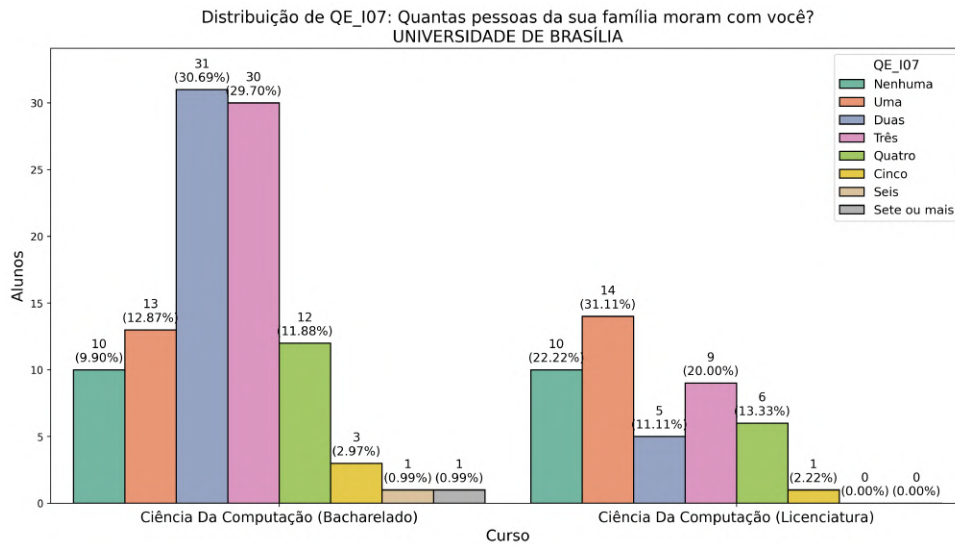


Figura 4.21: Gráfico da quantidade de membros que moram com os alunos de Ciência Da Computação.

A Figura 4.22 representa a distribuição da renda familiar dos alunos dos cursos, entender a renda familiar dos alunos é crucial para qualquer instituição educacional. Isso permite que a universidade compreenda melhor a diversidade socioeconômica de seu corpo discente.

Observamos que a maior proporção de alunos de bacharelado tem uma faixa de renda familiar de 10 a 30 salários mínimos. Em contraste, os alunos de licenciatura outras faixas de renda, como de 4 a 5 salários mínimos e de 6 a 10 salários mínimos têm uma representação considerável.

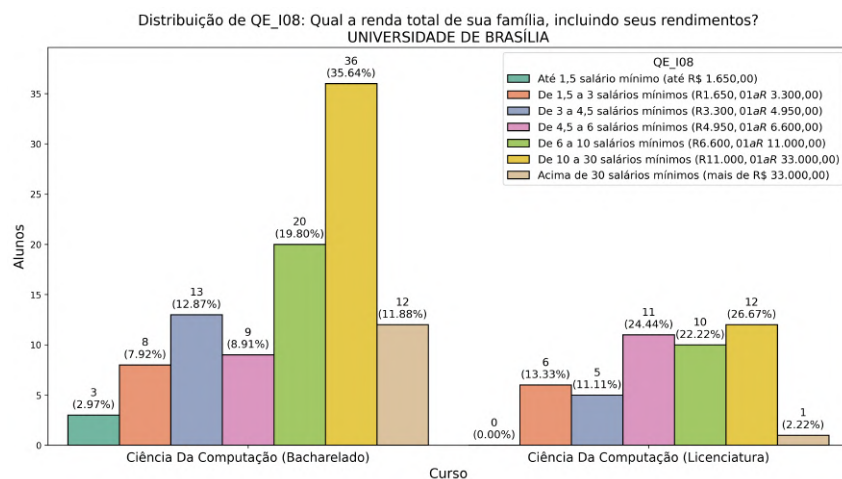


Figura 4.22: Gráfico da renda familiar dos alunos de Ciência Da Computação.



A Figura 4.23 e Figura 4.24 mostra a situação de renda e trabalho dos alunos, compreender essa informação é essencial para avaliar como o emprego pode afetar o desempenho acadêmico, a carga horária de estudo, e como os alunos conseguem conciliar trabalho e estudo.

