



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
PRÓ-REITORIA DE GRADUAÇÃO
DEPARTAMENTO DE CIÊNCIAS EXATAS E TECNOLOGIA DA INFORMAÇÃO
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO

CARLOS ALBERTO SIQUEIRA DA FONSECA

**ANÁLISE DESCRITIVA DOS DADOS EDUCACIONAIS SOBRE
A EVASÃO NOS CURSOS DE COMPUTAÇÃO DA UFERSA**

ANGICOS

2025

CARLOS ALBERTO SIQUEIRA DA FONSECA

**ANÁLISE DESCRITIVA DOS DADOS EDUCACIONAIS SOBRE
A EVASÃO NOS CURSOS DE COMPUTAÇÃO DA UFERSA**

Trabalho de Conclusão de Curso apresentado à
Universidade Federal Rural do Semi-Árido
como requisito para obtenção do título de
Bacharel em Sistemas de Informação.

Orientadora: Adriana Mara Guimarães de
Farias, Profa. Dra.

ANGICOS

2025

© Todos os direitos estão reservados a Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996 e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tomar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata. A mesma poderá servir de base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) sejam devidamente citados e mencionados os seus créditos bibliográficos.

F676a Fonseca, Carlos Alberto Siqueira da.
Análise descritiva dos dados educacionais sobre
a evasão nos cursos de computação da Ufersa /
Carlos Alberto Siqueira da Fonseca. - 2025.
39 f. : il.

Orientadora: Adriana Mara Guimarães de Farias.
Monografia (graduação) - Universidade Federal
Rural do Semi-árido, Curso de Sistemas
de Informação, 2025.

1. análise descritiva. 2. evasão. 3. dados
educacionais. 4. cursos. 5. python. I. Guimarães
de Farias, Adriana Mara, orient. II. Título.

Ficha catalográfica elaborada por sistema gerador automático em conformidade
com AACR2 e os dados fornecidos pelo(a).

Biblioteca Campus Angicos

Bibliotecário: Helder Romero Maia Duarte CRB: 15/673

O serviço de Geração Automática de Ficha Catalográfica para Trabalhos de Conclusão de Curso (TCC's) foi desenvolvido pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP) e gentilmente cedido para o Sistema de Bibliotecas da Universidade Federal Rural do Semi-Árido (SISBI-UFERSA), sendo customizado pela Superintendência de Tecnologia da Informação e Comunicação (SUTIC) sob orientação dos bibliotecários da instituição para ser adaptado às necessidades dos alunos dos Cursos de Graduação e Programas de Pós-Graduação da Universidade.

Carlos Alberto Siqueira da Fonseca

**ANÁLISE DESCRITIVA DOS DADOS EDUCACIONAIS SOBRE
A EVASÃO NOS CURSOS DE COMPUTAÇÃO DA UFERSA**

Trabalho de Conclusão de Curso apresentado à
Universidade Federal Rural do Semi-Árido
como requisito para obtenção do título de
Bacharel em Sistemas de Informação.

Defendida em: 25/07/2025.

BANCA EXAMINADORA

Adriana Mara Guimarães de Farias, Prof^ª. Dr^ª. (UFERSA)
Presidente

Joemia Leilane Gomes de Medeiros, Prof^ª. Dr^ª. (UFERSA)
Membro Examinador

Sairo Raoni dos Santos, Prof. Dr. (UFERSA)
Membro Examinador

Dedico este trabalho aos meus pais, meus irmãos, a minha esposa e ao meu filho, que amo incondicionalmente, eles todos são como uma válvula propulsora para minha busca incessante por conhecimento.

AGRADECIMENTOS

Agradeço primeiramente a Deus pela graça e misericórdia durante todas as viagens que me deslocava até a Universidade, livramentos e proteção para mim e minha família. Pelo amor e sabedoria concedidos a mim durante toda a minha trajetória na acadêmica e ainda na minha vida pessoal e profissional.

Agradeço aos meus pais, pelo seu esforço e dedicação em minha criação como ser-humano e indivíduo na sociedade. Pelos princípios e valores estabelecidos para que eu desempenhasse um papel de bom cidadão.

Agradeço aos meus irmãos pelo seu companheirismo e por sempre me ajudarem a trilhar o caminho educacional, pelo incentivo e apoio. Em especial ao meu irmão Joãozinho, que se dispôs de maneira grandiosa em momentos delicados da minha vida.

Agradeço à minha Orientadora Profa. Dr^a. Adriana Mara, por aceitar embarcar nessa trajetória que trilhamos juntos no meu TCC. Por seus direcionamentos e por estar empenhada em fazer com que os processos destravem, mostrando sempre outras abordagens que facilitam o ensino-aprendizagem.

Por fim, não menos importante, agradeço à minha esposa Patrícia e ao meu filho, Isaac, pelo amor, carinho e por me darem uma família abençoada por Deus. Através deles que busco sempre avançar nos meus estudos, para que assim, não apenas individualmente, crescamos em família, nas áreas financeiras e pessoal.

RESUMO

No atual cenário onde o uso correto dos dados para gerar informações úteis se faz presente, a análise descritiva dos dados se torna uma ferramenta poderosa para instituições de ensino, obtendo-se visões concretas e assertivas por meio dela sobre um contexto/problema. Nesse cenário, o trabalho aqui abordado consiste em realizar uma análise descritiva de dados educacionais sobre a evasão de graduandos dos cursos de Computação da Ufersa. Encontrar os cursos que se destacam nessa estatística negativa e em qual período houve maior expressividade/relevância. Utilizar a plataforma Google Colab para implementação dos códigos e a Linguagem de Programação Python, para geração de insights do problema em questão, chegando assim aos resultados esperados.

Palavras-chave: análise descritiva; evasão; dados educacionais; cursos; python.

ABSTRACT

In the current scenario, where the correct use of data to generate useful information is increasingly important, descriptive data analysis becomes a powerful tool for educational institutions, providing concrete and assertive insights into a given context or problem. In this context, the present work consists of performing a descriptive analysis of educational data related to student dropout in the Computer Science programs at Ufersa. The objective is to identify the programs that stand out in this negative statistic and determine the period in which dropout was most significant or relevant. The implementation of the code will be carried out using the Google Colab platform and the Python programming language, aiming to generate insights into the problem and thus reach the expected results.

Keywords: descriptive analysis; dropout; educational data; courses; python.

LISTA DE FIGURAS

Figura 1	– Evasão no Ensino Superior.....	17
Figura 2	– Taxas de permanência, conclusão acumulada e desistência acumulada.....	18
Figura 3	– Cursos Presenciais.....	18
Figura 4	– Cursos EaD.....	19
Figura 5	– Importando Bibliotecas Pandas e NumPy.....	22
Figura 6	– Carregamento da base de dados completa.....	22
Figura 7	– Base com as colunas excluídas.....	23
Figura 8	– Seleção dos cursos de computação a partir de 2015.....	23
Figura 9	– Valores de entrada do Dataframe antes da seleção dos cursos.....	24
Figura 10	– Informações do novo dataframe com a seleção dos cursos de Computação.....	24
Figura 11	– Separação dos cursos para análise individual.....	24
Figura 12	– Informações sobre os tipos dos dados no dataframe.....	25

LISTA DE GRÁFICOS

Gráfico 1	–	Sistemas de Informação (Campus Angicos).....	26
Gráfico 2	–	Computação e Informática (Campus Angicos).....	27
Gráfico 3	–	Ciências da Computação (Campus Mossoró).....	28
Gráfico 4	–	Engenharia da Computação (Campus Pau dos Ferros).....	28
Gráfico 5	–	Engenharia de Software (Campus Pau dos Ferros).....	29
Gráfico 6	–	Interdisciplinar em Tecnologia da Informação (Campus Pau dos Ferros).....	30
Gráfico 7	–	Sistemas de Informação (Campus Angicos).....	31
Gráfico 8	–	Computação e Informática (Campus Angicos).....	31
Gráfico 9	–	Ciências da Computação (Campus Mossoró).....	32
Gráfico 10	–	Engenharia da Computação (Campus Pau dos Ferros).....	32
Gráfico 11	–	Engenharia de Software (Campus Pau dos Ferros).....	33
Gráfico 12	–	Interdisciplinar em Tecnologia da Informação (Campus Pau dos Ferros).....	34

LISTA DE ABREVIATURAS E SIGLAS

IES	Instituições de Ensino Superior
EaD	Educação à Distância
KDD	Knowledge Discovery in Databases (Descoberta de Conhecimento em Bases de Dados)
BSI	Bacharelado em Sistemas de Informação
LCI	Licenciatura em Computação e Informática

LISTA DE SÍMBOLOS

% Porcentagem

SUMÁRIO

1 INTRODUÇÃO.....	14
1.1 JUSTIFICATIVA.....	14
1.2 OBJETIVOS.....	14
1.2.1 Objetivo Geral.....	15
1.2.2 Objetivos Específicos.....	15
2 REFERENCIAL TEÓRICO.....	15
2.1 TIPOS DE EVASÃO.....	16
2.1.1 Evasão por motivo de cancelamento.....	16
2.1.2 Evasão por motivo de trancamento de matrícula.....	16
2.1.3 Evasão por motivo de abandono escolar.....	17
2.2 TAXAS DE EVASÃO.....	17
2.2.1 Relação da evasão nos cursos de áreas tecnológicas.....	18
2.3 ESTRATÉGIAS DE COMBATE A EVASÃO.....	19
3 METODOLOGIA.....	20
4 PROCESSO DE ANÁLISE DESCRITIVA DOS DADOS 21.....	21
4.1 COLETA DOS DADOS.....	21
4.2 PRÉ-PROCESSAMENTO.....	23
4.3 TRANSFORMAÇÃO DOS DADOS.....	25
4.4 ANÁLISE DESCRITIVA.....	25
5 RESULTADOS E DISCUSSÃO.....	34
6 CONSIDERAÇÕES FINAIS.....	37
6.1 TRABALHOS FUTUROS.....	37
REFERÊNCIAS.....	38

1 INTRODUÇÃO

1.1 JUSTIFICATIVA

A justificativa para o desenvolvimento desta pesquisa volta-se para um assunto preocupante para a educação e mercado de trabalho, a evasão escolar, principalmente no ensino superior. Esse fenômeno tem grande relevância tanto nas áreas sociais quanto educacionais de nossa sociedade, pois traz consigo consequências negativas, tanto para as instituições de ensino, como também para as demandas do mercado de trabalho. Quando um aluno não conclui seu curso de nível superior, há um desperdício dos recursos e investimentos por parte do governo (com suas políticas públicas na educação), como também, o mercado de trabalho sofre com a escassez de profissionais qualificados para atuar em suas inúmeras áreas.

A evasão escolar no ensino superior tem impactos diretos na vida dos estudantes. Muitas vezes, os motivos que levam à evasão estão relacionados a dificuldades financeiras, problemas de saúde mental, falta de apoio acadêmico ou questões familiares. Abordar esse tema nos possibilita, com o uso das ferramentas e procedimentos corretos, entendermos a dimensão, através dos dados coletados, do que pode ser desenvolvido como estratégias educacionais para se conseguir reter o aluno até o término do seu curso.

O tema em questão é complexo e exige uma abordagem multifacetada e colaborativa. Os resultados gerados pelo abandono escolar evidenciam o aumento nas desigualdades sociais e econômicas, devido a pouca empregabilidade e oportunidades profissionais limitadas. Investir na prevenção e no combate a esse fenômeno não apenas beneficia os estudantes individualmente, mas também, contribui para o desenvolvimento social, econômico e cultural de uma nação na totalidade.

