



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE TECNOLOGIA
CURSO DE ENGENHARIA DE COMPUTAÇÃO

LEONARDO JUNIO ALVES DE QUEIROZ

**FERRAMENTA PARA ANÁLISE DO DESEMPENHO
ACADÊMICO, UTILIZANDO COMO CASO DE TESTE O
PERÍODO SOB INFLUÊNCIA DO COVID-19**

Natal/RN

2022

Leonardo Junio Alves de Queiroz

Ferramenta para análise do desempenho acadêmico, utilizando como caso de teste o período
sob influência do covid-19

Trabalho de Conclusão de Curso na
modalidade Monografia, submetido como
parte dos requisitos necessários para conclusão
do curso de Engenharia de Computação pela
Universidade Federal do Rio Grande do Norte.

Orientador: Prof(a). Dr(a). Isabel Dillmann
Nunes

Natal/RN

2022

Universidade Federal do Rio Grande do Norte - UFRN
Sistema de Bibliotecas - SISBI
Catalogação de Publicação na Fonte. UFRN - Biblioteca Central Zila Mamede

Queiroz, Leonardo Junio Alves.

Ferramenta para análise do desempenho acadêmico, utilizando como caso de teste o período sob influência do covid-19 Trabalho / Leonardo Junio Alves de Queiroz. - 2022.

74 f.: il.

Monografia (Gradação) - Universidade Federal do Rio Grande do Norte, Centro de Tecnologia, Curso de Engenharia de Computação. Orientadora: Prof. Dra. Isabel Dillmann Nunes.

1. Coronavírus - Monografia. 2. Desempenho acadêmico - Monografia. 3. Análise de dados - Monografia. I. Nunes, Isabel Dillmann. II. Título.

RN/UF/BCZM

CDU 519.2:616.2

Leonardo Junio Alves de Queiroz

Ferramenta para análise do desempenho acadêmico, utilizando como caso de teste o período sob influência do covid-19

Trabalho de Conclusão de Curso na modalidade Monografia, submetido como parte dos requisitos necessários para conclusão do curso de Engenharia de Computação pela Universidade Federal do Rio Grande do Norte.

Aprovado em 08 de Janeiro de 2022

Prof (a). Dr(a). Isabel Dillmann Nunes – Orientador

UFRN

Prof(a). Dr(a). Luiz Affonso H. Guedes de Oliveira – Examinador interno

UFRN

Prof(a). Dr(a). Apuena Vieira Gomes – Examinador interno

UFRN

Natal/RN

2022

AGRADECIMENTOS

À minha família, por todo apoio, ensino e educação em toda a minha trajetória, além de sempre acreditarem em mim, sempre me motivando a tentar melhorar e me mostrando que a educação e a perseverança é o melhor caminho para fazer a vida das pessoas melhor.

À minha noiva, por todas as vezes que me têm apoiado e incentivado. Por ter acreditado em mim e por todo amor, companheirismo e carinho.

À minha professora orientadora Isabel Dillmann Nunes, por toda a disponibilidade, conselhos, paciência e dedicação. Por ter me orientado e compartilhado conhecimentos em todas as partes de desenvolvimento deste trabalho.

Aos meus colegas da Superintendência de Informática e da Inovall, por serem compreensivos, apoiadores e uma boa companhia de trabalho. Por terem me auxiliado no meu desenvolvimento profissional e intelectual, expandindo meu conhecimento durante meu percurso.

A todos os professores da Escola de Ciências e Tecnologia e do Departamento de Engenharia de Computação e Automação por cada conhecimento compartilhado, além de auxiliar minha trajetória em todo o curso.

RESUMO

Ferramenta para análise do desempenho acadêmico, utilizando como caso de teste o período sob influência do covid-19

O desempenho acadêmico dos estudantes da Universidade Federal do Rio Grande do Norte (UFRN) através dos semestres é uma informação importante para mensurar a qualidade de ensino e o impacto socioeconômico no mesmo, se tornando ainda mais importante em tempos de pandemia. Com o surgimento do vírus Covid-19, uma doença causada pelo agente coronavírus identificado em dezembro de 2019 na China, onde expandiu-se a todos os demais países, se mostrou ainda mais importante uma ferramenta para analisar o desempenho e a influência de fatores externos no regime acadêmico. Diante disso, o objetivo deste trabalho é propor uma ferramenta para estudo da performance dos estudantes, utilizando o Software Anaconda em conjunto com a linguagem de programação Python e conceitos da área de ciência de dados, desde a coleta dos dados até a interpretação e análise dos dados e considerando como caso de teste os períodos de pandemia, verificando assim o sucesso de estudantes em disciplinas acadêmicas entre semestres, considerando os status referente a aprovações, reprovações, trancados e desistências, além de dados socioeconômicos e um mockup. Com isso, mostrar a importância desta ferramenta, podendo contribuir bastante para o ensino na universidade ajudando tanto discentes como docentes, além de analisar o impacto causado pela Covid-19 no ensino.

Palavras-chave: Coronavírus, desempenho acadêmico, análise de dados.

ABSTRACT

Tool for analyzing academic performance, using the period under the influence of covid-19 as
a test case

The academic performance of students at the Universidade Federal do Rio Grande do Norte (UFRN) throughout the semesters is important information to measure the quality of teaching and the socioeconomic impact, becoming even more important in times of pandemic. With the emergence of the Covid-19, a disease caused by the coronavirus agent identified in December 2019 in China, where it expanded to all other countries, proved an even more important tool to analyze performance and influence of factors external in the academic regime. Given this, the objective of this work was to propose a tool for the study of student performance, using the Anaconda Software with Python programming language and concepts in the area of data science, from data collection to interpretation and analysis of the data and considering the pandemic periods as a test case, verifying student success in academic disciplines between semesters, considering the statuses regarding approvals, disapprovals, locked and dropouts, in addition to socioeconomic data and a mockup. With this, show the importance of this tool, being able to contribute a lot to the teaching in the university helping students and professors, in addition to analyzing the impact caused by Covid-19 in education.

Keywords: Coronavirus, academic performance, data analysis.

LISTA DE ILUSTRAÇÕES

FIGURA		PÁGINA
1	Processos para análises de dados	18
2	Cálculo da média	19
3	Etapas de KDD	21
4	Áreas de conhecimento de uma EDM	22
5	Passos da análise de dados	25
6	Refinamento dos dados	26
7	Histograma	34
8	Gráfico de barras	34
9	Gráfico de setores	35
10	Mapa de calor	35
11	Série ou pontos conectados	36
12	Fluxo de atividades	40
13	Script para consulta e gráfico	45
14	Quantidade geral de discentes em DIS0001	46
15	Quantidade de discentes em DIS0001	47
16	Média das notas em DIS0001	47
17	Média das faltas em DIS0001	48
18	Quantidade de aprovações por status em DIS0001	48
19	Quantidade de reprovações por status em DIS0001	49
20	Quantidade de trancamentos em DIS0001	49
21	Quantidade de desistências em DIS0001	50
22	Quantidade de discentes gerais em DIS0002	51
23	Quantidade de discentes em DIS0002	51
24	Média de notas em DIS0002	52
25	Média de faltas em DIS0002	52
26	Quantidade de aprovações por status em DIS0002	53
27	Quantidade de reprovações por status em DIS0002	53
28	Quantidade de trancamento em DIS0002	54
29	Quantidade de discentes gerais em DIS0003	54
30	Quantidade de discentes em DIS0003	55

31	Média de notas em DIS0003	55
32	Média de faltas em DIS0003	56
33	Quantidade de aprovações por status em DIS0003	56
34	Quantidade de reprovações por status em DIS0003	57
35	Quantidade de discentes gerais em DIS0001 e DIS0002	58
36	Quantidade de discentes em DIS0001 e DIS0002	58
37	Média de notas em DIS0001 e DIS0002	59
38	Média de faltas em DIS0001 e DIS0002	59
39	Quantidade de aprovações por status em DIS0001 e DIS0002	60
40	Quantidade de reprovações por status em DIS0001 e DIS0002	60
41	Quantidade por escola ensino médio em DIS0002	61
42	Quantidade de bolsas em DIS0002	61
43	Classe econômica em DIS0002	62
44	Quantidade de alunos por quantidade de matrículas	62
45	Sucesso por um discente ao longo dos semestres	63
46	Filtro e tabela de dados no Mockup	67
47	Gráficos no Mockup	68
48	Exemplo de filtros com dados da disciplina DIS0002	68
49	Exemplo de informações obtidas referente a disciplina DIS0002	69

LISTA DE TABELAS

TABELA	PÁGINA
1 Comparação entre turmas	64

LISTA DE ABREVIATURAS E SIGLAS

SIGLA	SIGNIFICADO
UFRN	Universidade Federal do Rio Grande do Norte
SIGAA	Sistema Integrado de Gestão de Atividades Acadêmicas
OKF	<i>Open Knowledge Foundation</i>
KDD	<i>Knowledge Discovery in Databases</i>
EDM	<i>Educational Data Mining</i>
LA	<i>Learning Analytics</i>
PPP	Projeto Político-Pedagógico
UEMs	<i>Usability Evaluation Methods</i>
UX	<i>User Experience</i>
STI	Superintendência de Tecnologia da Informação
SGBD	Sistema de Gerenciamento de Banco de Dados
SQL	<i>Structured Query Language</i>
API	<i>Application Programming Interface</i>
CSV	<i>Comma-Separated Values</i>

SUMÁRIO

CAPÍTULO	PÁGINA
1 INTRODUÇÃO	14
1.1 Objetivos	15
1.2 Organização do documento	15
2 FUNDAMENTAÇÃO TEÓRICA	17
2.1 Análise descritiva	17
2.2 Análise de dados educacionais	19
2.2 Desempenho dos alunos	27
2.4 Visualização de dados	31
2.5 Mockup	36
3 METODOLOGIA DE PESQUISA	39
3.1 Fluxo de atividades	39
3.2 Ferramentas utilizadas	40
3.1.1 MySQL	40
3.1.2 Python	41
3.1.3 Anaconda	41
3.1.4 Bibliotecas	41
3.3 Convenções utilizadas	42
3.4 Dados	42
4 DESENVOLVIMENTO E RESULTADOS	44
4.1 Requisição e armazenamento das informações	44
4.2 Gráficos e informações obtidas	45
5 PROPOSTA DA FERRAMENTA	65
5.1 Mockup	66
6 CONCLUSÃO	70

1. INTRODUÇÃO

No final do semestre de 2019.1, diferente dos cenários comuns entre os períodos acadêmicos, foi descoberto na China uma doença causada por um novo tipo de coronavírus, denominado SARS-CoV-2, também chamada de Covid-19. No semestre de 2020.1 foram registrados os primeiros casos no Brasil, mudando o estilo de vida da população para um novo normal. Uma das áreas afetadas foi a educação, onde redes de ensino tiveram que paralisar o semestre e mudar para o regime de ensino remoto, como foi o caso da Universidade Federal do Rio Grande do Norte (UFRN).

Mesmo com o decorrer do tempo, ainda sente-se os efeitos causados por este novo estilo de vida e de ensino, o qual gerou o seguinte questionamento: Como foi o desempenho dos alunos durante o ensino remoto comparado com o ensino tradicional, ou seja, o ensino presencial?

Com a ausência de uma ferramenta para auxiliar a comunidade acadêmica a comparar o desempenho dos estudantes e observar quais os fatores possíveis para um desempenho acima ou abaixo da média (comparado aos outros períodos estudantis) e indicar melhorias no ensino e da administração dos cursos, a análise e verificação do desempenho dos alunos se torna uma experiência complicada, sendo estes itens mencionados de extrema importância tanto para auxiliar os alunos como para auxiliar todos os professores, coordenadores e gestores promovendo insights para a gestão e coordenação de cursos.