Muitos alunos trabalham 40 horas ou mais por semana e uma proporção significativa contribui para o sustento da família, o que pode sugerir que esses alunos possuem maiores responsabilidades financeiras e, possivelmente, menos tempo para se dedicar exclusivamente aos estudos. Mas boa parte dos alunos de Bacharelado não trabalham atualmente, o que pode indicar que eles têm mais tempo para se dedicar aos estudos.

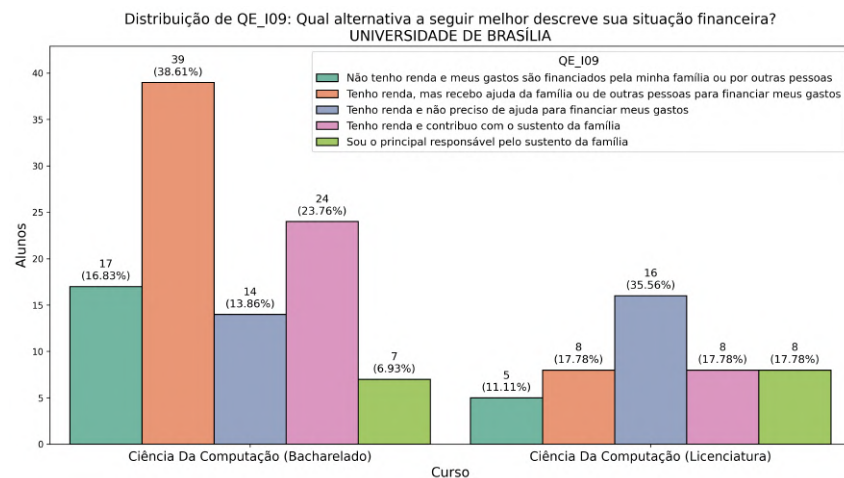


Figura 4.23: Gráfico da situação de trabalho dos alunos de Ciência da Computação.

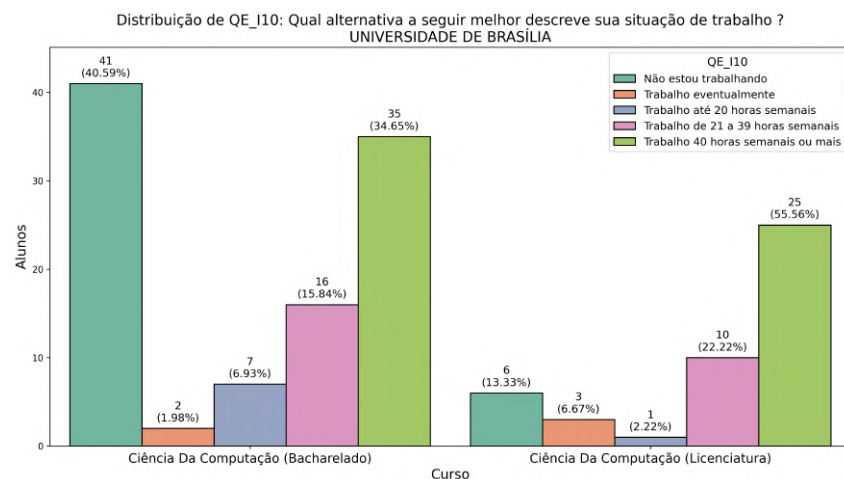


Figura 4.24: Gráfico da situação financeira dos alunos de Ciência da Computação.

#### 4.4.6 Análise do perfil educacional nos cursos de Computação da UnB

As Figura 4.25 e Figura 4.26 representam a unidade da federação onde os estudantes concluíram o ensino médio. Esses dados ajudam a entender o alcance geográfico da instituição, indicando se a universidade é uma opção atraente para estudantes de todo o país ou se atende principalmente à região local.

Comparando ambos os cursos, percebe-se uma tendência similar na distribuição geográfica dos estudantes com uma predominância clara do Distrito Federal. Isso pode ser atribuído à localização da universidade, facilitando o acesso para aqueles que já residem na região. A presença de estudantes de outros estados em ambas as modalidades também indica que a UnB tem a capacidade de atrair alunos de fora, embora em menor escala quando comparado aos residentes locais.