Analisar e comparar todos os cursos de Computação da UFERSA, destacando os cursos e períodos (anos) com maior incidência de alunos evadidos, auxilia na tomada de decisão, ajudando a elaborar assim, estratégias que visam a mudança na estatística negativa que a própria evasão escolar no nível superior traz. Através desses resultados, conseguimos, também, uma visão panorâmica do curso e campus com menor taxa de concluintes, assim buscando a melhor política para o melhoramento desses índices.

1.2 OBJETIVOS

As subseções a seguir contém os objetivos geral e específicos deste trabalho.

1.2.1 Objetivo Geral

Analisar e comparar os números referentes a quantidade de alunos evadidos nos cursos de Computação da Ufersa, nos últimos 5 anos, revelando assim os resultados durante a pandemia e identificando qual o curso com maior taxa de evasão e em qual período se deu essa realidade, de modo a entender de forma geral, onde se obteve as estatísticas negativas mais expressivas e com isso, buscar estratégias de contenção e melhoramento desses índices, usando para isso uma análise descritiva dos dados educacionais, coletados no site de dados aberto da Ufersa.

1.2.2 Objetivos Específicos

- Realizar pesquisa bibliográfica acerca do tema evasão escolar no nível superior.
- Coletar dados acerca das taxas de evasão nos cursos de Computação da Ufersa.
- Analisar os dados coletados, identificando quais os cursos e períodos onde apresenta os números mais expressivos para esse trabalho.
- Identificar através dos resultados os cursos onde a incidência de evasão se destaca.

2 REFERENCIAL TEÓRICO

Um tema com bastante relevância no meio acadêmico é o da evasão escolar, a definição usada por Baggi e Lopes (2011) é o da saída do aluno da instituição antes da conclusão do curso. Para ANDIFES et al. (1996), a evasão pode ser caracterizada como o ato do estudante se desligar do curso superior em situações como: o deixar de matricular-se, desistência oficial do curso e transferência ou reopção de curso.

Tal problema afeta de maneira diferente e significativa tanto a rede pública quanto privada de ensino. Segundo Lobo (2012), existe uma relação ruim entre as Instituições de Ensino Superior Brasileiras (IES) e a evasão, pois quando o aluno evade, existe o desperdício de recursos das universidades públicas, enquanto para as privadas é uma perda de receita por consequência da diminuição de mensalidades.

Neri (2009) argumenta que a educação constitui o verdadeiro custo de oportunidade para a sociedade: nenhuma nação cresce sem ela. Portanto, fazer um levantamento e trabalho

de informações, de maneira a encontrar os cursos da Ufersa que os alunos mais desistem ao longo da vida acadêmica, mostra a grande importância no assunto.

A evasão no Ensino Superior tem sido tema de pesquisas acadêmicas que, na sua maioria, buscam compreender as causas do aumento significativo do número de alunos que entram no Ensino Superior e evadem, muitas vezes após o primeiro semestre do curso (SOUZA, SILVA e GESSINGER, 2017). De acordo com Silva Filho (2007), é imprescindível a investigação das causas de evasão nas instituições de ensino superior.

2.1 TIPOS DE EVASÃO

Para CAROLINA (2023) existem três tipos de evasão acadêmica que ocorrem por razões diferentes, mas todas levam à interrupção dos estudos. Cada um deles pode ter impactos diferentes tanto para o aluno quanto para a instituição de ensino.

2.1.1 Evasão por motivo de cancelamento

Esse tipo de evasão ocorre quando o aluno cancela sua matrícula na instituição de ensino devido a razões específicas. As razões podem variar e incluem diversos motivos, como, por exemplo: questões financeiras, mudanças de planos, insatisfação com o curso ou com a instituição, etc.

2.1.2 Evasão por motivo de trancamento de matrícula

O trancamento de matrícula acontece quando um aluno solicita a interrupção temporária de seus estudos. Geralmente, é uma escolha pessoal do estudante e pode ser por motivos diversos, como problemas de saúde, questões familiares, necessidade de trabalhar, entre outros.

É importante lembrar que durante esse período, o aluno não participa das atividades escolares ou acadêmicas, mas sua vaga continua reservada na instituição. Portanto, considerando os 3 tipos de evasão de alunos, essa é a que a Instituição de Ensino possui a maior chance de ter o retorno do aluno.

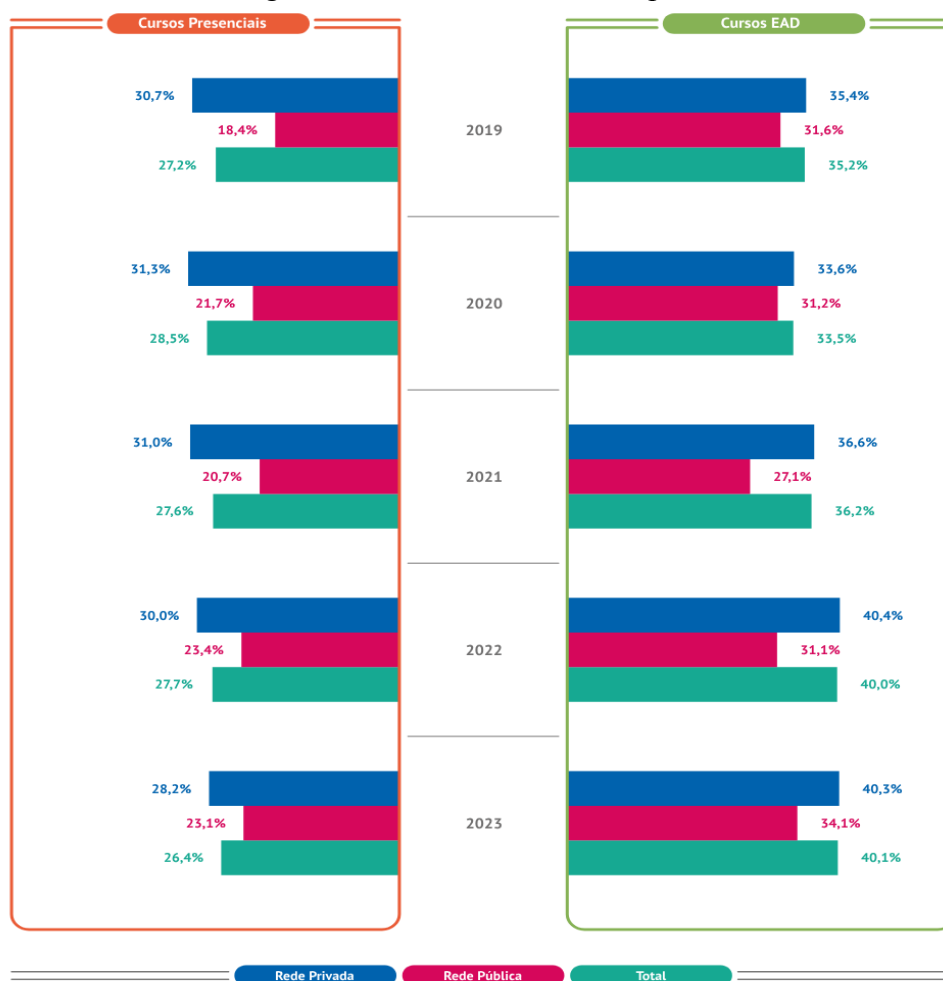
2.1.3 Evasão por motivo de abandono escolar

Este tipo de evasão ocorre quando o aluno deixa de frequentar as aulas e atividades sem formalizar o cancelamento da matrícula ou o trancamento. Dessa forma, o abandono escolar é caracterizado pela ausência prolongada do estudante sem justificativa, na instituição de ensino.

2.2 TAXAS DE EVASÃO

Segundo o Censo da Educação Superior, desenvolvido pelo Instituto SEMESP entre os anos de 2019 e 2023, os principais resultados encontrados sobre as taxas de evasão no nível superior são:

Figura 1 – Evasão no Ensino Superior.



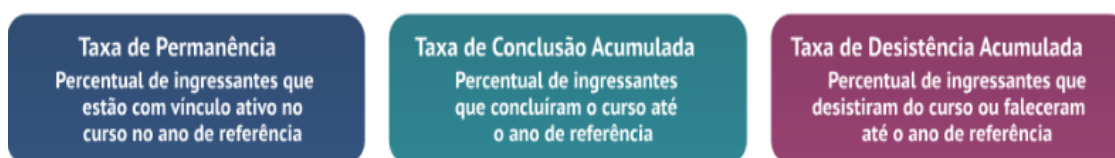
Fonte: Instituto Semesp | Base: INEP

Analisando a Figura 1, fica evidenciado uma maior taxa de evasão no formato de Educação à Distância (EaD) (se comparado a modalidade de ensino presencial), havendo uma certa progressão ano a ano. Já quando as taxas de evasão são do ponto de vista das instituições públicas e privadas, percebe-se que, há uma porcentagem maior de alunos evadidos em instituições privadas de ensino.

De forma geral, os dados apresentados nos mostram em destaque que alunos da rede privada de ensino, na modalidade EaD, estão desistindo ou abandonando seus cursos de nível superior, em contrapartida, alunos da rede pública, na modalidade presencial, são os que apresentam a menor taxa de evasão.

2.2.1 Relação da evasão nos cursos de áreas tecnológicas

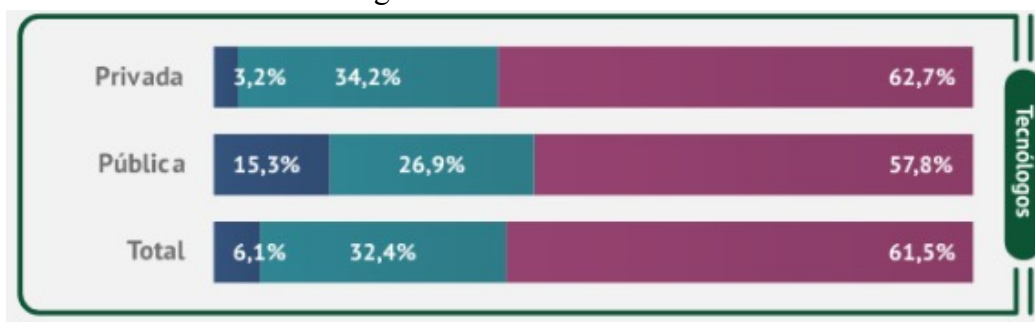
Figura 2 – Taxas de permanência, conclusão acumulada e desistência acumulada.



Fonte: Instituto Semesp | Base: INEP - Ciclo 2019-2023

Fazendo uso da Figura 2, se pode entender melhor o panorama apresentado (nas figuras a seguir) da relação entre os Cursos Presenciais e EaDs, entre a rede pública e privada de ensino em nosso país, panorama esse voltado para os cursos das áreas tecnológicas. As informações foram extraídas do Instituto Semesp, em pesquisa divulgada no “Mapa do Ensino Superior no Brasil” na sua 15ª edição.

Figura 3 – Cursos Presenciais



Fonte: Instituto Semesp | Base: INEP - Ciclo 2019-2023

Figura 4 – Cursos EaD



Fonte: Instituto Semesp | Base: INEP - Ciclo 2019-2023

Com base na relação apresentada nas figuras acima, vemos que os cursos voltados para as áreas tecnológicas na forma EaD, na rede privada, têm a maior taxa de evasão (desistência). Acompanhando esta mesma análise, os cursos na modalidade EaD, da rede pública, apresentam uma menor porcentagem de evasão dentre as relações.