Em virtude disso, foi pensado em uma ferramenta para que se possa averiguar o impacto no desempenho dos estudantes da UFRN, ou seja, se os discentes de um determinado período possuem índices positivos ou negativos quando comparado a outros períodos acadêmicos. Como caso de teste, será utilizado o período situado em pandemia comparando aos outros semestres e assim, verificar a forma que fatores como a Covid-19, influenciou no ensino dos alunos da graduação e da pós-graduação. Além disso, foi projetado um mockup como exemplo de como será feita a visualização das informações e como poderia ficar o visual do produto.

O uso dessa ferramenta irá se basear nos dados fornecidos pelo Sistema Integrado de Gestão de Atividades Acadêmicas (SIGAA). Com esta ferramenta seria possível qualificar e verificar as diferenças de desempenho entre estudantes e com isso saber como alguns aspectos externos influenciaram no ensino como um todo, auxiliando todos os envolvidos da área de ensino a formular um plano de ensino mais adequado para futuros semestres.

O SIGAA é um sistema que gerencia e informa as informações e procedimentos da comunidade acadêmica da UFRN, desde atividades de níveis de ensino a projetos de pesquisa e extensão.

Com isso, os principais beneficiários dessa pesquisa serão os alunos, os professores e os gestores da UFRN.

Os beneficiários diretos serão os docentes, coordenadores e gestores, para que com essas análises proposta, auxilie na preparação das aulas e construção de um currículo mais adequado ao semestres e que possa compreender melhor a diferença de desempenho dos alunos com base na disciplina ou curso como um todo, assim impactando positividade no planejamento e realização das atividades pedagógicas.

Como beneficiário indireto podemos considerar os discentes e os outros pesquisadores da instituição, pois espera-se que se possa melhorar a construção do currículo acadêmico, sendo mais assertivo e assim construindo um ambiente de aprendizagem mais positivo; além de ser um processo de melhoria contínua por estar relacionado aos dados educacionais dos estudantes, assim auxiliando para que as decisões institucionais sejam mais assertivas, beneficiando o ensino.

1.1. Objetivos

O objetivo deste trabalho é propor uma ferramenta para análise do desempenho dos alunos da graduação e pós-graduação entre períodos acadêmicos. Com isso, temos como objetivos específicos:

- Investigar análise exploratória de dados;
- Desenvolver análises com base nos dados de turmas de graduação e pós-graduação dos períodos 2016.1 até 2020.2;
- Propor uma ferramenta para comparar o desempenho dos alunos da graduação e pós-graduação entre períodos acadêmicos, dentre estes o período em situação de pandemia, com base em dados econômicos, notas e frequências;
- Criar um mockup de software para melhor análise e visualização das informações;

1.2. Organização do documento

O presente documento está organizado na seguinte estrutura:

- Capítulo 1: Possui a introdução, contexto, a problemática proposta e os objetivos deste trabalho;
- Capítulo 2: Apresenta os principais temas e conceitos relacionados a esta pesquisa;
- Capítulo 3: Exibe as ferramentas utilizadas e a metodologia de pesquisa deste trabalho;
- Capítulo 4: São discutidos os dados obtidos e apresentando a ferramenta proposta;
- Capítulo 5: Apresenta a conclusão e a sugestão de um trabalho futuro como continuação;

2. FUNDAMENTAÇÃO TEÓRICA

Considerando os objetivos deste trabalho, nesta seção é apresentada a fundamentação teórica desta pesquisa.

Na seção 2.1, são apresentados os conceitos e abordagens relacionados à análise de dados educacionais. Na seção 2.2, é exemplificado sobre o item referente ao desempenho do aluno. Na seção 2.3, são apresentados a importância e a utilização de um formato visual para a análise de dados. Por fim, na seção 2.4, são exibidos os conceitos e justificado a importância de um mockup.

2.1. Análise descritiva

Segundo Reis, E.A., Reis I.A. (2002), a análise descritiva é a fase inicial do processo de estudo de uma coleta de dados estatísticos. Ela pode ser utilizada para as primeiras manipulações em um estudo quantitativo para que se possa organizar, resumir, explorar e analisar o comportamento dos dados, observando os aspectos importantes de um conjunto de características ou comparar tais características entre conjuntos, sendo assim possível tentar abordar as causas e estipular possíveis consequências de alguma situação.

Ela pode ser abordada como um dos ramos de estatística, onde segundo Echeveste, Bittencourt, Bayer e Rocha (2005), é uma ciência que se dedica ao desenvolvimento e ao uso de métodos para a coleta, resumo, organização, apresentação e análise de dados.

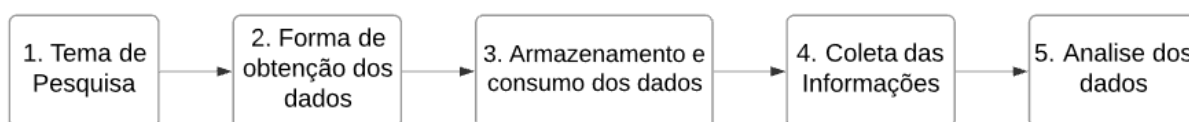
Quando abordado sobre essas áreas, é importante citar os seus principais princípios, onde auxilia nas etapas e se torna mais prático a análise:

- Universo estatístico: também chamado de população, abrange todos os elementos da pesquisa abordada, onde a definição deste, depende do objetivo do estudo.
- Amostra: é um subconjunto de dados com base no universo estatístico. Com ela, se faz segregações relacionadas ao tema e permite que a análise seja feita com base em grupos.
- Dado estatístico: é um elemento que está incluso ao conjunto da população, e por consequência, relacionado com o tema da pesquisa.
- Variável: toda pesquisa precisa estar relacionada a um tema que se pretende estudar, ou seja, precisa ser definido um objeto de estudo, que denominamos de variáveis, também conhecida como observação.

As variáveis podem ser classificadas como quantitativa ou qualitativa. As variáveis quantitativas são referentes a quantidade, como por exemplo idade, notas acadêmicas e afins, podendo ser divididas em discretas (possuem aspecto mensurável para assumir um número finito e geralmente são resultados de contagens) e contínuas (possuem aspecto mensurável de uma escala contínua, podendo abordar uma classificação de indicativos ordinais), enquanto que as variáveis qualitativas são aquelas que realizam a medição referente a qualidade e podem ser separadas em categorias, como por exemplo nível de escolaridade, modalidade de ensino e outras semelhantes, podendo ser classificadas em nominais (não possuem ordenação atrelada às categorias) e ordinais (possuem ordenação atrelada às suas categorias).

Independente da proposta do objeto de pesquisa, é importante como primeiro passo definir o tema da pesquisa, pois essa será a direção que toda a análise se baseará, como pode ser observado pela figura 1. Após isso, pensar na forma de obtenção dos dados necessários para a pesquisa, assim confirmando se os dados escolhidos estão relacionados ao tema e se estes podem ser capturados, além de estruturar onde eles serão armazenados e consumidos (ato de seleção das informações para montar gráficos e tabelas) quando iniciar a análise. Tendo definido estes passos, a quarta etapa é a coleta de todas as informações necessárias para que possa ser realizado o estudo. Com todas estas etapas concluídas, pode-se iniciar o processo de análise de dados para que com base na interpretação destes, possa-se chegar a um resultado.

Figura 1 - Processos para análises de dados



Fonte: Autoria própria

Segundo Ferreira (2015), o estudo relacionado a estatística pode ser dividido em descritiva e inferencial. A análise descritiva tem como pontos principais a organização, análise e apresentação de dados, assim fornecendo uma descrição de forma direta de qualquer conjunto de informações. Algumas de suas ferramentas são as medidas de posição, tais como média, desvio padrão, mediana, variância e moda e tendo que o estudo contém etapas bem definidas, desde a organização até a análise das informações, a apresentação de dados tende-se a ter um maior grau de confiabilidade. A análise inferencial tem como pontos principais o estudo de uma amostra de um universo, ou seja, um conjunto de dados menor para ser estudado e chegar a uma dedução significativa sobre o conjunto de dados maior, neste modelo é importante se atentar na escolha da amostra, para que a população seja bem representada.

De acordo com Ignácio (2010), os métodos estatísticos foram desenvolvidos como uma mistura de ciência, tecnologia e lógica para a solução e investigação de problemas em várias áreas do conhecimento.

Os processos relacionados à estatística sem o auxílio da tecnologia, pode ser tornar lentos ou até mesmo cansativos, assim a utilização de ferramentas podem auxiliar e acelerar o resultado, inclusive poderá prover de dados mais confiáveis, uma vez que pode ser estruturados algoritmos para o processamento destas informações de maneira mais automatizada.

Segundo Bulhões (2011), as medidas podem ser diferenciadas em de posição e de dispersão. As medidas de dispersão descreve o quantos os dados são semelhantes ou distintas de um valor central, como por exemplo variância e medida padrão. As medidas de posição são referentes a resumir um conjunto de dados a uma tendência central, como por exemplo: média, moda e mediana.

- Média: é o valor que representa a concentração dos dados de uma amostra. Ela pode ser calculada como a somatória de todas as observações da amostra (X_i) dividida pela a quantidade total de amostras (N), como pode ser observado na figura 2:

Figura 2 - Cálculo da média

$$\bar{X} = \frac{\text{Soma de todas as observações da amostra}}{\text{Tamanho da amostra}} = \frac{\sum X_i}{N}$$

Fonte: Autoria própria

- Moda: é o valor que representa o dado que mais aparece em uma amostra, ou seja, o valor mais frequente em um conjunto de dados.
- Mediana: é o valor que representa o centro de um conjunto de valores ordenados, dividindo assim em dois conjuntos de mesmo tamanho. Caso a quantidade de amostras seja par, é necessário realizar a média entre os dois valores do meio da amostra.

Segundo Silvestre (2007), o objetivo último da análise estatística é o de tirar conclusões a partir dos dados por forma a enriquecer o nosso conhecimento da realidade. Para ser exibido os resultados obtidos através da coleta de dados, pode ser utilizado como forma de representação a descrição tabular, que se baseia na utilização de tabelas e descrição gráfica, que realiza o uso de gráficos como por exemplo de linha e de coluna.

2.2. Análise de dados educacionais

Segundo Enise Teixeira (2003), a análise de dados é o processo de formação de sentido além dos dados, com o objetivo de organizar e resumir os dados de tal forma que possibilitem o fornecimento de respostas ao problema proposto para investigação. Com isso, pode ser considerada como uma técnica em absorver, tratar e transformar dados através de procedimentos matemáticos (estatísticos) em informação para que se possa averiguar e determinar possíveis utilizações.

A partir de dados desestruturados, são realizados procedimentos para estes serem compilados e definidos para que possam ser trabalhados, conseguindo captar conhecimento referente a marketing, relacionamento com o cliente dentre outras funções e áreas de atuação. Quando processados, essas informações podem ser analisadas com auxílio da tecnologia de maneira técnica e considerando que atualmente estamos em uma era digital, é fundamental realizar análises eficazes para gerar mais oportunidades de crescimento para instituições de âmbito público e privado, independente do seu porte ou sua atividade econômica.

Estes dados podem ser obtidos por meio de dados abertos, que são dados livres para serem utilizados, modificá-los e compartilhados para qualquer propósito, talvez mediante a possíveis exigências, conforme definido pela *Open Knowledge Foundation* (OKF).

Como dito por Tiago Nascimento (2019), “Os dados são o novo petróleo do século XXI”, além de serem o primeiro passo para se iniciar uma análise de dados. Eles podem ser utilizados em diversos tipos de análises, como por exemplo de desempenho acadêmico de uma rede de ensino, fiscalização de informações providas pelo governo, desenvolvimento de serviços inovadores, o estudo para criação de negócios inovadores e a produção de novos conhecimentos ao combinar diferentes fontes de dados e padrões.

Assim sendo possível utilizá-los independente da finalidade, estando no máximo sujeitos à exigência de creditar a autoria e preservar a licença aberta, conforme citado no sítio do Governo Federal.

A abertura de dados de universidades é de fundamental importância para produção de conhecimento e análises, podendo auxiliar em análises de ensino ou até mesmo de contexto social, tornando mais visível para a sociedade informações econômicas e sociais que à cerca e a produção de novas ferramentas.