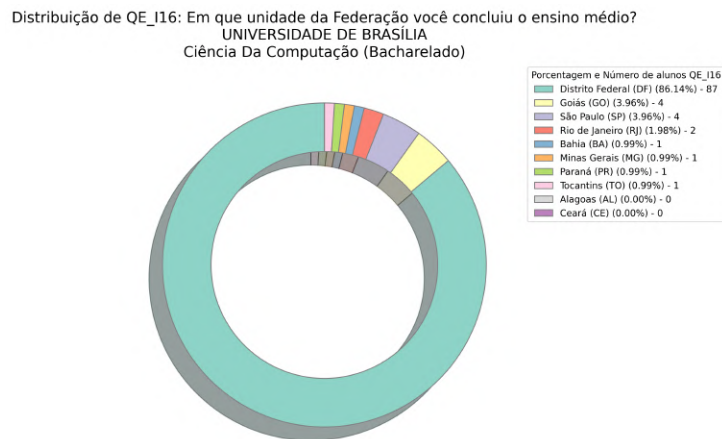


Figura 4.25: Gráfico da unidade federativa onde o aluno concluiu o ensino médio (Bacharelado).

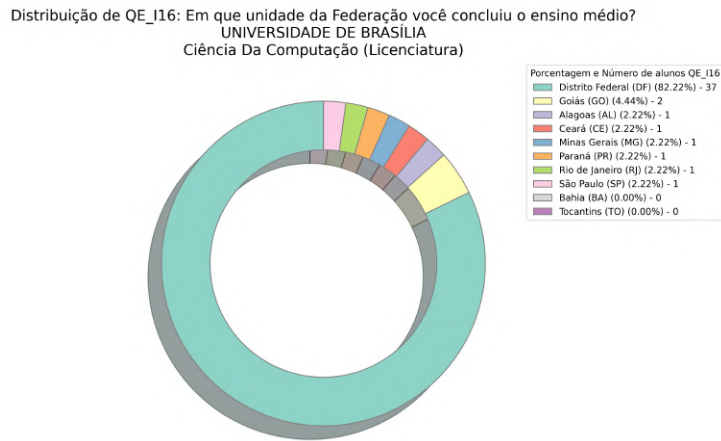


Figura 4.26: Gráfico da unidade federativa onde o aluno concluiu o ensino médio (Licenciatura).

Complementado essa informação, a Figura 4.27 mostra o tipo de escola onde cursaram o ensino médio. A origem educacional pode ter implicações na preparação acadêmica dos estudantes, no acesso a oportunidades e recursos, e nas políticas de inclusão e diversidade da instituição.

O fato de a maioria dos alunos ter cursado o ensino médio no Distrito Federal e em escolas privadas pode indicar que a Universidade de Brasília atrai principalmente alunos de um certo perfil socioeconômico da região local.

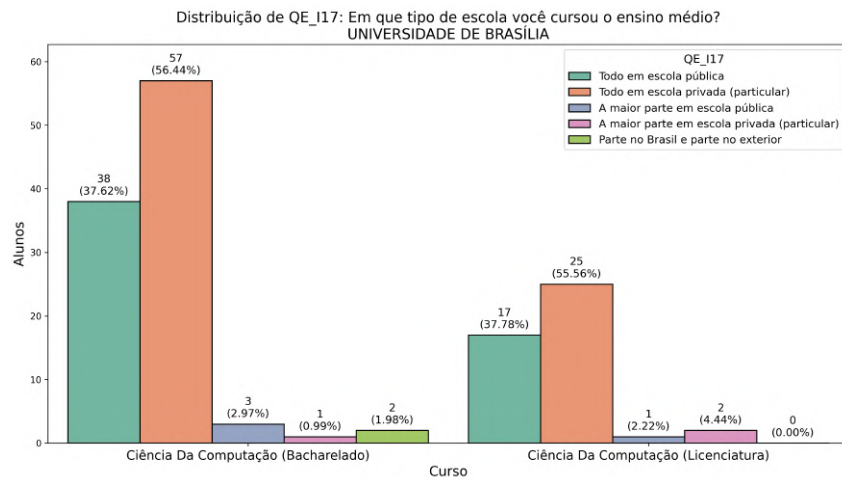


Figura 4.27: Gráfico do tipo de escola dos alunos de Ciência da Computação.