2.3 ESTRATÉGIAS DE COMBATE A EVASÃO

Com o uso tecnológico de forma direcionada, é possível analisar dados existentes sobre os cursos de computação da Ufersa, para assim, extrairmos informações relevantes. Com um levantamento assim, se torna possível realizar uma previsão, sabendo qual curso será mais propenso a desistência, com isso buscando métodos e ações mais direcionadas a esse público específico.

As estratégias que poderão ser utilizadas vai depender de como os índices se apresentam. Essas estratégias podem englobar **apoio financeiro** para alunos de outras cidades, de modo a trazê-lo para o envolvimento do ambiente acadêmico, por meio de uma maior oferta de bolsas de estudo remuneradas para os Cursos de Computação e parcerias com empresas privadas em programas de trainee.

Outras estratégias que ajudam é o **acompanhamento acadêmico** mediante uso de sistemas de alerta precoce para identificar alunos com risco de evasão. Outro meio são os **currículos mais conectados ao mercado**, através do desenvolvimento de soft skills e orientação de carreira, aumentando a motivação e percepção de utilidade, como também, melhorando a relação com o mercado de trabalho.

Por último, **pesquisas de satisfação e engajamento** que podem identificar padrões por curso, semestre, faixa etária, etc. e entender que afasta os alunos, assim ajudando no monitoramento e análise de dados.

3 METODOLOGIA

Inicialmente, para se chegar aos resultados esperados deste trabalho, será realizada a revisão de literatura. Após essa etapa, serão selecionadas as informações relevantes sobre o tema evasão escolar no nível superior, a partir de critérios como a pesquisa em artigos que exploram o tema direta ou indiretamente no período entre 2019 e 2024, e a confiabilidade das plataformas (sites) como o Instituto Semesp, o gov.br e a plataforma de dados aberto da Ufersa, onde se consegue de maneira clara e objetiva as informações que evidenciam e alimentam esse trabalho, gerando uma visão da macro para micro do contexto a ser explorado.

Esta pesquisa será desenvolvida de forma descritiva, onde os dados serão tratados de maneira quantitativa, a fim de se entender a dimensão sobre o tema em questão. A revisão de literatura visa identificar, por meio de artigos científicos, o que os pesquisadores já desenvolveram de relevante sobre o assunto evasão escolar no nível superior. Com isso, evidenciando os fatores que se destacam para o entendimento desse fenômeno em sua totalidade.

Em seguida, trazer dados estatísticos das plataformas disponíveis sobre o tema, como o **gov.br** através do Censo da Educação Superior, **Semesp** com o 15º Mapa do Ensino Superior, tanto nas áreas de ensino público, como privado, e também o método de educação presencial e a distância. Por último, fazendo uso da **Plataforma de Dados Abertos da Ufersa** com os quantitativos de alunos dos cursos de graduação, assim comparando quais desses itens em questão têm um maior destaque e relevância.

Logo após esse processo, será feita a coleta de dados, de maneira focalizada na pesquisa em questão. Dados esses dos cursos de Computação da Ufersa, dos últimos cinco (5) anos. Informações para se entender a quantidade de alunos evadidos e retidos (que não concluem o curso).

Por último será implementado as etapas que cabem ao processo de análise descritiva dos dados, tais como: **coleta dos dados**, que consiste em identificar e extrair os dados mais relevantes de uma ou mais fontes, focando apenas naquilo que é útil para a análise. Essa etapa garante que o processo comece com informações adequadas ao problema a ser resolvido.

Logo em seguida é implementada a etapa de **pré-processamento**, onde envolve a limpeza e preparação dos dados, tratando valores ausentes, inconsistências e ruídos. É fundamental para garantir a qualidade dos dados e evitar distorções nos resultados. A terceira etapa é a de **transformação dos dados**, onde os dados são convertidos para formatos apropriados para a análise, o que pode incluir normalização, agregações ou criação de novos atributos. Essa etapa facilita a identificação de padrões durante a análise posterior. Por fim, a última etapa implementada é a de **análise descritiva**, onde se aplicam técnicas e algoritmos para extração das informações pertinentes ao objetivo da análise em si (onde ocorre a extração efetiva de conhecimento). Será apresentado no tópico de **resultado e discussão** a etapa de **Interpretação e Avaliação**, etapa que é fruto das informações extraídas através da análise anterior, onde padrões encontrados são avaliados quanto à sua validade, utilidade e clareza para o usuário final. Essa análise garante que os resultados façam sentido e possam ser confiavelmente utilizados.

Fazendo uso das ferramentas Google Colab para implementação dos códigos e a Linguagem de Programação Python para o desenvolvimento serão extraídas as informações mais relevantes desses dados que são a quantidade de matrículas **ativas, trancadas, canceladas e concluídas**, para se entender o problema de forma mais direta e clara, assim gerando insights, e com isso, ajudando na tomada de decisão, visando o desenvolvendo de planos e estratégias para inverter esse quadro da evasão escolar nos cursos estudados.

4 PROCESSO DE ANÁLISE DESCRITIVA DOS DADOS

4.1 COLETA DOS DADOS

A seleção (coleta) dos dados para análise foi a partir da Base de Dados Abertos da Ufersa através do link <https://dadosabertos.ufersa.edu.br/dataset/quantitativos-de-alunos-graduacao/resource/c8112a24-d88e-4a4b-83c6-24f05d4ac6e0>, de onde foi carregada a base completa referente a todos os cursos presenciais ofertados pela própria instituição. Foi carregado inicialmente, como mostra a Figura 5, as bibliotecas NumPy e Pandas, a primeira para se trabalhar com funções matemáticas encontradas na base e a segunda para manipular dados em tabelas (Dataframes). Para trabalhar a base de dados coletada, foi utilizada a plata-

forma Google Colab, sendo escolhida essa por ser gratuita, fácil de usar, não exigir instalação (podendo rodar em qualquer máquina por ser uma plataforma online), facilidade em compartilhar os códigos desenvolvidos e por ser integrada ao Google Drive, o que facilita salvar e acessar projetos e a linguagem de programação Python, ferramentas estas, de suma importância durante todas as etapas.

Figura 5 – Importando Bibliotecas Pandas e NumPy



```

▼ Importando bibliotecas PANDAS e NUMPY

[1] 1 import pandas as pd
     2 import numpy as np

[2] 1 base = pd.read_csv('/content/drive/MyDrive/TCC 2/Arquivos .csv/estatisticasCursosGraduacaoCSV.csv')
  
```

Fonte: Autoria própria

Na Figura 5, há a atribuição de toda a base selecionada para um Dataframe chamado “Base”, conforme visto no bloco 2 de codificação. Logo após essa atribuição, a base é carregada, como exibido na Figura 6, mostrando todos os cursos (sem distinção específica de área). Na figura percebemos que é apresentado apenas os primeiros e os últimos cursos (para bases extensas), por padrão da plataforma Google Colab.

Figura 6 – Carregamento da base de dados completa

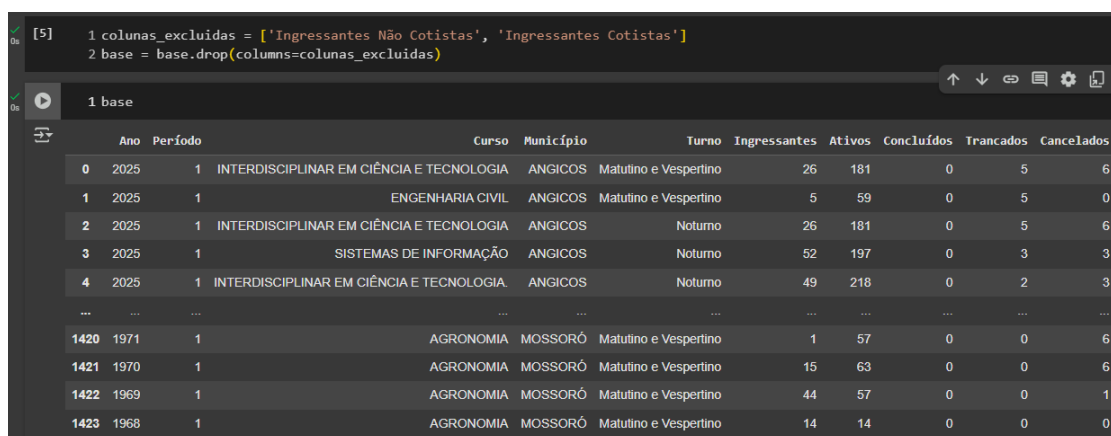
	Ano	Período	Curso	Município	Turno	Ingressantes	Ingressantes Não Cotistas	Ingressantes Cotistas	Ativos	Concluídos	Trancados	Cancelados
0	2025	1	INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA	ANGICOS	Matutino e Vespertino	26	26	0	181	0	5	6
1	2025	1	ENGENHARIA CIVIL	ANGICOS	Matutino e Vespertino	5	5	0	59	0	5	0
2	2025	1	INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA	ANGICOS	Noturno	26	26	0	181	0	5	6
3	2025	1	SISTEMAS DE INFORMAÇÃO	ANGICOS	Noturno	52	29	23	197	0	3	3
4	2025	1	INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA	ANGICOS	Noturno	49	25	24	218	0	2	3
...	...	...	...	...	...	...	...	...	...	...	...	...
1420	1971	1	AGRONOMIA	MOSSORÓ	Matutino e Vespertino	1	1	0	57	0	0	6
1421	1970	1	AGRONOMIA	MOSSORÓ	Matutino e Vespertino	15	15	0	63	0	0	6

Fonte: Autoria própria

4.2 PRÉ-PROCESSAMENTO

Na etapa de Pré-Processamento, houve limpeza dos dados irrelevantes, contidos nas colunas “Ingressantes Não Cotistas” e “Ingressantes Cotistas”, dados que já se resumiam na coluna “Ingressantes”, etapa essa que não gerou dano algum aos processos seguintes.

Figura 7 – Base com as colunas excluídas



```
[5] 1 colunas_excluidas = ['Ingressantes Não Cotistas', 'Ingressantes Cotistas']
     2 base = base.drop(columns=colunas_excluidas)
```

1 base

	Ano	Período	Curso	Município	Turno	Ingressantes	Ativos	Concluídos	Trancados	Cancelados
0	2025	1	INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA	ANGICOS	Matutino e Vespertino	26	181	0	5	6
1	2025	1	ENGENHARIA CIVIL	ANGICOS	Matutino e Vespertino	5	59	0	5	0
2	2025	1	INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA	ANGICOS	Noturno	26	181	0	5	6
3	2025	1	SISTEMAS DE INFORMAÇÃO	ANGICOS	Noturno	52	197	0	3	3
4	2025	1	INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA	ANGICOS	Noturno	49	218	0	2	3
...	...	...	...	...	...	...	...	...	...	...
1420	1971	1	AGRONOMIA	MOSSORÓ	Matutino e Vespertino	1	57	0	0	6
1421	1970	1	AGRONOMIA	MOSSORÓ	Matutino e Vespertino	15	63	0	0	6
1422	1969	1	AGRONOMIA	MOSSORÓ	Matutino e Vespertino	44	57	0	0	1
1423	1968	1	AGRONOMIA	MOSSORÓ	Matutino e Vespertino	14	14	0	0	0

Fonte: Autoria própria

Seguindo ainda nessa etapa, foi criado um novo dataframe (tabela) com o nome “baseAtualizada” (Figura 8) onde foi atribuído um refinamento, para separar os cursos de computação dos demais. A partir da função “.isin”, os dados são selecionados de uma coluna específica, no caso, a coluna “curso” foi usada para essa seleção, juntamente com a coluna “ano” selecionando os anos a partir de 2015, como ano mínimo para se entender como os números são apresentados nos últimos 10 (dez) anos e ver as variações ao longo do tempo, entre o período Pré-Pandemia, como também, durante a Covid-19 com as mudanças nos formatos de ensino, passando a serem desenvolvidos remotamente e o período Pós-Pandemia quando os alunos voltaram de forma presencial a Universidade.