Os dados para serem acessíveis ao público, devem conter os seguintes requisitos, conforme declarado pelo site do banco central:

- Serem livres de sigilo;

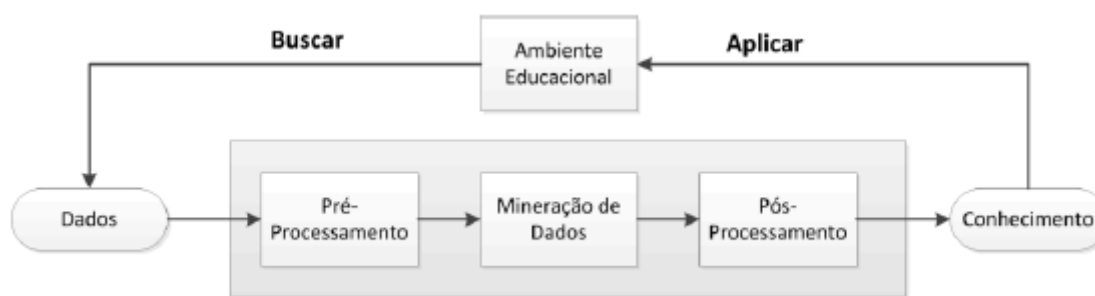
- Representados em meio digital;
- Estruturados em formato aberto;
- Processáveis por máquina;
- Referenciados na internet;

Disponibilizados sob licença aberta que permita sua livre utilização, consumo ou cruzamento. Contendo os requisitos acima, os dados podem ser utilizados pelo público para produzir novas análises e ferramentas que poderão contribuir com o desenvolvimento de um instituição ou até mesmo do país, considerando que poderá haver aprimoramento da máquina estatal em virtude da construção de processo e soluções mais eficientes e inovadoras.

Para realizarmos as análises necessárias, é de suma importância a etapa de prospecção ou mineração de dados, que é uma área interdisciplinar, mobilizando principalmente conhecimentos de análise estatística de dados, aprendizagem de máquina, reconhecimento de padrões e visualização de dados (Costa, Bajer, Amorin e Marinho, 2012). Ela é utilizada principalmente para analisar um grande volume de dados, também conhecido como *Big Data*, e captar destes dados relevantes as informações. *Big Data* é um conjunto de dados e informações coletados digitalmente onde são interpretados com o auxílio de um sistema eletrônico.

Ela pode ser vista como uma das etapas de um processo conhecido como descoberta de conhecimento em bases de dados (*Knowledge Discovery in Databases*, KDD), das quais as outras etapas são o pré-processamento de dados, que realiza a preparação e tratamento dos dados e o pós-processamento, que envolve a visualização e interpretação dos dados, como pode ser observado na figura 3.

Figura 3 - Etapas de KDD



Fonte: Colpani (2018).

Segundo Costa, Bajer, Amorin e Marinho (2012), KDD é o processo não trivial de identificação de padrões, a partir de dados, que sejam válidos, novos, potencialmente úteis e compreensíveis. Para definir se o conhecimento é relevante, pode-se trabalhar com metas,

sendo elas definidas em 2 tipos, verificação ou descoberta, variando de acordo com o objetivo da aplicação.

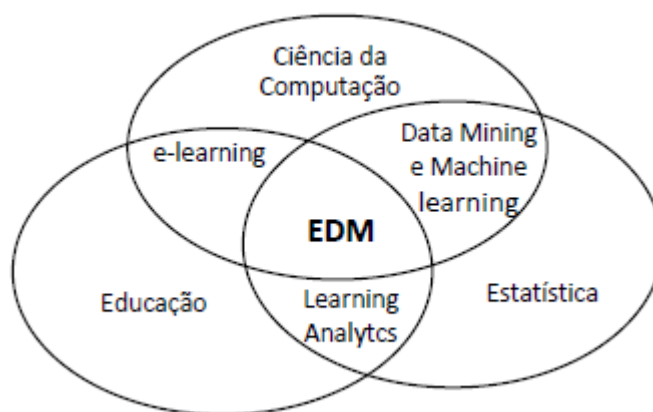
- Verificação: o sistema está limitado a verificar hipóteses definidas pelo usuário;
- Descoberta: o sistema encontra novos padrões de forma autônoma;

Com o uso de ambientes virtuais de aprendizagem (sistemas que armazenam informações e ferramentas de cursos para uma comunidade virtual) e de sistemas de informações educacionais (*software* para gestão integrada para desenvolvimento e acompanhamento das atividades das instituições de ensino¹) por instituições de ensino, são gerados grande quantidade de dados através dos usos e registro de conteúdos contínuos por todos atores de sistemas educacionais, desde discentes à gestores e demais.

Com este alto volume de dados, pode-se obter mais do que informações gerenciais e ampliar até relatórios mais completos e realizar a descoberta de novos conhecimentos relevantes nesses dados. Com isso, é visível que o uso de técnicas de mineração de dados educacionais (*Educational Data Mining*, EDM) tem se tornando mais essencial para a evolução dos ambientes de ensino.

O processo de EDM consiste em converter os dados brutos de sistemas educacionais em informações úteis para serem trabalhadas e refinadas e tem se tornado uma área de pesquisa cada vez mais relevante em diversas áreas, a fim de resolver diversas questões. Segundo Colpani (2018), a EDM é a combinação de três principais áreas de conhecimento: Computação, Educação e Estatística, como pode ser observada na figura 4, ela é a interseção de diversas áreas.

Figura 4 - Áreas de conhecimento de uma EDM



Fonte: Colpani (2018).

¹<https://npd.ufes.br/sistema-de-informação-para-o-ensino-sie>

Com as informações obtidas a partir da mineração de dados educacionais, é importante citar o conceito de *Learning Analytics* (LA), que consiste em procedimentos com abordagem analíticas com o objetivo de realizar análises sobre o desempenho educacional de estudantes com base em dados de ensino. Com isso, ela pode coletar, organizar e analisar os dados de forma mais profunda dos estudos de estudantes, auxiliando a perceber os impactos na aprendizagem dos mesmos e aos docentes, pensar estratégias de acordo com as dificuldades encontradas que podem influenciar no aprendizado.

Para os servidores de ensino, as ferramentas de Learning Analytics servirão para otimizar o tempo que é utilizado para realizar a análise do ensino e propor formas de melhorá-lo, pois devido a grande quantidade de disciplina, turmas e discente, uma ferramenta de fácil acesso e que possua essas informações, pode não só diminuir o tempo gasto para tais análises como melhorar a abordagem e auxiliar para melhores decisões.

A análise de dados consiste em extrair sentido dos dados de texto e imagem. Envolve preparar os dados para análise, conduzir análises diferentes, aprofundar-se cada vez mais no entendimento dos dados, fazer representação dos dados e fazer uma interpretação do significado mais amplo dos dados (Creswell, 2007).

A análise de dados pode ser divididos em 4 tipos diferentes, são eles:

- **Análise descritiva:** também chamado de estatística descritiva, é a análise de dados baseado em fatos. Tende-se a ter execução mais rápida por a avaliação dos dados é feita a partir de cálculos padronizados, podendo ser inclusive automatizados, auxiliando a compreender tendências ou padrões que possam ser mapeados. São bastantes utilizados em indicadores de desempenho, aplicações de métricas e relatórios para orientar a construção de estratégias tomando as melhores decisões em relação ao tema analisado, além disso, apresenta a vantagem de ser imparcial nas informações. Como por exemplo, um dos temas que se utiliza bastante este tipo de análise é o desempenho acadêmico.
- **Análise preditiva:** é a análise de dados baseada em formulação de previsão de cenários futuros, com base em padrões encontrados nas informações provenientes dos dados, assim sendo possível determinar uma relação de tendência ou probabilidade. Importante citar que a proposta não é prever o que vai acontecer, mas o que será possível de acontecer com base em determinadas condições.

- **Análise prescritiva:** é a análise de dados baseada em projeção, ou seja, a potencialidade de algo previsto baseado em padrões e fatores de influência, podendo assim realizar estudos e abordar a melhor possibilidade de sucesso para obtenção dos melhores resultados. Por serem análises que fruto de dados e predições, eles podem constantemente ser alterados, assim são geralmente construídas com auxílio de tecnologia de inteligência artificial e algoritmos para definição das melhores estratégias a serem seguidas.
- **Análise diagnóstica:** é a análise de dados baseada em fatos, assim como a análise descritiva, mas com foco em determinar relações de causa e efeito de determinados acontecimentos com base em probabilidades, pois conhecendo as características dos dados referente ao acontecimento, pode-se formular possíveis fatores que levaram à ocorrência da situação, além de analisar o alcance e o impacto de uma ação.

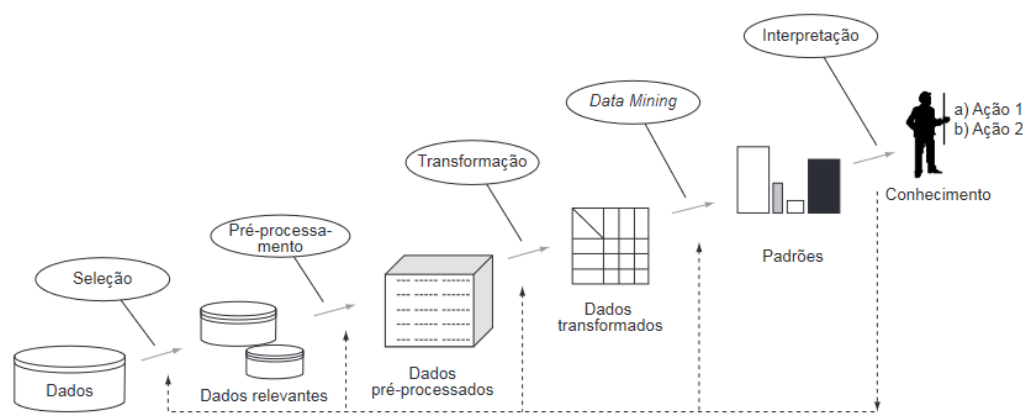
Assim, para auxiliar na obtenção dos melhores resultados de uma análise, pode-se utilizar os seguintes procedimentos como base, desde da seleção das informações até a produção de conhecimento:

- **Coleta de dados:** se refere a captura dos dados a serem utilizados na análise, podendo conter dados que não são úteis à pesquisa. É importante determinar quais informações poderão ser coletadas e a forma que serão armazenadas.
- **Processamento de dados:** etapa referente à organização dos dados, realizando a categorização das informações e relacionando-as. Nesta etapa pode-se agrupar em categorias e atribuir a códigos com significado específico para auxiliar na análise.
- **Limpeza de dados:** tratamento para que os dados possuam as informações nos formatos adequados e válidos, assim no ato de realizar a análise preliminar dos dados, terá menor probabilidade de conter dados indesejados e mais confiabilidade no que será exibido.
- **Análise exploratória de dados:** análise preliminar dos dados para entender melhor os padrões e os dados obtidos. Nesta etapa pode-se recorrer a gráficos e tabelas como forma de visualização para auxiliar a análise e na interpretação dos dados, utilizando os conhecimentos teóricos sobre o assunto abordado.
- **Produto de dados:** após as etapas anteriores, é realizado o desenvolvimento de um produto, seja este um software ou conhecimento, através da interpretação provenientes da análise dos dados. Pode ser utilizado um software para facilitar

a análise e a apresentação dos dados. Inclusive, pode-se apontar prováveis aspectos para serem abordados em futuras pesquisas.

Para um interpretação visual dos passos citados, pode-se visualizar a figura 5, a qual podemos perceber que o procedimento se inicia com um conjunto geral de dados, onde se realiza uma seleção de dados relevantes para a pesquisa; com isso se realiza o pré-processamento em conjunto com a transformação em dados úteis com limpeza e processamento dados, para que estes sejam organizados e suas informações possuam formato e valores válidos para a análise e por fim realizar a produção de conhecimento.