A Figura 4.28 mostra a distribuição dos alunos dos cursos segundo a modalidade de ensino médio que concluíram. A diversidade nas modalidades de ensino médio concluídas pode refletir na capacidade da universidade de atrair um corpo discente diversificado.

Podemos observar que a maioria dos alunos de ambos os cursos vem de um contexto de ensino médio tradicional, o que pode estar alinhado com a conclusão anterior de que muitos estudantes vêm de escolas privadas.

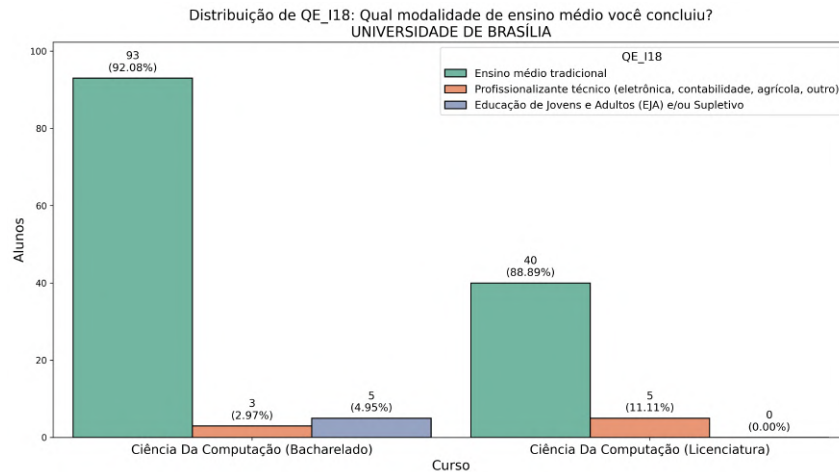


Figura 4.28: Gráfico da modalidade de ensino médio dos alunos de Ciência da Computação.

A Figura 4.29 categoriza os alunos com base na quantidade de livros lidos. Este pode ser um indicador importante do envolvimento com a leitura fora do currículo acadêmico, que pode ser um fator no desenvolvimento de habilidades críticas de pensamento e compreensão.

Podemos observar que os alunos de ambos os cursos dedicam tempo à leitura, embora o número de livros lidos seja relativamente modesto. Isso pode estar relacionado com o tempo disponível após estudos e trabalho, como mencionado nas análises anteriores.

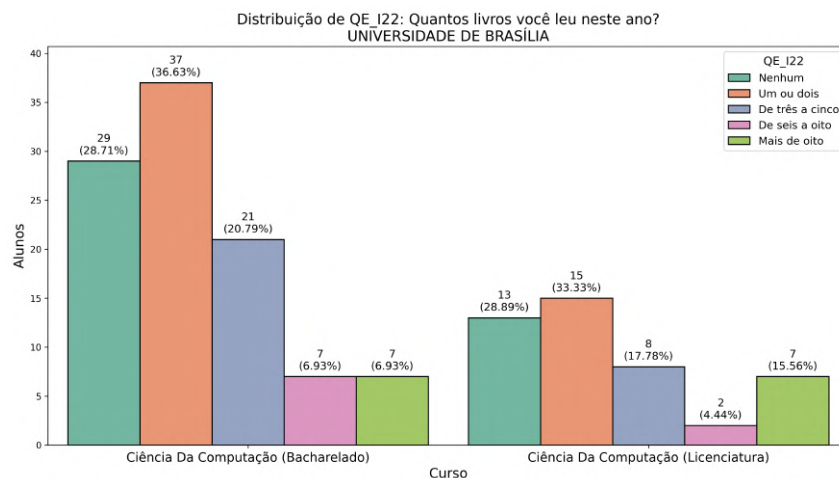


Figura 4.29: Gráfico da quantidade de livros lidos pelos alunos de Ciência da Computação.