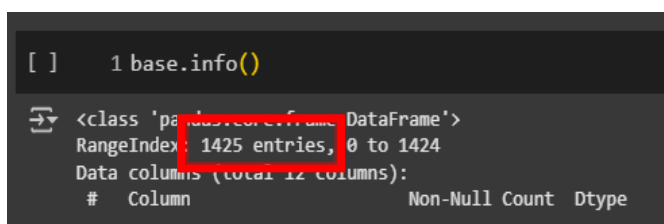
Figura 8 – Seleção dos cursos de computação a partir de 2015

```
1 baseAtualizada = base[
2     (base['Curso'].isin([
3         'COMPUTAÇÃO E INFORMÁTICA',
4         'SISTEMAS DE INFORMAÇÃO',
5         'CIÊNCIA DA COMPUTAÇÃO',
6         'ENGENHARIA DE COMPUTAÇÃO',
7         'ENGENHARIA DE SOFTWARE',
8         'INTERDISCIPLINAR EM TECNOLOGIA DA INFORMAÇÃO'
9     ])) & (base['Ano'] >= 2015) # Selecionando apenas com início do curso acima de 2015.
10 ]
```

Fonte: Autoria própria

Nas imagens abaixo, podemos comparar o antes e depois da seleção dos dados para o novo dataframe, mostrando a quantidade de registro entre eles, com os cursos selecionados.

Figura 9 – Informações dos valores de entrada do Dataframe antes da seleção dos cursos



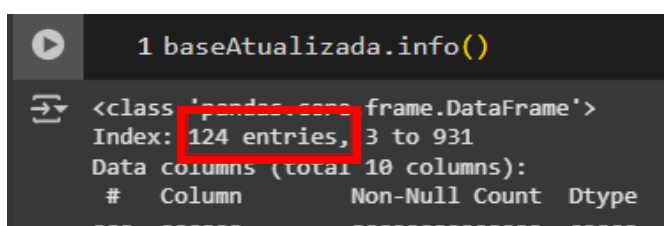
```
[ ] 1 base.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1425 entries, 0 to 1424
Data columns (total 12 columns):
#   Column                Non-Null Count  Dtype
---

```

Fonte: Autoria própria

Figura 10 – Informações do novo dataframe com a seleção dos cursos de Computação



```
1 baseAtualizada.info()

<class 'pandas.core.frame.DataFrame'>
Index: 124 entries, 3 to 931
Data columns (total 10 columns):
#   Column                Non-Null Count  Dtype
---

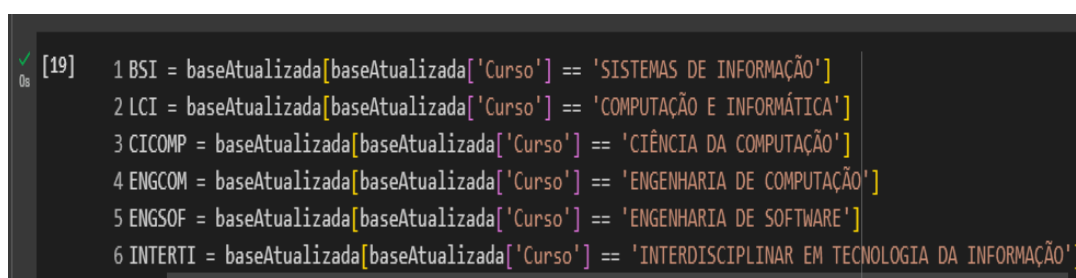
```

Fonte: Autoria própria

Nota-se pela seleção em cor vermelha (em cada imagem) a diferença de valores entre a base completa com 1.425 entradas (linhas) e a selecionada apenas com os dados pertinentes com 124 entradas, preparando dessa forma, a separação por curso para análise individual dos mesmos.

Foram criados vários dataframes, de modo a separar todos os cursos de computação da Ufersa, para análise individual, melhorando a visualização e extração de informações. A figura 11 mostra a nova nomenclatura que cada dataframe recebeu e a linha de código utilizada para esta seleção.

Figura 11 – Separação dos cursos para análise individual



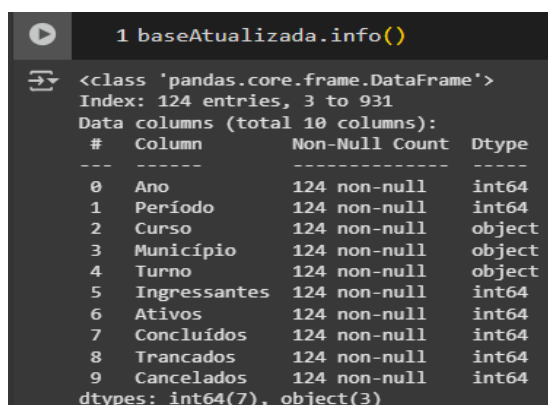
```
[19] 1 BSI = baseAtualizada[baseAtualizada['Curso'] == 'SISTEMAS DE INFORMAÇÃO']
      2 LCI = baseAtualizada[baseAtualizada['Curso'] == 'COMPUTAÇÃO E INFORMÁTICA']
      3 CICOMP = baseAtualizada[baseAtualizada['Curso'] == 'CIÊNCIA DA COMPUTAÇÃO']
      4 ENGCOM = baseAtualizada[baseAtualizada['Curso'] == 'ENGENHARIA DE COMPUTAÇÃO']
      5 ENGSOFT = baseAtualizada[baseAtualizada['Curso'] == 'ENGENHARIA DE SOFTWARE']
      6 INTERTI = baseAtualizada[baseAtualizada['Curso'] == 'INTERDISCIPLINAR EM TECNOLOGIA DA INFORMAÇÃO']
```

Fonte: Autoria própria

4.3 TRANSFORMAÇÃO DOS DADOS

Não houve necessidade de utilização da etapa de transformação dos dados, visto que estão no formato ideal para o objetivo proposto no projeto. A partir da função “.info()” aplicada ao dataframe “baseAtualizada” é gerada a imagem abaixo, nela vemos de forma detalhada as informações, que asseguram a manipulação dos dados em relação aos resultados que se esperam. Dados do tipo inteiro (numéricos), expressando as quantidades de alunos ativos, trancados, cancelados e concluídos e o ano e período de cada curso de computação, como também, o nome do curso e município do mesmo, essas duas últimas no formato de texto (object).

Figura 12 – Informações sobre os tipos dos dados no dataframe



```
1 baseAtualizada.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 124 entries, 3 to 931
Data columns (total 10 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Ano              124 non-null    int64
1   Período          124 non-null    int64
2   Curso            124 non-null    object
3   Município        124 non-null    object
4   Turno            124 non-null    object
5   Ingressantes     124 non-null    int64
6   Ativos           124 non-null    int64
7   Concluídos       124 non-null    int64
8   Trancados        124 non-null    int64
9   Cancelados       124 non-null    int64
dtypes: int64(7), object(3)
```

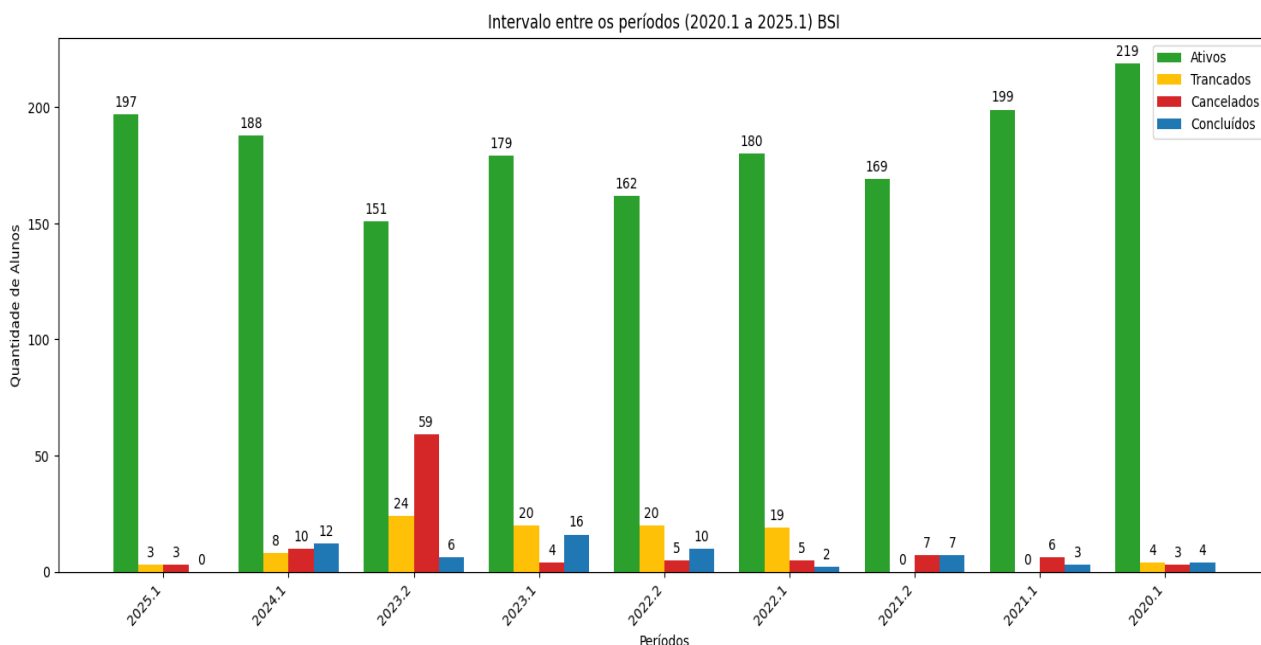
Fonte: Autoria própria

4.4 ANÁLISE DESCRITIVA

Na parte de análise descritiva dos dados, foi aplicado um algoritmo em Python, devido a plataforma ser projetada principalmente para rodar os códigos nesta linguagem, especialmente aqueles relacionados à **análise de dados**, por esse motivo é que foi utilizado essa linguagem. O algoritmo foi usado para a criação de novos dataframes separando os cursos de computação, gerando gráficos como resultados, facilitando assim, a descoberta de padrões e geração de insights. Nos gráficos de 1 a 6 vemos a relação entre os alunos ativos, trancados, cancelados e concluídos por período, nos últimos 5 anos. Foi utilizado o recorte dos últimos 5 anos devido a melhor visualização dos gráficos em barras para esse documento,

visto que, quando se implementou a seleção dos dez últimos anos, ficava ruim a leitura e visualização nítida das informações pertinentes.