Figura 5 - Passos da análise de dados



Fonte: scielo²

Conforme Creswell (2007), a informação a ser analisada pode ser melhorada com o uso de programas de computador com software qualitativo. Com isso, os processos de filtragem e as análises se tornam mais simples e mais confiáveis.

Segundo Mozzato e Grzybovski (2011), às análises pode ser algo fastidioso quando feito manualmente. As etapas da análise podem ser complicadas e demoradas, além de monótonas caso não sejam utilizadas técnicas e ferramentas tecnológicas que auxiliam as etapas. Utilizando-as pode se obter não só melhores resultados, como também deixar o procedimento mais prático, para caso tenha alguma alteração nos dados e seja necessário refazer.

Segundo Lopes (2008), a educação para cidadania, um dos grandes objetivos da educação de hoje, exige uma ‘apreciação’ do conhecimento moderno, impregnado de ciência e tecnologia. Assim mostrando a importância de incorporar o uso de ferramentas tecnológicas à análise de dados. Um ponto importante na frase de Lopes é a citação ‘educação para cidadania’ como objetivo, assim podendo ser afirmado que a tecnologia está diretamente

² <https://www.scielo.br/j/gp/a/gC9RkgLD8B8FffPsNhBYWkB/?lang=pt>

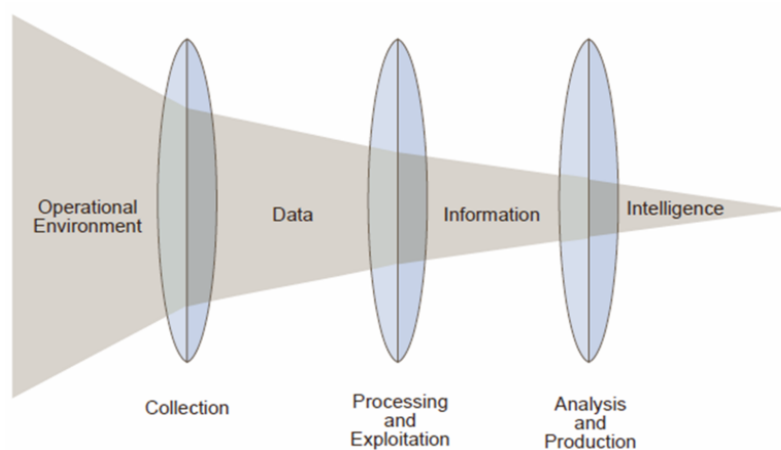
relacionada à educação das pessoas em geral, seja na passagem de conhecimento e métodos de ensino, seja em utilizá-la na análise de desempenho dos estudantes.

Tendo os dados já transformados, é realizada uma análise preliminar dos dados para ter uma amostra das informações e assim planejar como eles serão trabalhados e qual o significado das informações obtidas. Como última etapa, pode-se utilizar algum software ou ferramenta para auxiliar na análise e realizar as devidas interpretações das informações referidas à pesquisa e realizar a construção do conhecimento, que pode ser considerado como produto da análise.

Conforme Mozzato e Grzybovski (2011), o objetivo da análise de conteúdo é compreender criticamente o sentido das comunicações, seu conteúdo manifesto ou latente e as significações explícitas ou ocultas. Um ponto importante da análise das informações obtidas é a relação com a teoria do assunto e o que aquelas informações poderão influenciar na tomada de decisões e planejamento em situações na realidade, ou seja, no contexto além da pesquisa, pois a mesma pode conter aspectos que não são explícitos no momento que visualiza os informações obtidas.

Com isso, podemos perceber que a quantidade de dados a cada etapa tende-se a diminuir e tornar-se mais condizente com a pesquisa que está sendo realizada, devido a ser refinado por cada procedimento, gerando inteligência com as análises realizadas. Para exemplificar o refinamento dos dados gerais até a análise, pode-se observar a figura 6, onde que após cada etapa de processamento, a quantidade dados tem sua quantidade reduzida, entretanto resulta-se em informações mais assertivas, gerando assim a possibilidade de uma análise mais confiável e de um produto que possa abordar de melhor forma o tema proposto.

Figura 6 - Refinamento dos dados



Fonte: Wikipédia - Análise de dados³

³ https://pt.wikipedia.org/wiki/An%C3%A1lise_de_dados

Um ponto a ser citado é de Mozzato e Grzybovski (2011), o produto final da análise de uma pesquisa, por mais brilhante que seja, deve ser sempre considerada de forma provisória e aproximativa, ou seja, a análise do tema deve ser considerada provisória, considerando que os dados utilizados naquele momento compreende a aquele período de tempo e á aquele cenário social e para ser analisado em outros tempo ou até mesmo em cenários diferentes será necessário novas análises. Caso tenha sido utilizado ferramentas tecnológicas nas etapas de tratamento e análises, esta alteração tende-se a ser feita de forma simplificada e acelerando o ajuste do produto ou na geração de novos produtos com o conhecimento a ser passado atualizado. E com isso, sendo úteis para realizar atividades de tomada de decisão.

Diante diversos temas de atuação, a análise com base em um conjunto de dados referente a ensino ou meio acadêmico para fazer análises sobre o desempenho educacional de estudantes pode ser dita como *Learning Analytics* (LA) e em conjunto com a prática de mineração de dados, denominada *Educational Data Mining* (EDM). Esta área de pesquisa é de extrema importância pois pode auxiliar na análise mais adequada na forma que o aprendizado ocorre e sobre a abordagem e métodos de ensino utilizada em uma instituição. Pode-se observar que essas duas áreas são fundamentais para análises de desempenho acadêmico.

Segundo Lacerda (2018), "LA e EDM constituem um ecossistema de métodos e técnicas que coletam, processam, reportam e atuam periodicamente os dados legíveis de máquina de forma contínua para desenvolver o ambiente educacional e refletir sobre os processos de aprendizagem". Ou seja, após os procedimentos padrões da análise de dados, como resultado terá um produto que poderá auxiliar na tomada de decisões no ramo acadêmico e na interpretação de possíveis razões do rendimento dos estudantes.

2.3. Desempenho dos alunos

O acompanhamento e a mensuração do desempenho acadêmico de uma turma em uma disciplina, ou até mesmo em um curso, é um ponto essencial para que se possa avaliar a influência interna (por exemplo a metodologia de ensino utilizada em sala de aula) e externa (fatores políticos-socioeconômicos) no aprendizado. Ainda que considerando o fato de o processamento da informação no ser humano ser um processo dinâmico e complexo (Boruchovitch, 1999), não permitindo uma análise específica sobre o desempenho e sim mais geral, tentando abordar todos os aspectos de influência, vale-se tentar mensurar e propor

hipóteses a partir de dados provenientes de análises das disciplinas. Com os estudos sob essas informações, pode-se tentar propor aprimoramentos viáveis para melhorar os resultados da relação de ensino entre docente e discente.

Segundo Boruchovitch (1999), as estratégias de aprendizagem podem ser consideradas como qualquer procedimento adotado para a realização de uma determinada tarefa. Entretanto muitas vezes a estratégia formada pela coordenação de ensino podem não obter os resultados esperados e sem uma ferramenta que auxilie na análise deste item, torna-se complicado para um docente e até mesmo para a gestão administrativa identificar uma possível causa do resultado referente aos discente não ter saído como planejado pelo projeto pedagógico do curso, considerando que disciplinas podem ter diversas turmas e docentes que a lecionam, tornando assim o processo de análise dependente dos fatores externos, mas também da metodologia individual do docente utilizada. Sem uma ferramenta para abordar de maneira prática a similaridade e diferenças entre turmas e professores, torna-se um trabalho árduo tentar identificar possíveis erros na metodologia e propor melhorias.

Cada curso em uma universidade possui seu Projeto Político-Pedagógico (PPP), onde segundo Kelly Drumond (2020), define os objetivos, diretrizes e as ações para que a instituição de ensino consiga obter as melhores capacidades de ensino em função da absorção de conhecimento pelos discentes. Ele possui 3 pontos principais, que são:

- Proposta curricular: Referente ao ensinamento do docente, estabelecendo os conteúdos abordados e como eles são abordados em ambiente acadêmico, além da forma de avaliação, carga horária e o local escolar necessário.
- Organização dos professores: Organização da equipe de professores para exercer as atividades necessárias pelas intenções educativas para que não sobrecarregue algum dos envolvidos.
- Gestão administrativa: A gestão tem como principal função viabilizar os itens necessários para que os pontos acima sejam exercidos da melhor forma e que os meios se condizem pelas práticas da rede de ensino.

Para isso, é preciso compreender a relação e a importância entre o conhecimento compartilhado da relação docente-discente e os fatores de contexto social, para que as metodologias e a equipe estruturada possam trabalhar em conjunto e convergir para que o conhecimento seja passado da melhor forma possível.

Avaliar o rendimento ou desempenho escolar é tarefa árdua tanto para professores quanto para técnicos que enfrentam a falta de instrumentos de medida desenvolvidos e apropriados à nossa cultura, além da complexidade dos processos envolvidos (Capellini,

Tonelotto e Ciasca, 2004). Ou seja, além das dificuldades sociais e de formular a melhor forma de ensino dentre diversos estilos acadêmicos de estudantes, a falta de instrumento de medida ou de análise tem a tendência de dificultar a formulação de um bom projeto pedagógico.

Nessa conjuntura, é construído um currículo efetivo contendo as disciplinas para os cursos e a progressão necessária para a conclusão dos mesmos e atribuídos à equipe de professores a responsabilidade de adaptar a prática pedagógica às particularidades da turma, seja fatores políticos, culturais e/ou econômicos.

Segundo Capellini, Tonelotto e Ciasca (2004), entendem que a dificuldade de aprendizagem refere-se a qualquer dificuldade observável vivenciada pelo aluno para acompanhar o ritmo de aprendizagem de seus colegas da mesma idade, independentemente do fator determinante da defasagem. Com isso, poderá ter uma análise a nível mais assertivo, pois apesar de ainda ser uma análise geral, tende a fornecer ao indivíduo responsável pela análise uma noção sobre possíveis influências e diferenças entre turmas e afins.

Segundo Gomes (2018), os fatores que influenciam no desempenho podem ser diferenciados entre internos e externos.

Pode ser considerado como influência interna a adaptação necessária atribuída ao docente da turma para que possa atender e passar o devido conhecimento aos diferentes estilos de alunos, assim determinado pela metodologia utilizada pelo docente. Importante citar que em caso de resultados abaixo do esperado, não significa que é proveniente somente da metodologia escolhida pelo professor, pois tal metodologia pode não ter sido adequada diante fatores externos daquele período.

E como influência externa, pode ser considerado os fatores sociais, políticos e econômicos, como por exemplo o fator Coronavírus, que será utilizado como caso de teste. Devido o fator externo poder contar com fatores diários distintos mas que podem impactar a um grupo de pessoa, tende-se a ser menos complexo a formulação de hipóteses, mas tem de ser bem analisado considerando que os estudantes podem ser afetados de diferentes formas pela mesma influência externa.

Ou seja, a influência interna e externa tem de ser analisada em conjunto para determinar um possível motivo do rendimento acima ou abaixo do planejado diante a proposta do projeto pedagógico.

Para Capellini, Tonelotto e Ciasca (2004), apesar de as dificuldades se apresentarem no aluno, apenas podem ser compreendidas quando se analisa o processo ensinar-aprender e todas as variáveis envolvidas, mas o fenômeno escolaridade é extremamente complexo, e

falhas ou problemas identificados neste período não podem ser atribuídos a uma causa única, geralmente são multideterminadas.

Obtendo um parecer sobre a influência de outros fatores no ensino acadêmico, pode-se iniciar o estudo para melhorar, ou até mesmo realizar o desenvolvimento de uma nova abordagem atual da instituição em prol de um melhor aproveitamento dos estudantes. Capellini, Tonelotto e Ciasca (2004) diz que a característica principal de uma dificuldade de aprendizagem é o baixo rendimento, ou desempenho, em atividades de leitura, escrita ou cálculo matemático apresentado por escolares em relação ao que se poderia esperar de acordo com sua inteligência e oportunidades, sendo este podendo ser composto por um ou mais fatores.