A Figura 4.30 apresenta um gráfico que compara a quantidade de horas dedicadas aos estudos por semana pelos alunos. Este dado pode refletir o equilíbrio entre estudo e outras atividades, como trabalho ou lazer, e pode ser fundamental para identificar necessidades de apoio acadêmico ou ajustes na carga de trabalho do curso.

Constatamos que os alunos de ambos os cursos se comprometem com o estudo fora do tempo de aula, mas com uma tendência ligeiramente maior de dedicação no curso de Bacharelado. Isso pode estar relacionado à natureza do curso ou às demandas acadêmicas específicas. Além disso, a minoria de alunos que não estuda fora das aulas pode precisar de atenção especial para garantir que estão acompanhando o conteúdo do curso.

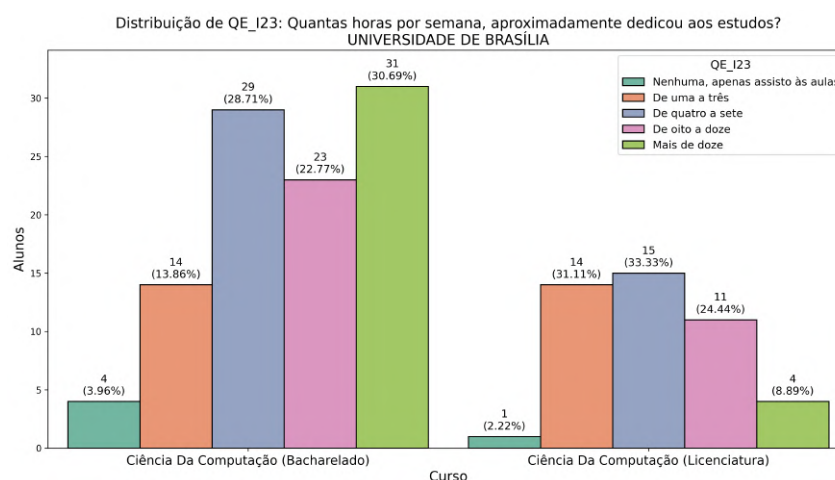


Figura 4.30: Gráfico da quantidade de livros lidos pelos alunos de Ciência da Computação.

Finalmente, a Figura 4.31 compara a disponibilidade de oportunidades de aprendizado de idiomas estrangeiros para alunos. O acesso ao aprendizado de idiomas estrangeiros é um fator importante no desenvolvimento educacional dos alunos, pois pode melhorar as habilidades de comunicação global e aumentar as oportunidades de emprego, especialmente em campos como o da computação, que frequentemente requer colaboração internacional.

Fica claro que, em ambos os cursos, existe uma parcela de alunos que não teve oportunidade de aprendizado de idiomas dentro da instituição. Isso pode ser um ponto de atenção para a universidade, considerando a relevância de habilidades multilíngues na área de tecnologia e computação. Para os alunos que tiveram essa oportunidade, a modalidade presencial parece ser a mais comum.

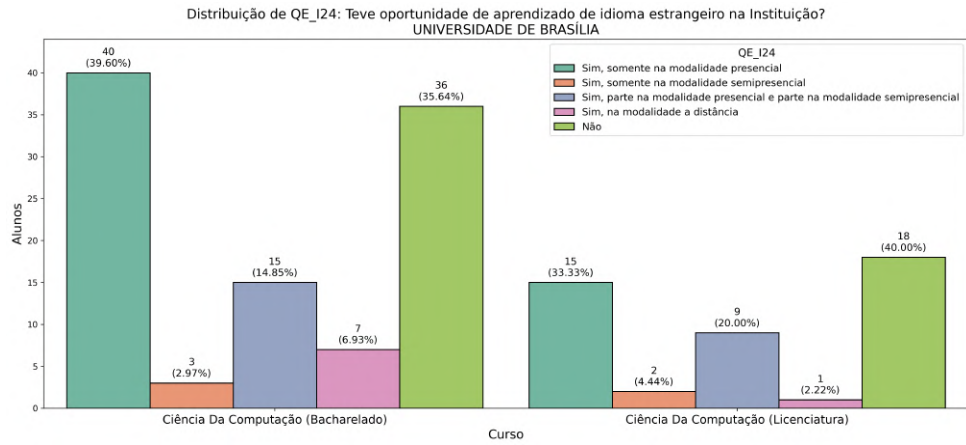


Figura 4.31: Gráfico do aprendizado de idiomas estrangeiros nos alunos de Ciência da Computação.