Gráfico 1 – Sistemas de Informação (Campus Angicos)



Fonte: Autoria própria

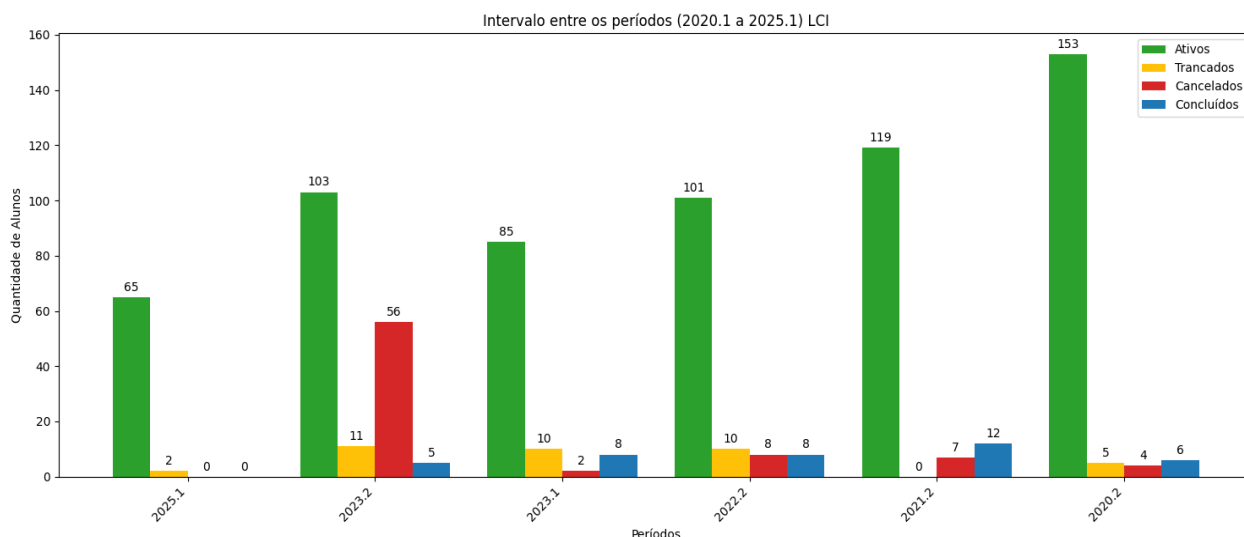
O gráfico acima apresenta a distribuição dos alunos do curso de **Bacharelado em Sistemas de Informação (BSI)** da Ufersa Campus Angicos, entre os períodos de 2020.1 a 2025.1, classificados nas categorias da legenda na parte interna do gráfico. Observa-se que, ao longo dos períodos, a maioria dos alunos permanece com matrículas ativas, com destaque para o período 2025.1, que apresenta o maior número de alunos ativos (197) em se falando de um período pós-pandemia.

O período com maior número de cancelamentos foi 2023.2, com 59 alunos **cancelados**, representando um pico atípico em comparação aos demais períodos, o mesmo também, com a menor quantidade de alunos **ativos**. A quantidade de alunos **trancados** e **concluídos** varia moderadamente, sendo mais expressiva nos períodos de 2022.1 a 2024.1. Nota-se, ainda, uma tendência de crescimento no número total de alunos ativos nos períodos mais recentes, indicando possível expansão ou maior retenção no curso.

O Gráfico 2 mostra a distribuição de alunos do curso de **Licenciatura em Computação e Informática (LCI)** do mesmo Campus Angicos, usando recorte de tempo/período do gráfico anterior. Percebe-se uma queda significativa no número de alunos

ativos no período mais recente (2025.1), com apenas 65 alunos, contrastando com o pico de 153 **ativos** em 2020.2. Assim como no curso de BSI, o período 2023.2 também registra um número elevado de **cancelamentos**, 56, o que sugere um padrão crítico naquele semestre para ambos os cursos. O curso de BSI apresenta aumento no número de ativos nos últimos períodos, o de LCI demonstra uma tendência inversa, com queda acentuada em 2025.1.

Gráfico 2 – Computação e Informática (Campus Angicos)

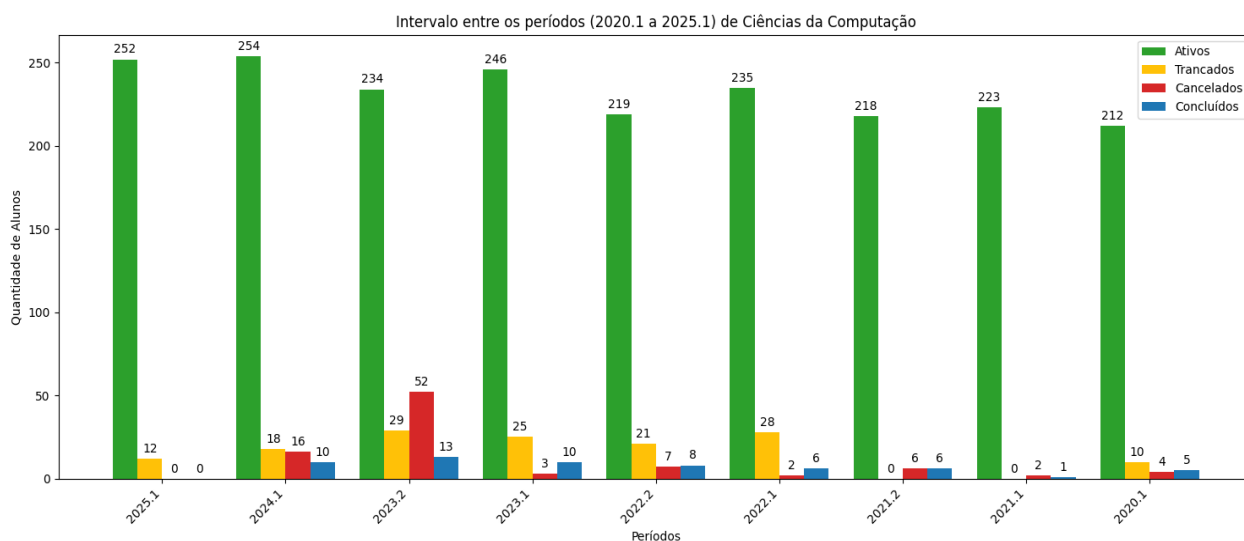


Fonte: Autoria própria

Para o Gráfico 3 (abaixo) nos é apresentado a evolução da quantidade de alunos do curso de **Ciência da Computação** da Ufersa Campus Central (Mossoró). O curso mencionado mantém um número elevado e estável de alunos ativos ao longo de todos os períodos, oscilando entre 212 e 254 alunos. Assim como observado nos cursos de **BSI** e **LCI** (gráficos anteriores, ambos do Campus Angicos), o período 2023.2 também representa um ponto crítico, com 52 **cancelamentos**, sendo o mais alto da série, o que reforça um padrão institucional ou externo de impacto naquele semestre.

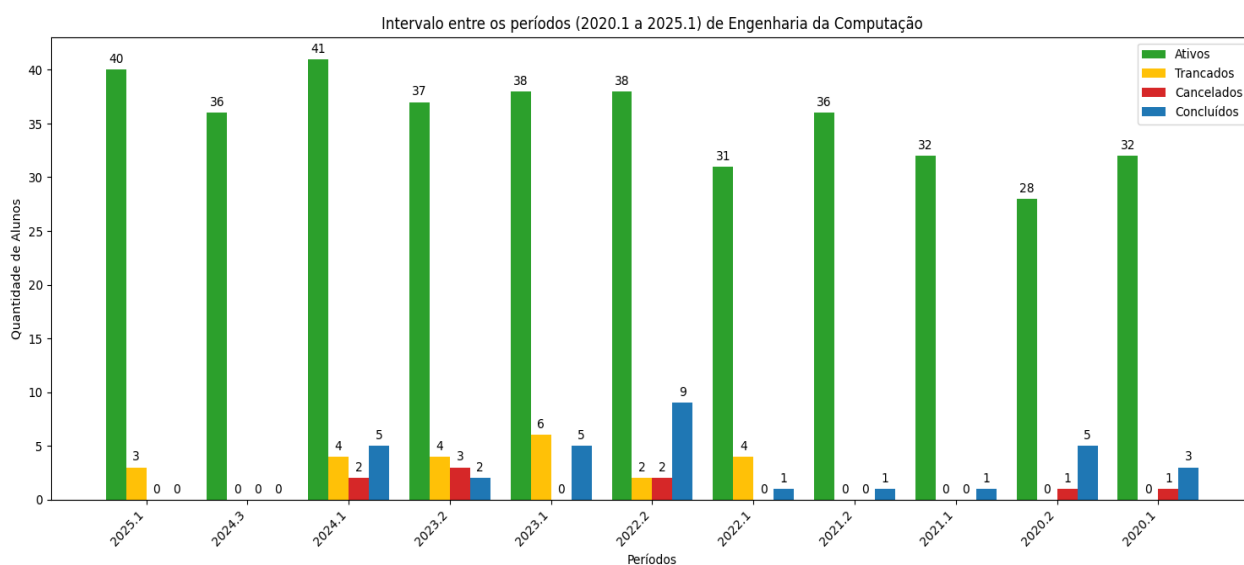
Os **trancamentos** têm valores oscilantes, mas, em geral, estão entre 10 e 29 alunos por semestre, sendo 2023.2 e 2022.1 os semestres com mais trancamentos (29 e 28, respectivamente). O número de alunos **concluintes** por semestre é baixo, não ultrapassando 13 alunos, o que pode indicar um tempo de formação superior ao previsto ou evasão antes da conclusão. A análise isolada demonstra a força e estabilidade do curso, mas também levanta pontos de atenção quanto à baixa taxa de conclusão por período.

Gráfico 3 – Ciências da Computação (Campus Mossoró)



Fonte: Autoria própria

Gráfico 4 – Engenharia da Computação (Campus Pau dos Ferros)

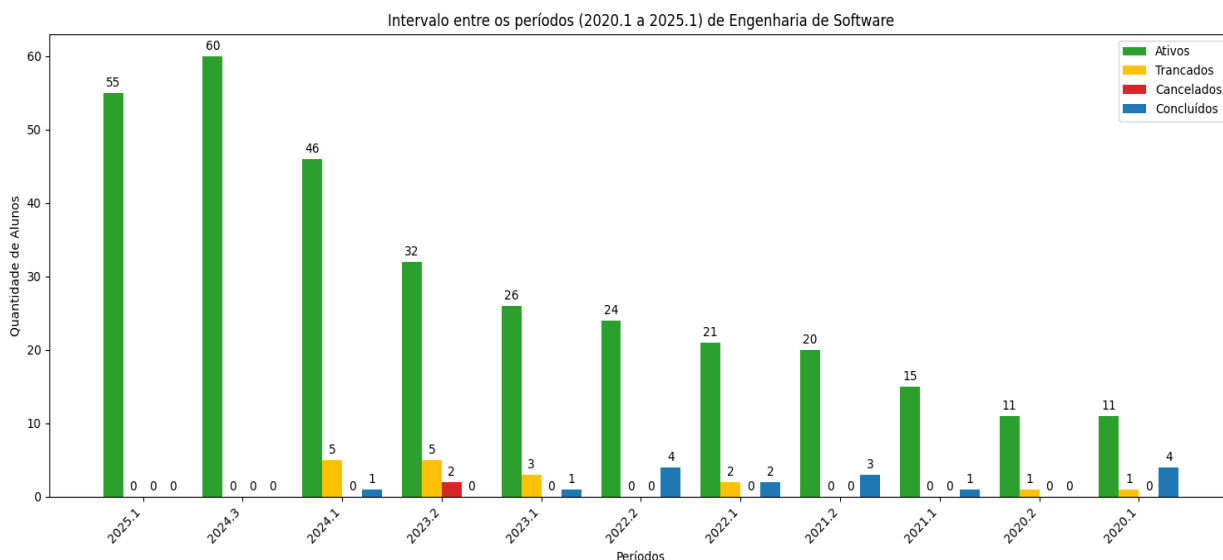


Fonte: Autoria própria

O Gráfico 4 nos exhibe os números de alunos (conforme legenda) do curso de **Engenharia da Computação** do Campus Pau dos Ferros, ao longo de 11 períodos, entre 2020.1 a 2025.1. Predominantemente, a categoria "Ativos" se mantém em alta, com a maioria dos períodos registrando mais de 30 alunos nessa situação, atingindo picos de 40 e 41 alunos em 2025.1 e 2024.1, respectivamente. As ocorrências de matrículas **trancadas** e **canceladas** são esporádicas e de baixo volume, sendo que em muitos períodos, essas categorias registram

zero alunos. Vale salientar que existem mais períodos com alunos **concluídos** do que trancados/cancelados.