Sendo assim, pode se dizer que as teorias recentes de aprendizagem têm se preocupado com a interação entre o material a ser aprendido e os processos psicológicos necessários para aprender, enfatizando o estudo sobre o modo pelo qual o aprendiz obtém, seleciona, interpreta e transforma a informação (Boruchovitch, 1999).

Um fator que pode estar relacionado ao desempenho é o material abordado durante o ensino ou a forma que um conhecimento é compartilhado. Mesmo que os professores tendem a seguir certos padrões, normalmente estes possuem estilos próprios no momento de ensino.

Assim, Boruchovitch (1999) esclarece que é possível se prescrever métodos instrucionais para apoiar cada um desses processos que compõem a inteligência e enfatiza que a instrução precisa ir além de prover única e exclusivamente a informação e fornecer o apoio necessário para o processamento cognitivo, sendo a quantidade de apoio aos processos cognitivos a ser inserida na instrução uma variável chave para a efetividade da mesma. Com isso, poderá abordar as situações que afetam negativamente o aluno e tentar prover uma melhoria para que se adeque melhor o ensino aos estudantes.

Atualmente, cada professor, coordenador e gestor de curso da UFRN possui acesso ao SIGAA, onde nele pode-se consultar as informações das suas disciplinas do semestre em questão e dos semestres anteriores, onde com isso ele pode realizar uma comparação entre estes em relação ao desempenho. Entretanto, além de ser mais trabalhoso e necessitar de alto tempo dedicado, a análise é feita somente com base nas notas e frequência dos estudantes de suas disciplinas. Assim é de fundamental importância uma ferramenta mais completa que além de otimizar o tempo no momento de fazer uma análise, fosse possível ter uma abordagem mais simples e completa para realizar estas operações, trazendo benefício desde do discente até o coordenador de um curso.

O compartilhamento de experiência entre profissionais é uma prática que pode ser utilizada para analisar o que pode ser melhorado na prática de ensino ou até mesmo, promover uma evolução na forma de ensino e planejamento de cada um. Pensando nisso, uma trilha de aprendizagem pode ser de grande importância para otimizar o desempenho dos estudantes.

Com base em Lourenço (2020), podemos citar trilhas de aprendizagem como caminhos alternativos e flexíveis para promover o desenvolvimento das pessoas, onde aprendem trabalhando e trabalham treinando. Ou seja, elas são um conjunto de atividades com o objetivo de promover uma maior capacitação de profissionais e o desenvolvimento de novas competências, com base em estilo de ensino e habilidades pessoais em conjunto com as expectativas da organização. As trilhas possuem módulos de aprendizagem, que são os passos que vão ser completados durante a mesma.

Segundo Pearson (2021), as trilhas podem ser diferenciadas em 2 modelos de organização:

- Linear: o formato de aprendizagem é feito em sequência, assim para avançar para outro módulo, tem de ser finalizado o anterior para avançar para o próximo nível. Dessa forma, o criador da trilha cria um percurso a ser seguido.
- Agrupado: o formato de aprendizado não possui uma ordem definida, mas sim um número mínimo de passos que precisam ser completados para que se possa concluir a trilha. Ao contrário do formato linear, este o estudante define a ordem e os módulos de aprendizado, ganhando mais autonomia.

As trilhas de aprendizado também podem ser utilizadas para melhorar o currículo acadêmico, tomando como base o desempenho dos estudantes nas disciplinas ao longo dos períodos que ela foi lecionada e analisando-os, podendo assim tentar identificar quais foram as dificuldades e os problemas que ocasionaram um mau desempenho para tentar melhorar para os outros semestres, não só referente a disciplina analisada, mas no curso como um todo.

Com isso, podemos considerar que a trilha de aprendizagem pode auxiliar o melhoramento do desempenho dos estudantes e aumentar a experiência dos envolvidos e assim auxiliá-los a evoluir em seus respectivos trabalhos, além de outras vantagens, como proporcionar um aprendizado mais inclusivo, torna o acesso ao conhecimento mais ágil e efetivo, promovem autonomia e apresenta meios alternativos de aprimoramento.

2.4. Visualização de dados

Realizando os procedimentos principais de análise de dados, desde a coleta de dados até o desenvolvimento de um produto, pode-se trabalhar para exibi-los de forma mais clara e de mais fácil interpretação. Essa apresentação visual dos dados é o que se denomina visualização de dados e auxilia na análise das informações. Uma boa forma de visualização facilita a análise e inicia-se a formulação de estratégias. A representação e apresentação de dados que exploram nossa habilidade de percepção visual com a finalidade de aumentar a compreensão (Ghidini, 2017).

Apresentar de forma mais legível e melhor forma visual para facilitar a compreensão das informações e das análises, apoiando as decisões baseadas nos dados. Segundo Zeviani (2019), “Visualizações dão a você respostas para questões que você não sabia que tinha.” Este é um dos pontos mais essenciais para a mineração de dados educacionais, pois quanto melhor e mais apropriada for a forma de os dados forem apresentados, melhores são as decisões por parte dos indivíduos responsáveis pelas análises.

Quando nos referimos a área de mineração de dados educacionais, os dados são separados em material e contexto de aprendizagem. A estrutura material são os termos referentes a habilidades, unidades e aulas, enquanto que no contexto de aprendizagem são os alunos, professores e escolas por exemplo. Segundo Costa, Bajer, Amorin e Marinho (2012), esta etapa tem duas motivações principais:

- Identificação - os dados são apresentados de forma que as pessoas possam identificar os padrões mais facilmente, que são difíceis de expressar formalmente;
- Classificação - a destilação de dados pode ser usada também para apoiar a modelos de predição. Neste caso, parte dos dados são exibidos para serem rotulados pelos usuários;

De acordo com Zeviani (2019), para determinar o tipo de visualização, é essencial verificar se a informação é de forma conceitual ou de acordo com os dados (orientada a dados) e o gráfico está declarando ou explorando algo. Para ser concebido de forma conceitual, tem de expor ideias e esquemas para simplificar ou transmitir algum conhecimento. Quando referente a orientada a dados, tem de ter foco em estatística e padrões para que com isso possa informar, ou esclarecer a informação. Além disso, tem de se verificar se o dado visual está de forma declarativa, ou seja, com foco em planejar ou estruturar algo para fazer afirmações; ou de forma exploratória, que é focado em prototipação e interação para fazer descobertas e definir hipóteses. Assim pode-se dizer que um sistema com o intuito

de melhorar o desempenho dos discentes seja de forma orientada a dados e exploratória.

A escolha desde as formas geométricas até às variáveis visuais, é de grande importância para que se possa ter a comunicação entre o gráfico e o leitor, para que o mesmo se processe a informação visual e maximizar a eficiência e efetividade da mesma.

Para que a visualização de dados tenha a melhor eficácia, é importante escolher um gráfico que melhor demonstra as informações que se deseja passar e para isso, é necessário verificar a escala de variáveis, o propósito da visualização (forma do gráfico) e a forma visual/estética do gráfico.

Em relação a escala, ela pode ser dividida entre qualitativa (níveis categóricos) e quantitativa (níveis métricos). Quando a escala possui níveis categóricos, ela pode ser nominal, onde consiste em dados que não tem ordenação ou ordinal, quando há presença de ordenação. Quando possui escala quantitativa, ela pode ser discreta, contendo dados de contagem e suporte a número inteiros ou pode ser contínua, onde dá suporte aos números reais.

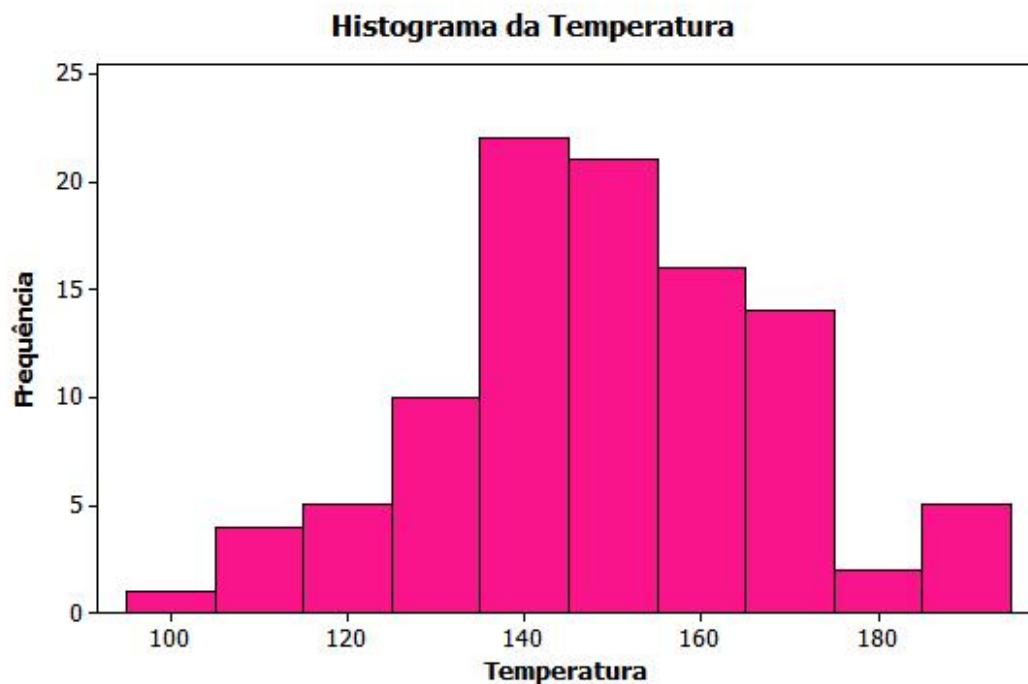
Em relação às funções de gráfico, relacionada aos dados a serem apresentados, podem ser divididas entre:

- Distribuição: quando a variável a ser apresentada possui uma distribuição de dados qualitativas ou quantitativas
- Relação: quando exibe a relação entre variáveis
- Ordenação: quando se precisa exibir os dados com base em uma classificação das informações
- Evolução: quando as informações possuem contexto cronológicos
- Mapa: quando os dados possui informações geográficas
- Organização: quando representam a estrutura de organização de coleções de elementos
- Fluxo: quando descrevem a conexão ou movimento entre elementos

Em relação a forma de um gráfico, onde se toma como base os aspectos estéticos, podem ser diferenciadas em diversos tipos. Algumas das principais formas são:

- Histograma: um gráfico do tipo histograma pode ser utilizado para descrever uma distribuição de dados por meio de barras. Como exemplo de um histograma, pode ser observado a figura 7.

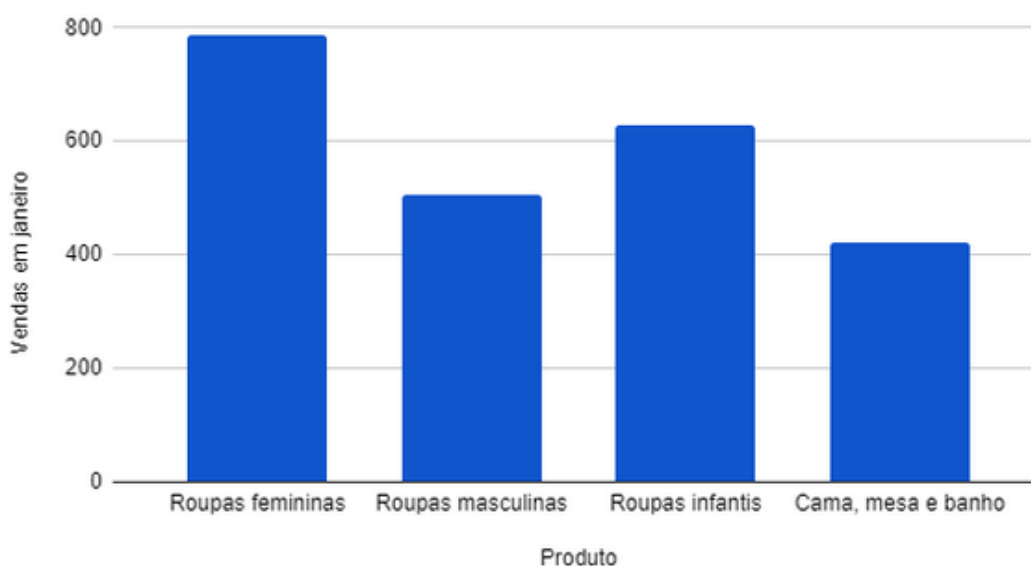
Figura 7 - Histograma



Fonte: lean six sigma brasil⁴

- Gráfico de barras: um gráfico de barras pode ser utilizado para exibir as frequências (absoluta ou relativa) de categorias. Como exemplo de um gráfico de barras, pode ser observado a figura 8.