# Capítulo 5

## Conclusões

Neste trabalho foram desenvolvidas análises sobre os diferentes dados do ENADE do ano 2021. O principal objetivo foi avaliar o desempenho dos alunos da UnB dos cursos Ciência da Computação (Bacharelado) e Computação (Licenciatura) em comparação com outras instituições.

Para cumprir os objetivos específicos, inicialmente foram coletados os microdados do ENADE do site do INEP. Inicialmente, a estrutura deles não era a adequada para fazer as diferentes análises. Por isso foi necessário baixar um arquivo auxiliar do site do MEC, que tem mais informações sobre cursos de graduação do Brasil, como o nome da instituição de ensino superior, para fazer o mapeamento com os arquivos dos microdados que tinham uma informação bem limitada. Também foi feito o mapeamento dos arquivos com a consulta do dicionário de arquivos e variáveis, para entender a estrutura de cada arquivo e fazer o mapeamento correspondente mediante *scripts* em Python específicos para cada arquivo. Depois de fazer o pré-processamento dos dados, foram feitas as análises mediante *scripts*, para inicialmente fazer uma análise geral dos alunos participantes do ENADE 2021, para saber quais cursos participaram, número de alunos inscritos e número de alunos que realizaram a prova.

Para apoiar na análise, optou-se pela geração de *heatmaps* para estes dados, com dados do sexo, estado civil, raça, renda familiar e informações sobre o ensino médio que cursaram os participantes. São eficazes para mostrar distribuições de dados e variações dentro de um conjunto de dados.

Realizou-se uma análise das notas dos estudantes dos cursos participantes, abrangendo a mediana das notas *NT\_FG*, *NT\_CE* e *NT\_GER*. Depois foi feita uma comparação específica de todos os alunos do ENADE com os alunos de Ciência da Computação (Bacharelado e Licenciatura). Complementando essa informação, foram gerados *boxplots* que fornecem um resumo visual de importantes indicadores estatísticos: mediana, quartis superior e inferior, e valores extremos. Para a análise dos vetores de notas, foram utilizados

os dados de uma tabela do INEP que tinha as notas já calculadas do “Conceito ENADE”. Na seleção de instituições para ser analisadas, foram escolhidas aquelas com as maiores pontuações na coluna “Conceito ENADE”, sendo a maioria universidades federais. Para essas 20 instituições, foi feita uma tabela para cada curso com os estudantes de Ciência da Computação Bacharelado e Licenciatura separadamente com os dados do número de alunos inscritos e número de alunos que realizaram a prova. Foram analisadas apenas as questões objetivas da prova, 8 questões do componente de formação geral e 27 questões do componente específico. Na tabela gerada com o percentual de acertos das 20 instituições, foi destacada a UnB para poder avaliar o desempenho em comparação com outras instituições. Finalmente, foi analisado o questionário do estudante, que tem informações sobre o perfil dos participantes. Foram analisados aspectos como sexo, estado civil, ambiente familiar e perfil educacional dos alunos.

Este trabalho permitiu a melhor visualização sobre os dados do ENADE e o desempenho dos participantes, incluindo a criação de tabelas e gráficos para então fazer uma análise mais detalhada e avaliar quais fatores podem afetar o desempenho dos alunos no exame.

Os materiais completos deste projeto estão disponíveis no repositório do *GitHub*<sup>1</sup>.

Para trabalhos futuros relacionados à análise de dados do ENADE, pode ser feito uma interface gráfica para que o usuário possa fazer a análise de dados mais acessível. Como exemplo, pode ser um *dashboard* dinâmico, o que ajudaria também na geração de informação. Além disso, seria bom que, na parte de geração de tabelas, o *script* possa identificar as questões que foram anuladas para cada curso, entre outras automatizações.