Gráfico 5 – Engenharia de Software (Campus Pau dos Ferros)



Fonte: Autoria própria

O Gráfico 5 mostra a evolução do número de alunos no curso de **Engenharia de Software**, situado no mesmo campus do gráfico 4 e abrangendo a mesma faixa de períodos. Em contraste com o gráfico de Engenharia da Computação, que apresentava um número de alunos **ativos** mais estável e ligeiramente inferior, a Engenharia de Software demonstra um crescimento notável nos alunos "**Ativos**" nos períodos mais recentes, atingindo picos de **60** em 2024.3 e **55** em 2025.1.

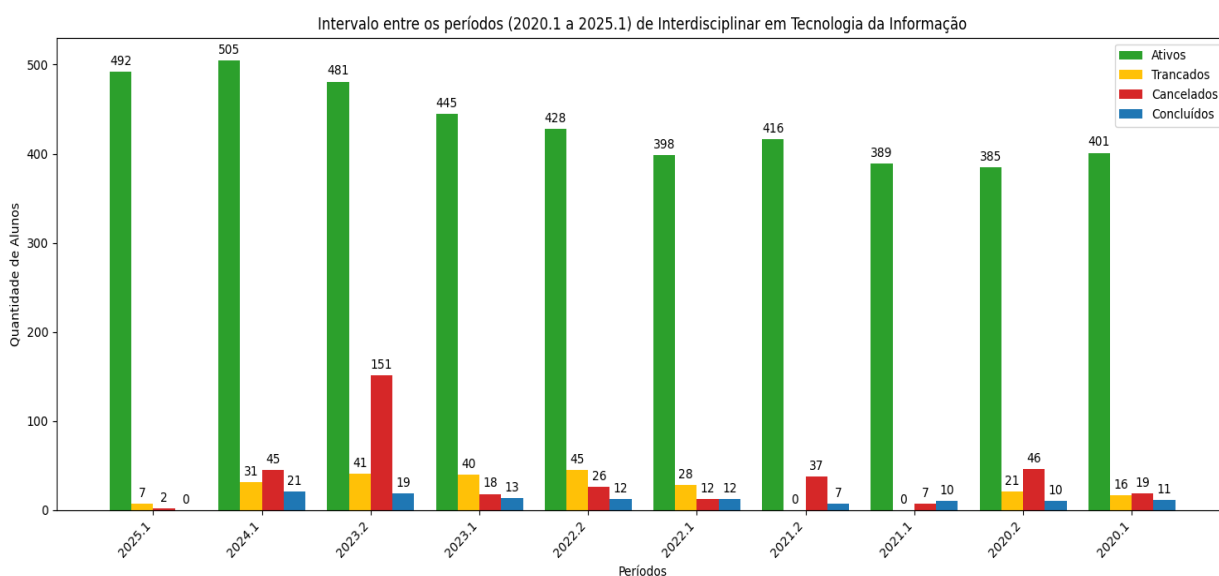
Observa-se uma tendência de aumento no número de alunos **ativos** ao longo do tempo, partindo de **11 alunos em 2020.1** e chegando aos mais de **50**. As categorias "Trancados", "Cancelados" e "Concluídos" permanecem com volumes baixos, semelhante ao curso de Engenharia da Computação. No entanto, em **Engenharia de Software**, há mais ocorrências de alunos "**Concluídos**" nos períodos mais antigos, como 4 em 2020.1, e em 2024.1, 5 alunos foram "**Trancados**". Essa diferença sugere que, embora ambos os cursos mantenham uma proporção alta de alunos ativos em relação às outras categorias, Engenharia de Software tem experimentado um crescimento mais expressivo na entrada de novos alunos nos últimos anos, enquanto Engenharia da Computação manteve um fluxo mais constante.

Por fim, o Gráfico 6 (ainda no Campus Pau dos Ferros) nos revela as quantidades de alunos do curso de **Interdisciplinar em Tecnologia da Informação**, abrangendo o período de

2020.1 a 2025.1. Diferentemente dos cursos de Engenharia da Computação e Engenharia de Software, este curso demonstra um volume significativamente maior de alunos em todas as categorias.

A quantidade de alunos **ativos** é alta em relação aos gráficos anteriores, superando 380 em todos os períodos e atingindo picos de 505 em 2024.1 e 492 em 2025.1. As demais categorias, embora em menor número que os **ativos**, são mais expressivas do que nos outros gráficos. Frisa-se mais uma vez o período 2023.2, com sua elevada quantidade de matrículas **canceladas** sendo 151 e o período 2022.2 com 45 **trancamentos**. Em resumo, o curso de Interdisciplinar em Tecnologia da Informação se destaca pela sua grande base de alunos ativos e por um número mais substancial de alunos trancados, cancelados e concluídos, indicando uma dinâmica de fluxo de alunos mais movimentada em comparação com os cursos de Engenharia da Computação e Engenharia de Software.

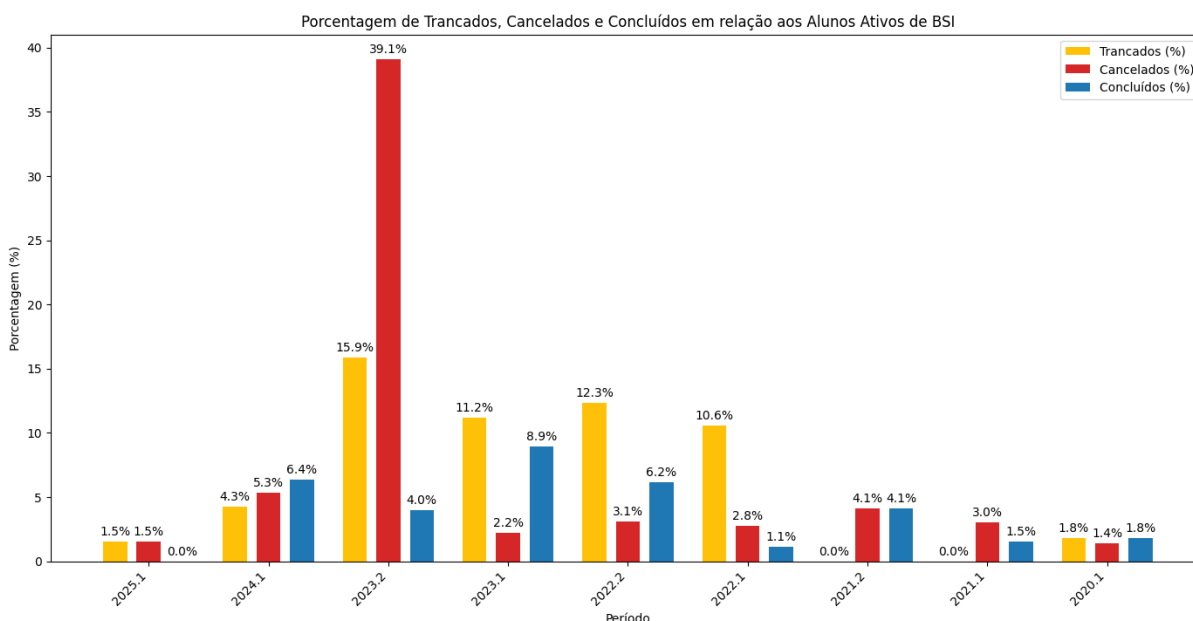
Gráfico 6 – Interdisciplinar em Tecnologia da Informação (Campus Pau dos Ferros)



Fonte: Autoria própria

Nos gráficos seguintes, do 7 ao 12, é mostrado, também por curso, a relação entre alunos trancados, cancelados e concluídos, usando como fator principal os alunos ativos do mesmo período. Essa relação é expressa em porcentagem, exemplo: 22.2% dos alunos ativos cancelaram a matrícula no período 2023.2 no curso de Ciências da Computação.

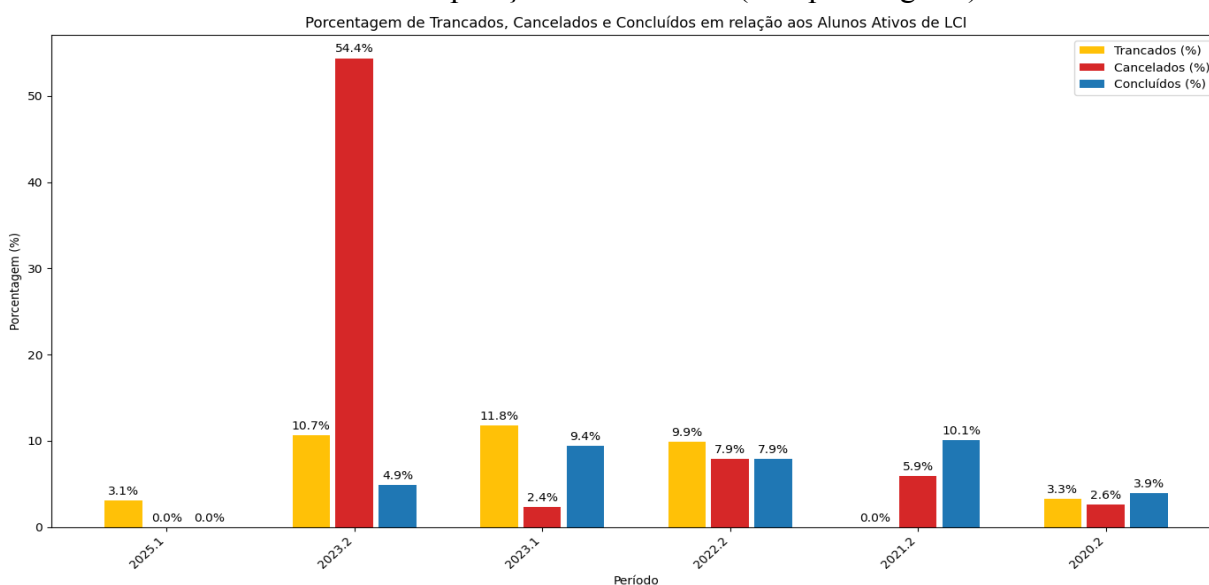
Gráfico 7 – Sistemas de Informação (Campus Angicos)



Fonte: Autoria própria

Os Gráficos 7 e 8 fornecem informações sobre a porcentagem nos cursos de **BSI** e **LCI** dos alunos trancados, cancelados e concluídos em relação aos alunos ativos. Para ambos os cursos, o **período 2023.2** como antes exposto, prova o pico alarmante para matrículas canceladas e número elevado para trancadas. A taxa de cancelamentos em BSI é de 39.1% e em LCI fica com impressionantes 54.4%, já a quantidade de trancamentos gira em torno de 15.9% e 11.8% para BSI e LCI respectivamente. A taxa de trancamentos de 11.8% em LCI já é destaque no período anterior 2023.1. As taxas de alunos **concluídos** em **BSI** e **LCI** com maior destaque são no período 2023.1 com 8.9% para o primeiro e em 2021.2 com 10.1% para o segundo curso.