Figura 8 - Gráfico de barras



Fonte: Escola Educacao - Como usar gráficos⁵

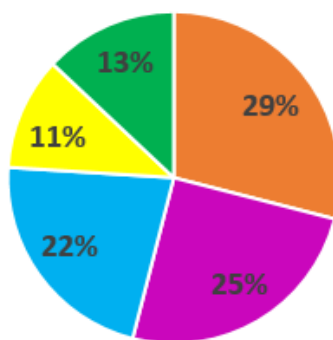
⁴ <https://leansixsigmabrasil.com.br/histograma/>

⁵ <https://escolaeducacao.com.br/como-usar-graficos/>

- Gráfico de setores: pode ser utilizado para exibir a frequência relativa de categorias. Como exemplo de um gráfico de setores, pode ser observado a figura 9.

Figura 9 - Gráfico de setores

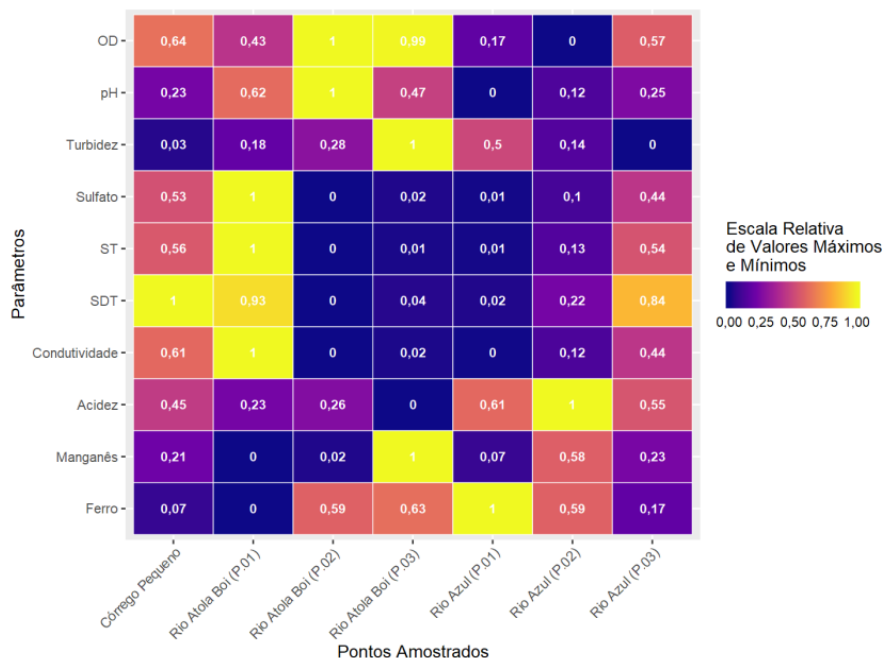
■ Segurança ■ Emprego ■ Saúde ■ Educação ■ Outros



Fonte: Escola Educacao - Gráfico de setores⁶

- Mapa de calor: pode ser utilizado quando tem como função comparar variável quantitativa entre categorias. Como exemplo de um mapa de calor, pode ser observado a figura 10.

Figura 10 - Mapa de calor



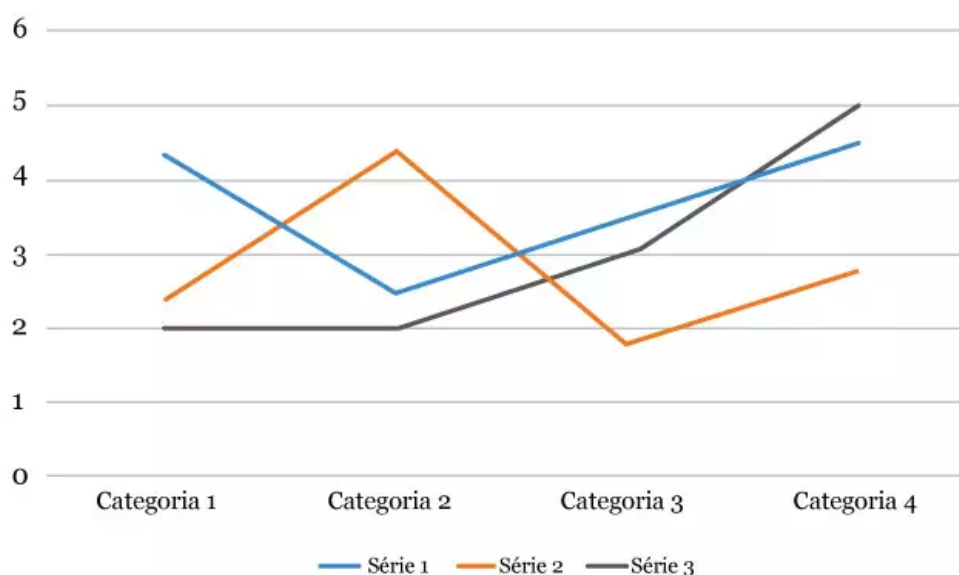
Fonte: 2 engenheiros - Mapa de calor⁷

⁶ <https://escolaeducacao.com.br/grafico-de-setores/>

⁷ <http://2engenheiros.com/2020/12/08/mapa-de-calor-r/>

- Série ou pontos conectados: pode ser utilizado para apresentar o percurso de parâmetro/variável ao longo do tempo, formando linhas ou pontos conectados. Como exemplo de uma série ou pontos conectados, pode ser observado a figura 11.

Figura 11 - Série ou pontos conectados



Fonte: Educa mais brasil - Tipos de Gráficos⁸

2.5. Mockup

Para um bom software ou website, é fundamental que se tenha um design em sua fase de planejamento, para que assim se possa ter uma visão de como o produto irá ficar, além de fornecer uma base à equipe responsável por desenvolver a parte visual. Tendo isso, foi construído um mockup para ser baseado ou até mesmo utilizado na íntegra.

Conforme citado por Rivero, Viana e Conte (2013), mockups são modelos de software, utilizados nas primeiras etapas do processo de desenvolvimento, onde imagens ou desenhos de como o software se veria após a sua implementação são preparadas.

Ou seja, um mockup é um exemplo gráfico a ser seguido no desenvolvimento de uma ferramenta, expondo de maneira mais elaborada a proposta visual do produto, pensando em como organizar os itens para melhor visualização e melhor usabilidade do usuário, trazendo mais facilidade no uso e melhor organização nos itens da tela.

⁸ <https://www.educamaisbrasil.com.br/enem/matematica/tipos-de-graficos>

Segundo Kawakami, Rivero e Conte (2014), o método de avaliação mais utilizado para identificar problemas de usabilidade em aplicações Web móveis é o teste de usabilidade, que foca na interação direta dos usuários com o sistema. No entanto, o custo deste tipo de avaliação é alto devido aos equipamentos e laboratórios necessários para sua aplicação.

Complementando, Kawakami, Rivero e Conte (2014) dizem também que existem duas categorias de métodos de avaliação de usabilidade (*Usability Evaluation Methods* - UEMs), testes com usuários e inspeções com especialistas. UEMs são procedimentos compostos por um conjunto definido de atividades que são usados para avaliar a usabilidade de uma aplicação.

Os testes com usuário são análises focadas no usuário, onde se obtém as respostas a partir de perguntas diretas como por exemplo questionários para avaliar a usabilidade do produto. Entretanto, Kawakami, Rivero e Conte (2014) afirmam que ele possui um custo elevado, devido ao fato que é necessário realizar o teste com diferentes perfis de usuários. Com isso, é importante considerar que o teste de usabilidade só pode ser aplicado quando já existe uma parte inicial do sistema pronta.

Ou seja, além do custo elevado, é necessário que parte da aplicação esteja pronta, para que assim o usuário possa utilizar e opinar como pode ser melhorada a ferramenta. Para auxiliar, é válido citar o uso do conceito de co-experiência, um conceito que orienta as atividades de observação e interpretação, mas sem ditar o resultado destas. Segundo Silva (2015), a co-experiência pode ajudar a estabelecer observações e identificar interpretações nas descobertas, especialmente quando o foco é amplo e difuso, como nos estágios iniciais do design do produto.

As inspeções com analistas são aquelas que a equipe de desenvolvimento e especialista em usabilidade de usuário analisam o produto formado com base em regras de usabilidade para assim tentar atribuir melhorias ao uso do produto.

Com a estrutura definida do mockup, ele pode ser utilizado para visualizar em escala real, como ficará o visual do produto, além de seu formato e outros aspectos semelhantes. Inclusive poderá trazer aos envolvidos como um usuário a plataforma e assim pensar na melhor estratégia para fornecer a melhor experiência a quem a está usando.

A preocupação em como usuário utilizará a tela e a melhor organização para facilitar o uso são alguns dos objetivos do estudo relacionado ao User Experience (UX). Com isso podemos definir como UX uma disciplina que avalia a experiência do usuário em um produto para que seja proposto ajustes e melhores com o propósito de estimular a melhoria do produto

e assim encantar e satisfazer o usuário, com quanto mais for prazeroso o uso do usuário, mais provável será a utilização do produto, assim conquistando a fidelização do mesmo.

Para isto, pode-se definir 3 pontos principais para que esse objetivo seja alcançado:

- Utilidade: Avalia o quanto o produto pode ser útil para o cliente, ou seja, o quanto ele prefere o seu produto quando comparado a outros;
- Facilidade de uso: Avalia o tempo e a facilidade que usuário utiliza para resolver algum problema e fazer suas análises com sucesso;
- Prazer: O quão prazeroso (divertido, interessante, recompensador, etc.) é usar o serviço e não as alternativas.

Além desses itens, é possível ajustar de maneira mais prática, sem necessariamente ir no ambiente de desenvolvimento para fazer estes ajustes, otimizando o tempo gasto para fazer estas correções e menos custo. Assim foi idealizada uma representação gráfica, podendo assim ser utilizada como um todo ou como uma base para formular o design a ser incorporado.

3. METODOLOGIA DE PESQUISA

A proposta deste trabalho é aplicar um estudo da análise a partir dos dados fornecidos pelo SIGAA para complementar as formas de análise de desempenho dos estudantes de graduação e pós graduação.