---

<sup>1</sup>Disponível em: [https://github.com/destandroid/TCC\\_ENADE2021](https://github.com/destandroid/TCC_ENADE2021)



# Referências

- [1] IA.BB AI ENTHUSIASTS, FRIENDS e MAINLY FROM BANCO DO BRASIL RESEARCHERS: *Crisp-dm — organizando seu projeto de mineração de dados*. <https://medium.com/@ia.bb/crisp-dm-organizing-your-data-mining-project-fdb0399cbd9b>. Publicado em 29 de Junho. Acesso em: 20 de agosto de 2023. ix, 9
- [2] EXAME: *Temos mais dados do que nunca. como usá-los a nosso favor?* <https://exame.com/carreira/dados-uso-favor/>, 2021. Publicado em 9 de junho de 2021 às 12h36. Última atualização em 7 de julho de 2021 às 14h16. Acesso em: 20 de agosto de 2023. 1
- [3] INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA: *Exame nacional de desempenho dos estudantes - enade*. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enade>, 2023. Acesso em: 18 de abril de 2023. 1, 4
- [4] VAN ROSSUM, G. e F. L. DRAKE: *Python 3 Reference Manual*. CreateSpace Independent Publishing Platform, 2019. 2, 10
- [5] INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA: *Nota técnica nº 20/2019 - cálculo da nota final do enade*. [https://download.inep.gov.br/educacao\\_superior/enade/notas\\_tecnicas/2019/nota\\_tecnica\\_n20-2019\\_CGCQES-DAES\\_calculo\\_NF\\_Enade.pdf](https://download.inep.gov.br/educacao_superior/enade/notas_tecnicas/2019/nota_tecnica_n20-2019_CGCQES-DAES_calculo_NF_Enade.pdf), 2019. Acesso em: 15 de agosto de 2023. 5
- [6] KITCHIN, R. e T. P. LAURIAULT: *Towards critical data studies: Charting and unpacking data assemblages and their work*. The Programmable City Working Paper, 2014. 7
- [7] DAVENPORT, T. H. e L. PRUSAK: *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Press, 1998. 7
- [8] HAN, J., M. KAMBER e J. PEI: *Data Mining: Concepts and Techniques*. 2011. 8, 15
- [9] PYLE, D.: *Data Preparation for Data Mining*. 1999. 8
- [10] SHMUELI, G. e N. R. PATEL: *Data Mining for Business Intelligence*. 2010. 8

- [11] IBM: *Cross-industry standard process for data mining (crisp-dm)*. [https://inseaddataanalytics.github.io/INSEADAnalytics/CRISP\\_DM.pdf](https://inseaddataanalytics.github.io/INSEADAnalytics/CRISP_DM.pdf). Acesso em: 13 de maio de 2023. 8
- [12] DATAGEEEKS: *Pre-processamento de dados*. <https://www.datageeks.com.br/pre-processamento-de-dados/>, 2023. Acesso em: 15 de agosto de 2023. 8, 15
- [13] PROJECT JUPYTER: *Jupyter documentation*, 2023. <https://docs.jupyter.org/en/latest/>, Acesso em: 13 de maio de 2023. 10
- [14] MICROSOFT: *Visual Studio Code Documentation*. Microsoft, 2019. 11
- [15] PANDAS DEVELOPMENT TEAM: *pandas user guide*. [https://pandas.pydata.org/docs/user\\_guide/index.html](https://pandas.pydata.org/docs/user_guide/index.html), 2023. Acesso em: 15 de agosto de 2023. 11
- [16] MATPLOTLIB DEVELOPMENT TEAM: *Matplotlib user's guide*. <https://matplotlib.org/stable/users/index.html>, 2023. Acesso em: 15 de agosto de 2023. 11
- [17] SEABORN DEVELOPMENT TEAM: *Seaborn tutorial*. <https://seaborn.pydata.org/tutorial.html>, 2023. Acesso em: 15 de agosto de 2023. 11
- [18] UNIDECODE DEVELOPMENT TEAM: *Unidecode on pypi*. <https://pypi.org/project/Unidecode/>, 2023. Acesso em: 15 de agosto de 2023. 11
- [19] NUMPY DEVELOPMENT TEAM: *Numpy user guide*. <https://numpy.org/doc/stable/user/index.html>, 2023. Acesso em: 15 de agosto de 2023. 11
- [20] BRASIL: *Lei geral de proteção de dados pessoais (lgpd)*. Lei nº 13.709, 2018. 11
- [21] INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP): *Enem e enade têm novo conjunto de microdados publicados*. Disponível em: <https://www.gov.br/inep/pt-br/assuntos/noticias/institucional/enem-e-enade-tem-novo-conjunto-de-microdados-publicados>, 2022. 18 de abril de 2023. 12
- [22] OLIVEIRA, A: *Coleta de dados: Quais são os métodos e como fazer?* <https://mindminers.com/blog/coleta-de-dados-como-fazer/>, 2023. Acesso em: 10 de dezembro de 2023. 13
- [23] INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA: *Microdados enade*. <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enade>, 2022. Acesso em: 18 de abril de 2023. 14
- [24] EDUCAÇÃO, MINISTÉRIO DA: *Indicadores sobre ensino superior*. <https://dadosabertos.mec.gov.br/indicadores-sobre-ensino-superior>, 2022. Acesso em: 18 de abril de 2023. 14