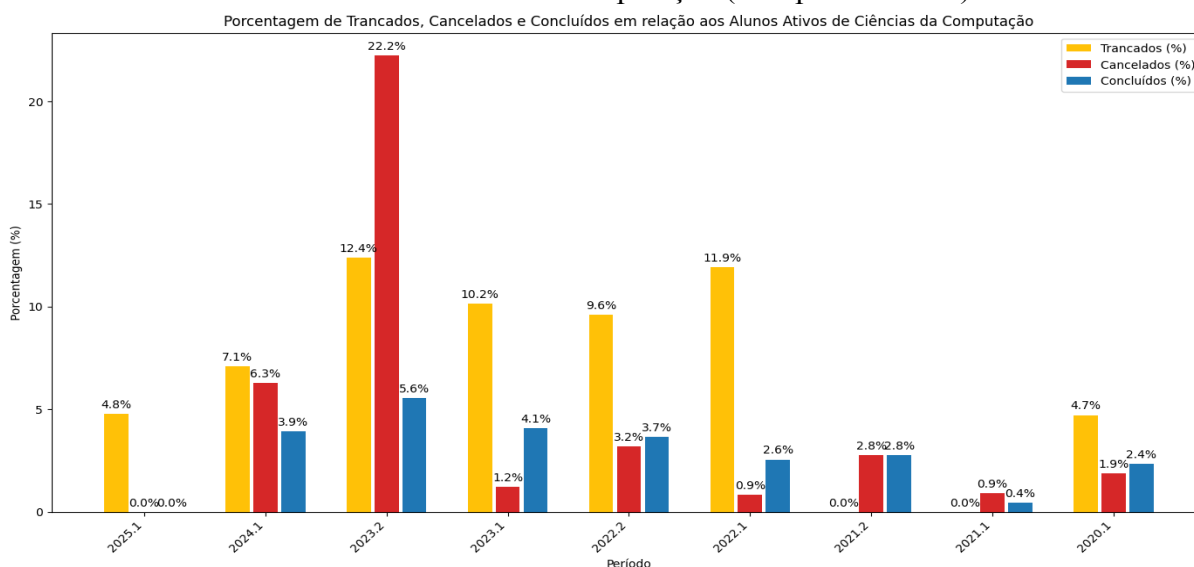
Gráfico 8 – Computação e Informática (Campus Angicos)



Fonte: Autoria própria

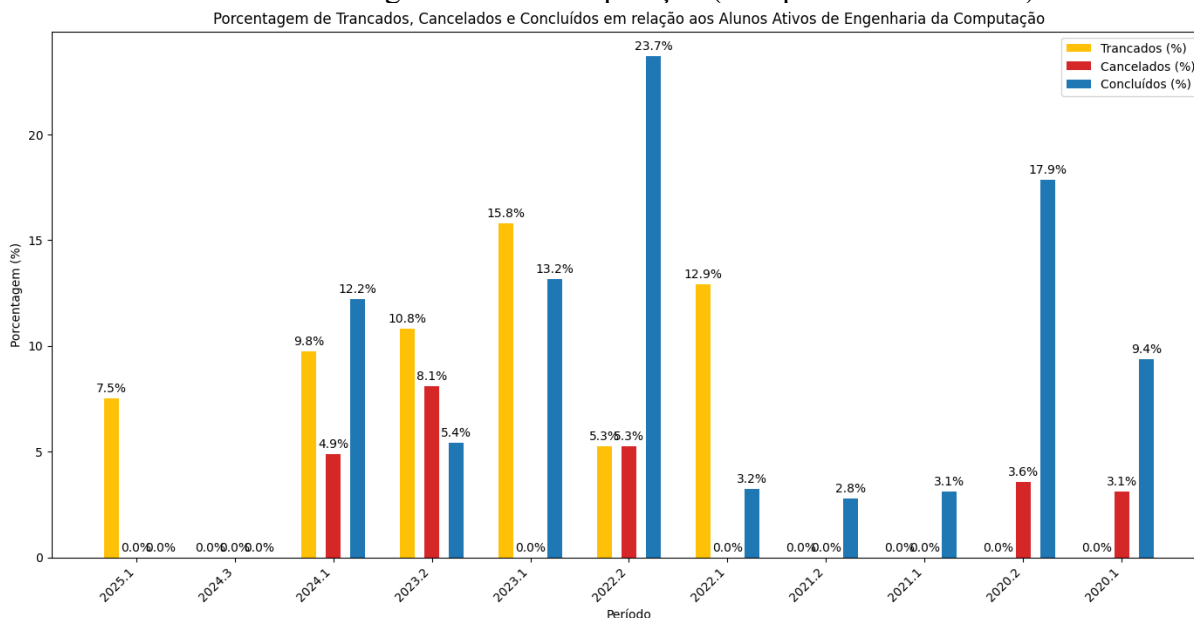
Com base no Gráfico 9 (abaixo), seguindo a mesma linha de análise que os anteriores, agora voltado para o curso de **Ciências da Computação** do Campus Mossoró, verifica-se novamente o destaque expresso por 22.2% de **cancelamentos e trancamentos** de 12.4% no período **2023.2**, indicando que uma parcela bem significativa dos alunos ali que estavam ativos, optaram por sair ou pausar o curso. Há uma flutuação considerável entre os períodos **2022.1** e **2023.2** em se falando das matrículas **trancadas**. Em contrapartida os **cancelamentos** são geralmente baixos, fora do pico ocorrido, sendo na maioria das vezes, abaixo de 5% e chegam a 0.0% em 2025.1.

Gráfico 9 – Ciências da Computação (Campus Mossoró)



Fonte: Autoria própria

Gráfico 10 – Engenharia da Computação (Campus Pau dos Ferros)

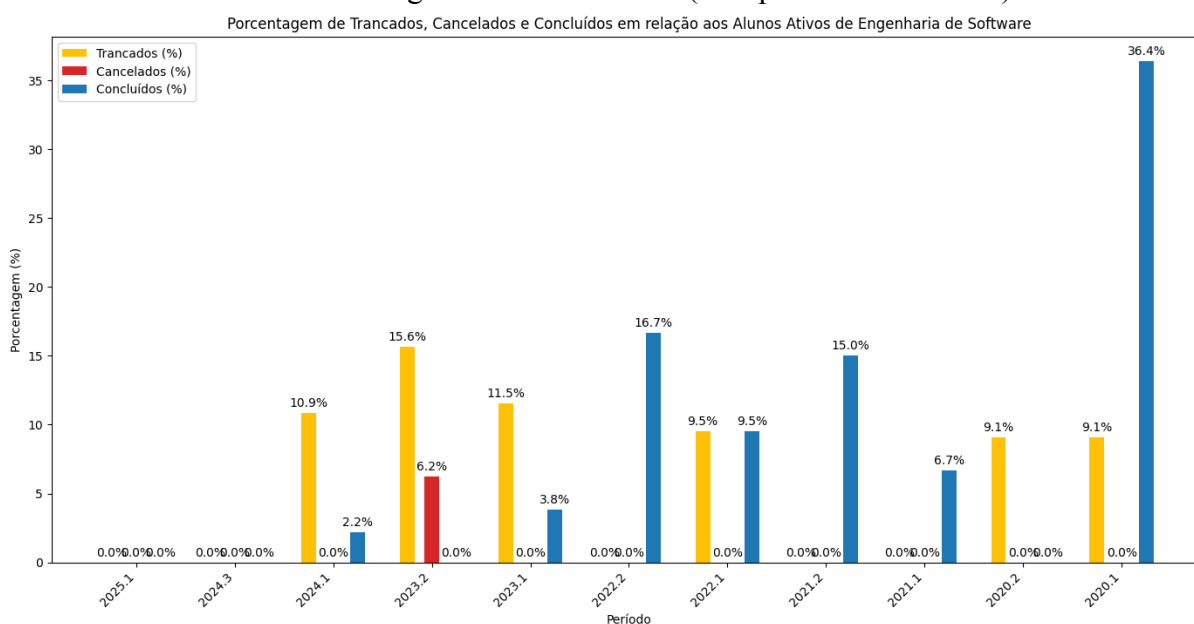


Fonte: Autoria própria

O Gráfico 10 demonstra que o curso de Engenharia da Computação tem uma alta porcentagem de alunos **concluídos**. O pico mais notável é em **2022.2, com 23.7%**, e outro ponto alto em **2020.2, com 17.9%**. Nas barras referente a trancamentos percebe-se o início dessa estatística a partir do período 2022.1 (período ainda em pandemia). Na maioria dos períodos, as taxas de **cancelamentos** ficam em 0.0% ou abaixo de 5%. No entanto, há um pico em **2023.2, com 8.1%**, que, embora não seja tão alto quanto os picos de trancamento ou conclusão, é o mais significativo para cancelamentos no gráfico.

Em geral, a porcentagem de matrículas trancadas tende a ser mais alta e mais frequente do que a de canceladas, sugerindo que os alunos do curso de **Engenharia da Computação** tendem mais a pausar os estudos do que a abandoná-los definitivamente. Embora haja uma flutuação considerável nos trancamentos, os cancelamentos são, em sua maioria, baixos ou inexistentes, com exceção de um pico em 2023.2.