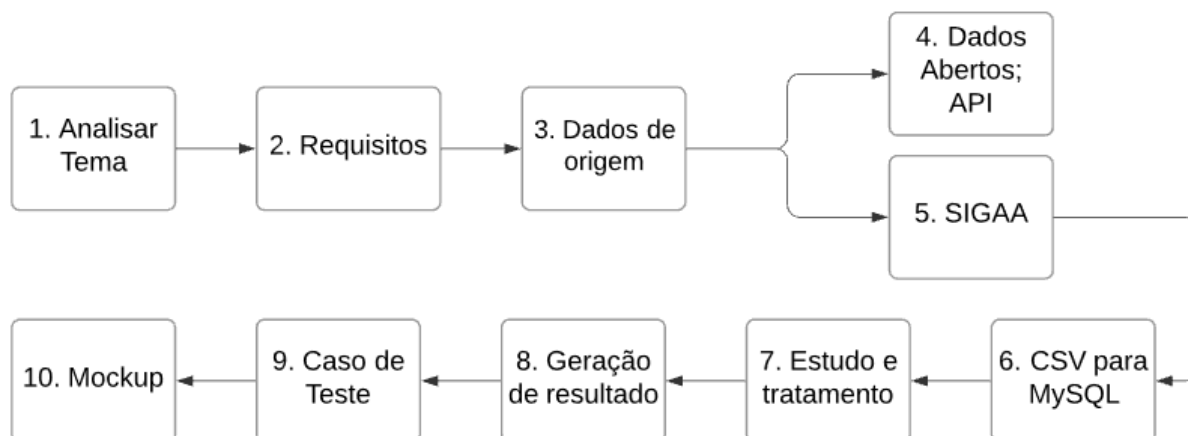
Os estudos são do tipo descritivo-analítico, pois serão realizados com base em análise dos dados utilizando métricas e técnicas estatísticas para entender as informações, podendo assim averiguar pelos dados e visualização gráfica obtida a performance dos alunos e, como estamos utilizando como caso de teste o período de pandemia, verificar o impacto causado por este quando comparado aos outros semestres estudantis.

3.1. Fluxo de Atividades

Para a produção deste trabalho de conclusão de curso, as atividades foram feitas de acordo com o seguinte fluxo apresentado na figura 12. Para melhor explicação, tem-se a legenda a seguir:

1. Caracterizar o cenário atual da UFRN, identificando um problema a ser resolvido;
2. Identificar os requisitos e necessidades para esta proposta;
3. Dados de origem, considerando a melhor fonte com base na qualidade das informações;
4. Análise dos dados abertos e das informações fornecidas pela API;
5. Solicitação à equipe do SIGAA os dados referente às turmas, disciplinas, matrículas dos discentes e dados socioeconômicos;
6. Importação do CSV (Comma-Separated-Values) para o banco MySQL;
7. Estudo e tratamento das informações;
8. Estruturação de buscas em SQL e geração de ferramentas visuais (gráficos e tabela);
9. Avaliação das informações obtidas com base no caso de teste;
10. Construção do Mockup;

Figura 12 - Fluxo de atividades



Fonte: Autoria própria

3.2. Ferramentas utilizadas

Para realizar a análise de dados, foi utilizado a linguagem de programação Python⁹, por ser uma das linguagens mais utilizadas na área de ciências de dados pela variedade de bibliotecas que ajudam na transformação, análise e visualização das informações e sua escalabilidade.

Com isso, foi realizada a migração dos dados fornecidos pela Superintendência de Tecnologia da Informação (STI) para o banco de dados MySQL¹⁰, realizando a conexão a partir do Jupyter Notebook¹¹. Além disso, foi idealizado um mockup para que se possa se basear no desenvolvimento de um sítio web ou aplicativo para qualquer computador, tablet ou smartphone, para assim qualquer pessoa tenha a possibilidade de consultar.

3.2.1. MySQL

Para armazenar essas informações, por motivos de desempenho, foi realizada a escolha de guardar estas em um banco de dados, onde é uma forma de salvar dados eletronicamente em um sistema de computador. Dentre as opções foi escolhido o MySQL, por ser gratuito e de fácil utilização.

O MySQL é um Sistema de Gerenciamento de Banco de Dados (SGBD), que utiliza como interface a Linguagem de Consulta Estruturada (*Structured Query Language - SQL*).

⁹ <https://www.python.org/>

¹⁰ <https://www.mysql.com/>

¹¹ <https://jupyter.org/>

Devido a sua compatibilidade, estabilidade e desempenho, foi escolhido para ser o sistema onde armazena os dados deste trabalho, em conjunto com o MySQL Workbench, onde é uma ferramenta visual de design de banco de dados que provem modelagem de dados, desenvolvimento SQL dentre outras funcionalidades.

3.2.2. Python

Uma linguagem de programação é um conjunto de instruções e símbolos, composto por regras semânticas e sintáticas, para implementação de um código fonte, e assim ser interpretado para ser executado pelo sistema. Com isso, a linguagem escolhida para a análise foi o Python.

A linguagem Python é bastante utilizada no que se refere à análise e processamento de dados. Ela possui uma tipagem forte e dinâmica, e sendo de alto nível, possui uma sintaxe mais simplificada. Alguns dos fatores que a tornam bastante utilizada é devido ao seu número crescente de bibliotecas relacionadas à análise de dados, a sua escalabilidade e suas bibliotecas de visualização.

3.2.3. Anaconda¹²

Para auxiliar na análise, facilitando assim a obtenção visual das informações e tratamento dos dados, foi utilizado o Anaconda. Esta é uma das plataformas de ciência de dados mais conceituadas, contendo a distribuição da linguagem de programação Python e R, visando simplificar o uso de pacotes e a produção de análises.

O ambiente de desenvolvimento integrado escolhido foi o Jupyter Notebook, onde é um aplicativo web open-source, assim podendo ser utilizado para diversos fins. Dentre os possíveis usos, estão a limpeza de dados, visualização de gráficos e até textos e imagens explicativas.

3.2.4. Bibliotecas

Uma biblioteca Python pode ser considerada como um conjunto de scripts acessados por um programa Python para uma determinada função, assim facilitando o desenvolvimento

¹² <https://www.anaconda.com/>

da aplicação com seus recursos. Para realizar o uso destas, pode ser feita a importação no início do código de desenvolvimento.

Uma vez os dados disponíveis para serem trabalhados, duas bibliotecas principais foram utilizadas para tornar a análise mais simples, sendo elas o Pandas¹³ e o Matplotlib¹⁴.

Pandas é uma biblioteca escrita em Python para manipulação e análise de dados. Devido a sua facilidade de uso e a alta quantidade de ferramentas, é uma das bibliotecas mais usadas na área de ciência de dados.

Em conjunto com a Matplotlib, uma biblioteca para a construção de gráficos e visualização de dados, é possível ler, manipular e plotar dados para uma melhor análise das informações.

3.3. Convenções utilizadas

Para a análise de dados e o desenvolvimento da ferramenta de análise foram realizadas algumas convenções para realizar um agrupamento de dados do formato hierárquico divisivo.

Os semestres que foram selecionados são 2016.1, 2016.2, 2016.3, 2016.4, 2017.1, 2017.2, 2017.4, 2018.1, 2018.2, 2018.4, 2019.1, 2019.2, 2019.4, 2020.1, 2020.2, 2020.5 e 2020.6; sendo os níveis de ensino a Graduação e Stricto Sensu

Em relação a situação do discente na matrícula, foi considerado como de aprovação as que possuíam: ‘Aprovado’, ‘Aprovado por Nota’. Como situações de reprovação foi considerado ‘Reprovado’, ‘Reprovado por Nota’, ‘Reprovado por Faltas’, ‘Reprovado por Média e Faltas’ e ‘Reprovado por Nota e Falta’. E como situação de trancado e desistência, foi considerado ‘Trancado’ e ‘Desistência’, respectivamente.

3.4. Dados

Um conjunto de dados de uma instituição é uma fonte de grande riqueza de informação e importância devido a sua importância como sua quantidade.

Observando os dados disponíveis pelo Portal de Dados Abertos¹⁵, cujo tem como princípio a disponibilização dos dados referentes à instituição para toda a comunidade (regulamentado pela Lei de Acesso à Informação Pública - Lei nº 12.527/2011), foi verificado

¹³ <https://pandas.pydata.org/>

¹⁴ <https://matplotlib.org/>

¹⁵ <http://dados.ufrn.br/>

que algumas informações necessárias ainda não tinham sido disponibilizadas, com isso foi requisitado à Superintendência de Tecnologia da Informação (STI) os dados atualizados, em estrutura semelhante ao padrão dos dados abertos, do banco de dados do Sistema Integrado de Gestão de Atividades Acadêmicas, desenvolvido pela STI.

Com isso a análise e desenvolvimento da ferramenta foram disponibilizados em formato de CSV, um dos formatos mais utilizados para armazenar e compartilhar dados. Os arquivos continham as seguintes informações:

Referente a turmas: `id_turma`, `codigo`, `id_disciplina`, `nivel_ensino`, `campus`, `local`, `ano`, `periodo`, `data_inicio`, `data_fim`, `descricao_horario`, `total_solicitacoes`, `capacidade_aluno`, `tipo`, `distancia`, `data_consolidacao`, `agrupadora`, `id_turma_agrupadora`, `qtd_aulas_lancadas`, `situacao`, `id_convenio` e modalidade.

Referente as matrículas dos discentes: `id_discente`, `id_turma`, `descricao_horario`, `ano`, `periodo`, `id_disciplina`, `id_curso`, `unidade`, `nota`, `reposicao`, `faltas`, `media_final`, `numero_faltas`, `descricao`.

Referente a disciplinas: `id_disciplina`, `tipo_componente`, `codigo`, `nivel_ensino`, `nome`, `unidade`, `ch_teorico`, `ch_pratico`, `ch_estagio`, `ch_total`, `ch_dedicada_docente`, `ch_ead`, `cr_ead`, `permite_horario_flexivel`, `num_unidades`, `equivalencia`, `pre_requisito`, `co_requisito`, `ementa`, `bibliografia` e modalidade.

Referente aos dados socioeconômicos: `id_discente`, `ano`, `periodo`, `id_curso`, `renda_familiar_bruta_mensal`, `cidade`, `uf`, `segundo_grau_tipo_escola`, `possui_bolsa_pesquisa`, `possui_auxilio_alimentacao`, `possui_auxilio_transporte` e `possui_auxilio_residencia_moradia`.

Sendo o `id_disciplina`, `id_turma`, `id_disciplina`, `id_curso` e `campus` colunas com as informações criptografadas.

Como forma de consultar os dados, também foi realizada a tentativa de utilizar a interface de programação de aplicação (API - *Application Programming Interface*), onde consiste em um conjunto de regras e padrões de programação que funcionam como uma interface de acesso a dados. Entretanto, por haver a limitação de visualização, foi definido a utilização dos dados compartilhados pela STI para este trabalho.

4. DESENVOLVIMENTO E RESULTADOS

O objetivo deste projeto é propor quais os possíveis conteúdos que poderia-se ter em uma ferramenta de análise acadêmica e quais os seus benefícios, além de utilizar como caso de teste o período em pandemia. A seguir serão mostrados os gráficos para exemplificar possíveis análises que poderiam ser feitas com essa ferramenta.

4.1. Requisição e armazenamento das informações

A primeira etapa de desenvolvimento foi analisar se os dados disponíveis no portal de dados abertos atenderia a pesquisa, entretanto foi observado que os mesmos estavam desatualizados em relação às informações dos períodos em pandemia e por este motivo não foi utilizado este modo. Além do modo de dados abertos, foi realizada a utilização da API de sistemas da UFRN, através das credenciais de acesso fornecidas após o cadastro¹⁶ (que consiste em um formulário para inserir as informações referentes ao projeto, seus participantes e quais recursos iriam ser utilizados), analisando os dados obtidos por esta, foi percebido que as informações eram limitadas, não sendo possível obter os dados de todas as turmas e estudantes necessários para o estudo.

Com isso, foi realizada uma requisição dos dados à Superintendência de Tecnologia da Informação (STI) da UFRN, contendo as informações necessárias relacionadas às turmas, disciplinas, dados socioeconômicos e matrículas. Como resultado foram fornecidos arquivos no formato CSV (*Comma-Separated Values*) contendo as informações.

Para armazenar essas informações foi escolhido o MySQL. Para este processo, foram analisadas as informações contidas nos documentos recebidos e estruturadas as colunas das tabelas. Foram construídos tabelas referente às informações das turmas, dados socioeconômicos, disciplinas e matrículas com o código em SQL, com tipos de coluna int (tipo onde permite números inteiros), varchar (permite o registro de letras, palavras e textos), float (armazena número que ontem casa decimal), tinyint (uma forma que será usada para armazenar tipos booleanos) e text (armazena textos).