- [25] TEIXEIRA, INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO: *Dicionário de arquivos e variáveis dos microdados do enade 2021*. Anexo nos Microdados do ENADE 2021, 2022. Disponível em: <https://www.gov.br/inep/pt-br/area-de-atuacao/dados-abertos/microdados/enade>. 15, 18
- [26] PAUL, E: *Um detalhe simples para melhorar seu heatmap*. <https://medium.com/@eduardopaul/um-detalhe-simples-para-melhorar-seu-heatmap-2ccff36a4391>, 2021. Acesso em: 15 de agosto de 2023. 21
- [27] UNIVERSIDADE FEDERAL DO PARANÁ: *Média aritmética – média ponderada – moda – mediana*. [https://docs.ufpr.br/~prbg/public\\_html/ce003/central.pdf](https://docs.ufpr.br/~prbg/public_html/ce003/central.pdf). Acesso em: 15 de agosto de 2023. 31
- [28] SILVA, Fernando da: *Análise exploratória de dados com o gráfico de box-plot*. <https://analisemacro.com.br/econometria-e-machine-learning/analise-exploratoria-de-dados-com-o-grafico-de-boxplot/>, 2023. Acesso em: 15 de agosto de 2023. 35
- [29] INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA: *Nota técnica nº 7/2022 - metodologia de cálculo do conceito enade 2021*. [https://download.inep.gov.br/educacao\\_superior/enade/notas\\_tecnicas/2019/nota\\_tecnica\\_n\\_7\\_2022\\_CGCQES\\_DAES\\_metodologia\\_calculo\\_conceito\\_enade\\_2021.pdf](https://download.inep.gov.br/educacao_superior/enade/notas_tecnicas/2019/nota_tecnica_n_7_2022_CGCQES_DAES_metodologia_calculo_conceito_enade_2021.pdf), 2022. Acesso em: 15 de agosto de 2023. 38
- [30] INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA: *Conceito enade 2021*. <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/indicadores-de-qualidade-da-educacao-superior/outros-documentos>, 2022. Publicado em 12 de setembro de 2022. Acesso em: 15 de agosto de 2023. 38
- [31] NAKAGAWA, Shinichi e Innes C. CUTHILL: *Effect size, confidence interval and statistical significance: a practical guide for biologists*. <https://onlinelibrary.wiley.com/doi/10.1111/j.1469-185X.2007.00027.x>, 2007. Acesso em: 05/02/2024. 40
- [32] DAMÁSIO, B: *O que é correlação ponto-bisserial?* <https://psicometriaonline.com.br/o-que-e-correlacao-ponto-biserial/>, 2021. Acesso em: 18 de dezembro de 2023. 43
- [33] INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA: *Provas e gabaritos do enade*. <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enade/provas-e-gabaritos>, 2023. Acesso em: 15 de agosto de 2023. 48