Gráfico 11 – Engenharia de Software (Campus Pau dos Ferros)

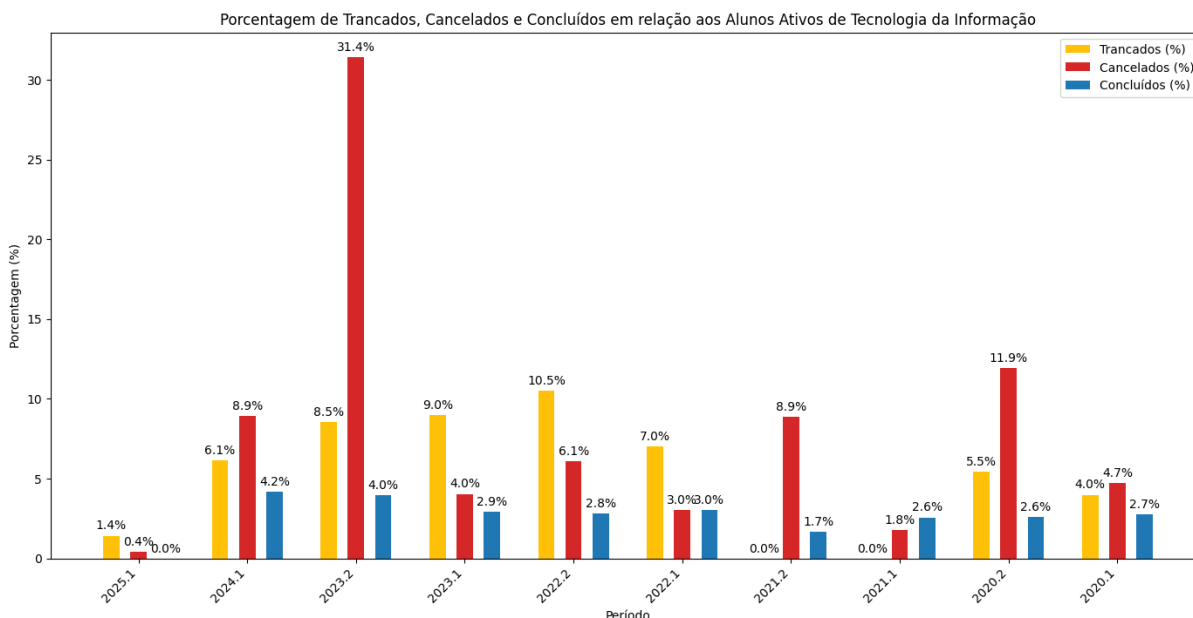


Fonte: Autoria própria

Neste Gráfico 11 encontramos os valores referentes aos alunos do curso de Engenharia de Software. Os dados mostram que a maior taxa de **conclusão** ocorreu em 2020.1 com 36,4%, enquanto a maior porcentagem de **trancamentos** foi em 2023.2 sendo de 15,6%. **Cancelamentos** foram registrados apenas em 2023.2 com o valor de 6,2%. Os períodos mais

recentes, como 2025.1 e 2024.3, não apresentaram registros de trancamento, cancelamento ou conclusão. As taxas de conclusão se mostraram positivas e percebe-se uma oscilação e depois queda nos números com o passar dos anos.

Gráfico 12 – Interdisciplinar em Tecnologia da Informação (Campus Pau dos Ferros)



Fonte: Autoria própria

Por fim, no gráfico 12 (acima), do curso de **Tecnologia da Informação**. Observa-se uma variação significativa entre os períodos, com o pico de **cancelamentos** atingindo notáveis 31,4% em 2023.2. As taxas de **trancamento** foram mais elevadas em 2022.2 10,5%. Em contraste, as porcentagens de **conclusão** são menores, com o maior índice de 4.2% registrado em 2024.1. De modo geral, o cancelamento se destaca como a categoria de maior impacto em diversos períodos.

5 RESULTADOS E DISCUSSÃO

Com todas as informações apresentadas nos gráficos, consegue-se perceber os cursos e períodos com maior destaque, em se falando dos números de evadidos (seja por trancamentos ou cancelamentos) e concluintes. O período 2023.2 destacou-se em todos os cursos por sua elevada porcentagem de cancelamentos em relação aos alunos ativos, período esse que as aulas já eram desenvolvidas 100% presencial, após a pandemia do Covid-19.

O período mencionado acima, voltou a acontecer a obrigatoriedade do aluno ter que se matricular em alguma turma, caso contrário, seu vínculo com a instituição seria cancelado, fato esse que pode ter colaborado para os números negativos expressivos, dado os alunos que não se atentaram para tal informação, como aconteceu no curso de Licenciatura em Computação e Informática, do Campus Angicos, com 54,4% das matrículas canceladas, sendo o curso com a maior taxa de cancelamentos. Pode ter ocorrido também uma possível discrepância entre os dados apresentados na base dos Dados Abertos da Ufersa em contraponto aos dados dos relatórios das coordenações de cada curso. O fato dos dados serem catalogados incorretamente, evidencia-se porque no carregamento da base não encontra-se o período 2024.1 do curso de LCI, período esse com um valor significativo de 10 (dez) alunos concluídos, segundo a própria coordenação, o que mudaria o quadro do período anterior.

Existem pontos positivos em relação aos reais cancelamentos de vínculos dos alunos com as instituições de nível superior, esse fato revela alunos que não se identificam com o curso, procrastinam durante as atividades curriculares, pessoas que se matricularam nos cursos analisados porque encontravam-se sem opções e arriscaram, evidenciando um certo desinteresse e por quererem também, seguir em outras áreas.

Ainda no mesmo Campus, o curso de Bacharelado em Sistemas de Informação teve o maior número de matrículas trancadas no período supracitado, sendo 15,9% do total de matrículas ativas. Já durante a pandemia, o curso que mais sofreu com a evasão por trancamentos foi o de Engenharia de Software, do Campus Pau dos Ferros, com 9,1% em ambos os períodos de 2020.1 e 2020.2. Seguindo por matrículas canceladas, o curso Interdisciplinar de Tecnologia da Informação, da mesma instituição já citada, obteve uma porcentagem de 11,9% no período 2020.2.

Os trancamentos revelam alunos que ao adiarem sua decisão de continuação no curso de forma eficiente, ficam estagnados nos relatórios de suas respectivas coordenações, ocupando vagas sem retorno para instituição, mercado de trabalho e governo federal, gastando recursos públicos sem real interesse em concluir o curso.

O curso com maior taxa de concluintes nos últimos cinco anos é o de Engenharia de Software, com 36,4% de alunos concluintes no período 2020.1, o mesmo acontecendo no início da pandemia.

Os cursos de Engenharia de Software e Engenharia da Computação são de segundo ciclo, os alunos que chegam até eles já passaram pelo curso de Interdisciplinar em Tecnologia da Informação, evidenciando a dedicação e capacidade dos discentes em seguir até o final dos cursos e confirmação os números expressivos de alunos concluintes apresentados nos gráficos. Embora a evasão seja um desafio em diversas áreas, alguns cursos conseguem levar uma parcela considerável de seus alunos até a formatura em determinados períodos.

6 CONSIDERAÇÕES FINAIS

Com base no que foi apresentado podemos perceber a grande importância em uma análise descritiva dos dados, evidenciando os indicadores sobre o tema evasão no ensino superior, como também, investigando os fatores internos e externos que influenciam os elevados números desse problema. O manuseio e tratamento correto dos dados traz um vislumbre para qualquer contexto/negócio e de forma patente, revela os pontos fracos e fortes a serem trabalhados. Vemos também que, com o auxílio de métodos e tecnologias eficazes, conseguimos de maneira potencial as informações pertinentes e importantes para a tomada de decisões que tornam mais eficientes os esforços empregados para a resolução do problema.

Diante dos números extraídos nos gráficos e apresentados acima, vemos como os cursos de Computação da Ufersa, enfrentam uma resistência ou pouca adaptação por parte dos discentes. Trabalhar com dados num tema como evasão escolar é um desafio que impacta significativamente os futuros debates nesse meio. Fatores (dados descritivos qualitativos) que não foram encontrados na base de dados original, poderiam ajudar a traçar perfis com características completas do discente, assim gerando um modelo de predição para atacar o problema de forma mais eficiente e eficaz.

6.1 TRABALHOS FUTUROS

No futuro é de suma importância o cruzamento dos dados carregados nas bases com os relatórios das coordenações de cada curso, de modo a se efetivar a veracidade dos mesmos e conferir se há algum possível problema de entrada de dados, buscando soluções para isso. Ao tornarmos as informações fidedignas, ao ponto de apresentar para a gestão da Universidade, trazemos a atenção dos mesmos para o tema tão pertinente e importante a ser discutido, visando um trabalho em conjunto em todos os níveis e buscando soluções satisfatórias que se perpetuem durante os anos seguintes. Também se faz pertinente o melhoramento na base de dados trazendo dados característicos que possam ajudar a elaborar um modelo de classificação/predição a partir de algoritmos para tais funções, usando para isso características como cor, gênero, classe social, base de ensino (pública ou privada), idade e a partir de então efetuar o cruzamento dessas características visando prever uma futura evasão por parte do discente, ajudando a nível de coordenações e progred a tomada de decisões.

REFERÊNCIAS

ANDIFES; ABRUEM; OTHERS. Diplomação, retenção e evasão nos cursos de graduação em instituições de ensino superior públicas: resumo do relatório apresentado a ANDIFES, ABRUEM e SESu/MEC pela Comissão Especial. Avaliação: Revista da Avaliação da Educação Superior, [S.l.], v.1, n.2, 1996. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/225423/000224712.pdf?sequence=1>.

Acesso em 27 de janeiro de 2024.

Baggi, C.; LOPES, D. Evasão e avaliação institucional no ensino superior: uma discussão bibliográfica. Avaliação – Campinas, julho, v. 16, n. 2, p. 355-374, 2011.

CAROLINA, Ana. Evasão escolar: causas, consequências e como evitar?. CRM Educacional. Outubro de 2023. Disponível em: <https://crmeducacional.com/evasao-escolar-por-que-acontece-e-entenda/>. Acesso em 02 de fevereiro de 2024.

LOBO, M. Panorama da evasão no ensino superior brasileiro: aspectos gerais das causas e soluções. Associação Brasileira de Mantenedoras de Ensino Superior. Cadernos, v. 25, 2012. Citado na página 14.

NERI, M. C. Motivos da evasão escolar. 2009. Citado 2 vezes nas páginas 14 e 18.

SESU, B. M. da Educação. Secretaria da E. retenção e evasão nos cursos de graduação em instituições de ensino superior públicas. Avaliação: Revista de rede de avaliação institucional da educação superior. Campinas, v. 1, n. 2, p. 55–65, 1996. Citado na página 18.

SILVA FILHO, Roberto Leal Lobo et al. A evasão no ensino superior brasileiro. Cadernos de pesquisa, v. 37, n. 132, p. 641-659, 2007.

SOUZA, C. T.; SILVA, C. DA; GESSINGER, R. M. Um estudo sobre evasão no ensino superior do Brasil nos últimos dez anos. Congressos CLABES, v. 0, n. 0, 9 out. 2017.

Gov.br. Censo da Educação Superior. Outubro de 2023. Disponível em <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-da-educacao-superior/resultados>. Acessado em 15 de março de 2024.

Evasão – Dados Brasil – 15º Mapa do Ensino Superior, Disponível em: <https://www.semesp.org.br/mapa/edicao-15/brasil/>. Acessado em 13 de maio de 2025.

Quantitativos de Alunos dos Cursos de Graduação – Dados Abertos Ufersa, Disponível em: <https://dadosabertos.ufersa.edu.br/dataset/quantitativos-de-alunos-graduacao/resource/c8112a24-d88e-4a4b-83c6-24f05d4ac6e0>. Acessado em 14 de maio de 2025.

Quantitativos consolidados – Dados Abertos Ufersa, Disponível em: https://dadosabertos.ufersa.edu.br/dataset/quantitativos-de-alunos-graduacao/resource/5160a2dd-a212-4de4-9734-13c3bcea6b32?inner_span=True. Acessado em 14 de maio de 2025.

Dicionário de dados - Quantitativos de Alunos – Dados Abertos Ufersa, Disponível em: https://dadosabertos.ufersa.edu.br/dataset/quantitativos-de-alunos-graduacao/resource/67b159a5-4f37-4093-98c9-dcdd8535ad59?inner_span=True. Acessado em 14 de maio de 2025.

Etapas do processo KDD (Fayyad et al. (1996)). Disponível em: https://www.researchgate.net/figure/Figura-1-Etapas-do-processo-KDD-Fayyad-et-al-1996_fig1_228435521. Acessado em 11/06/2025.