Com os dados sendo salvos no banco para serem consultados, foi utilizado o anaconda em conjunto com o Jupyter Notebook para realizar as consultas e a estruturação dos dados.

A consulta seguia no padrão do script exibido na figura 13, onde com a utilização do pandas, foi realizada uma consulta com um código SQL (`Sql_Busca`) para gerar um data

¹⁶ <https://api.ufrn.br/documentacao.html#cadastro>

frame (DataFrame) para ser trabalho. Caso seja necessário filtrar o resultado obtido com base em alguns filtros, foi estruturado data frames distintos (DataFrame_Filtro1, DataFrame_Filtro2) com cada filtro com base no data frame base e feito merge com base em uma coluna comum entre ambos (DataFrame_Merge). Após isso, iniciou a estruturação para o desenvolvimento do gráfico, com base no data frame resultante, selecionado o estilo de gráfico em linha e configurando o parâmetro que será atribuído ao eixo X (Coluna_Eixo_X) e ao eixo Y (Coluna_Eixo_Y), além do tamanho da figura, a presença da legenda e simbolizando que a legenda não terá rotação; em conjunto com esses procedimentos, atribuído às legendas na melhor localização. Após isso, para concluir foi definido o nome do eixo X (Legenda_Eixo_X) e do eixo Y (Legenda_Eixo_Y), além do salvar a imagem referente ao gráfico.

Figura 13 - Script para consulta e gráfico

```
DataFrame = pd.read_sql_query(f"Sql_Busca", engine)

DataFrame_Filtro1 = DataFrame.query("Parametro_1 == '" + Filtro_1 + "'")
DataFrame_Filtro2 = DataFrame.query("Parametro_2 == '" + Filtro_2 + "'")

DataFrame_Merge = pd.merge(DataFrame_Filtro1, DataFrame_Filtro2, how='outer', on=['Coluna_Merge'])

Grafico_DataFrame = DataFrame_Merge.plot.line(x='Coluna_Eixo_X', y='Coluna_Eixo_Y', figsize=(10,6),
                                             legend=True, rot=0)
Grafico_DataFrame.legend([Legendas_Graficos], loc='best')

plt.ylabel('Legenda_Eixo_Y')
plt.xlabel('Legenda_Eixo_X')
plt.savefig('Caminho_Nome_Arquivo', bbox_inches='tight')
```

Fonte: Autoria própria

4.2. Gráficos e informações obtidas

Inicialmente foram realizadas as consultas de desempenho com 2 disciplinas de graduação (escolhidas com base em poderem ser consideradas complexas, além de apresentarem resultados diferentes nas análises realizadas) e 1 de stricto (escolhido por apresentar dados completos e resultado da análise ser diferente das disciplinas de graduação) para que fosse possível ter uma base do que o sistema poderia fornecer.

As disciplinas selecionadas e os professores foram trocados por nomes distintos para a análise:

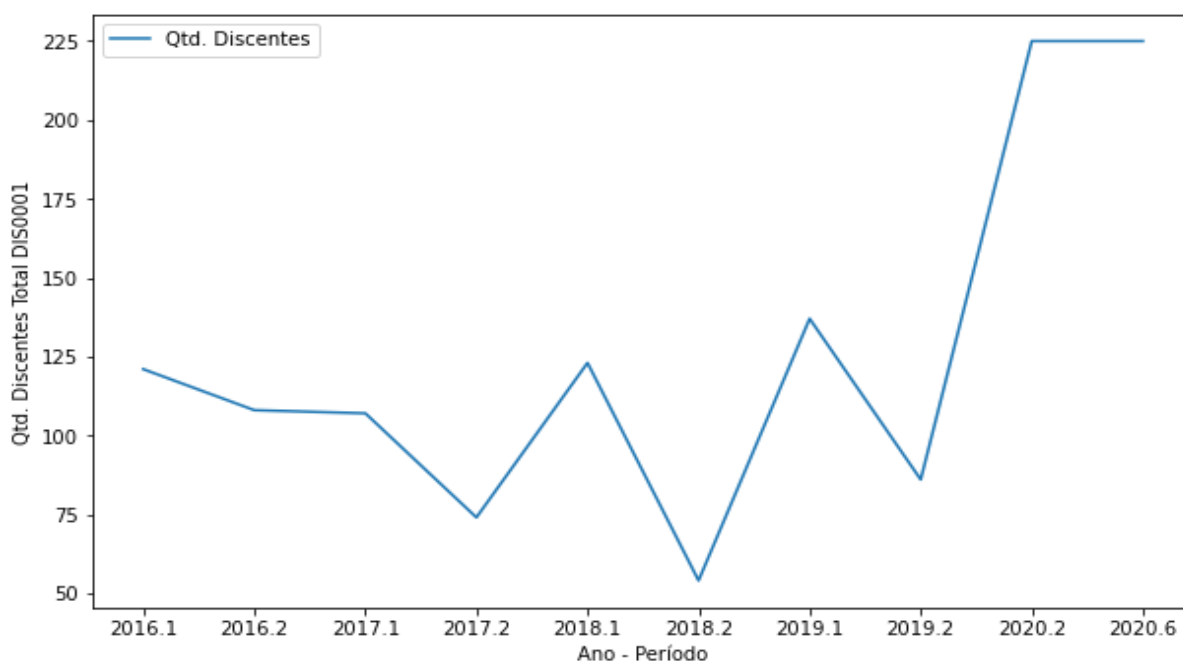
- DIS0001 - Disciplina de graduação I
- DIS0002 - Disciplina de graduação II (onde consiste em uma disciplina equivalente a DIS0001, entretanto sendo fornecida por outra unidade)
- DIS0003 - Disciplina de Pós graduação I

Antes de exibir o resultado, é válido salientar alguns pontos:

1. Removemos os dados das disciplinas de férias devido a quantidade de discentes ser consideravelmente abaixo das demais e não ser possível uma análise mais detalhada em relação ao nosso caso de teste, que é o período de pandemia.
2. Algumas disciplinas apresentam ocorrências em todos os períodos letivos, entretanto algumas outras possuem ocorrência em alguns períodos específicos, como por exemplo, só no primeiro semestre do ano.
3. O caso de teste a ser considerado será o período de pandemia, onde as disciplinas firmam a distância, para que se possa verificar o impacto no desempenho dos estudantes.

Com isso, as informações que obtivemos com base a disciplina DIS0001 referente a quantidade de discentes gerais, ou seja, aqueles que realizaram a inscrição na disciplina, pode ser observado na figura 14, onde verifica-se que nos semestres 2020.2 e 2020.6 que foi lecionado a disciplina DIS0001, a quantidade de matrícula foi muito acima dos semestres anteriores.

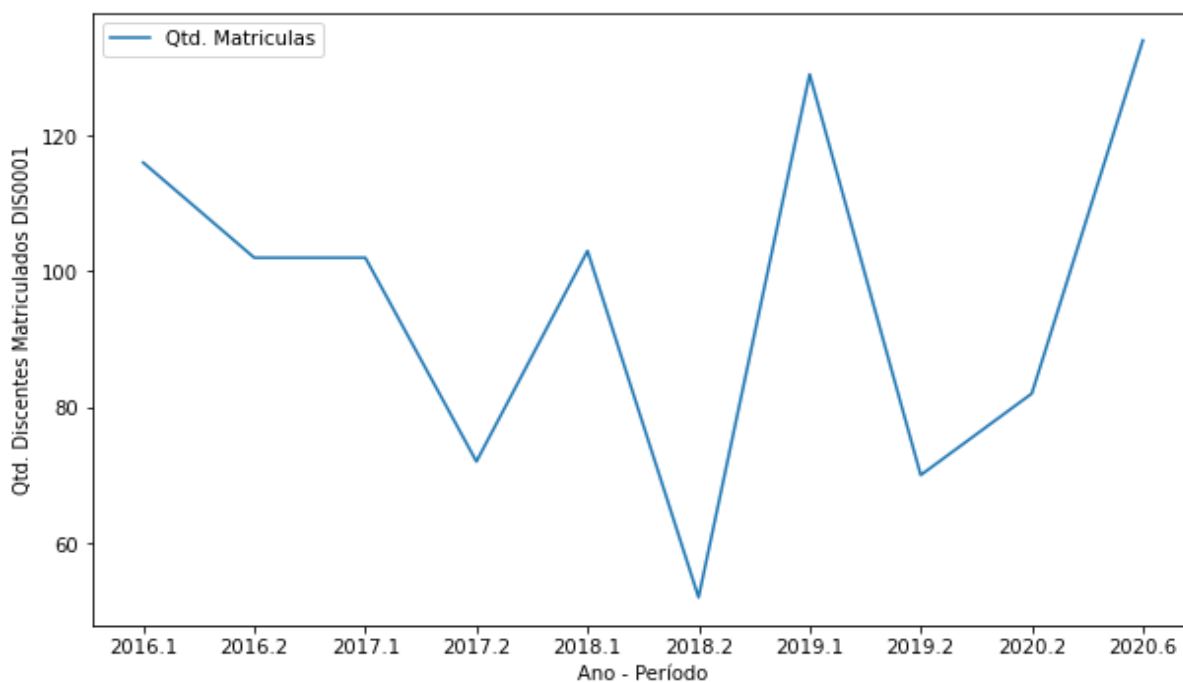
Figura 14 - Quantidade geral de discentes em DIS0001



Fonte: Autoria própria

Entretanto, podemos observar na figura 15 que a quantidade de discentes com matrículas efetivadas no semestre de 2020.2 foi abaixo do semestre 2020.6, o qual esta foi semelhante ao semestre 2019.1.

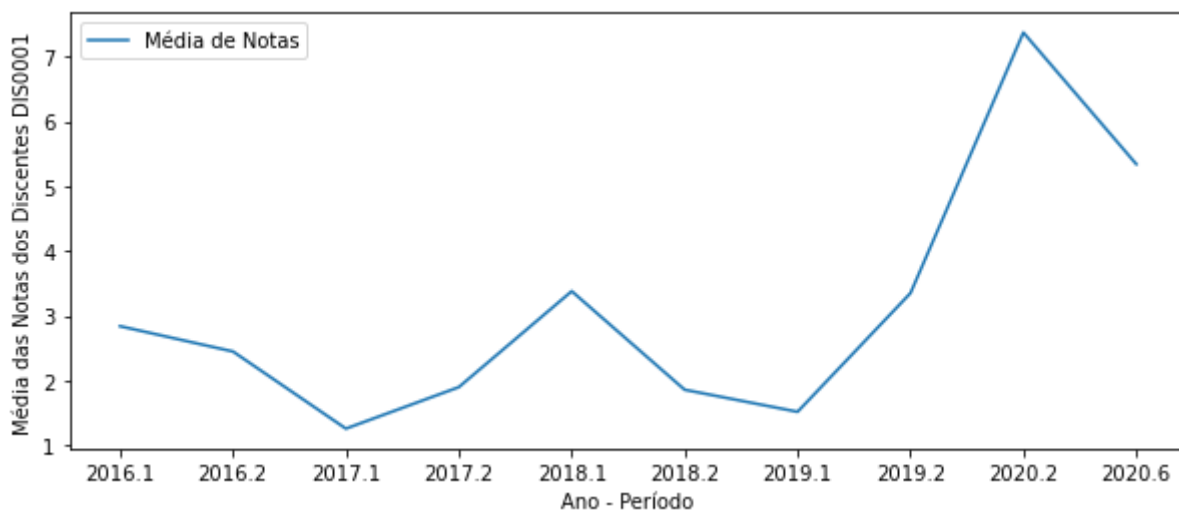
Figura 15 - Quantidade de discentes em DIS0001



Fonte: Autoria própria

Para verificarmos o desempenho dos estudantes de forma mais geral, assim obtivemos a média das notas e a média das faltas em cada período. Na figura 16 é exibida as informações referente a notas, onde se pode observar que possui uma diferença considerável em relação a nota média nos semestres de pandemia, sendo bem acima dos demais semestres.

Figura 16 - Média das notas em DIS0001

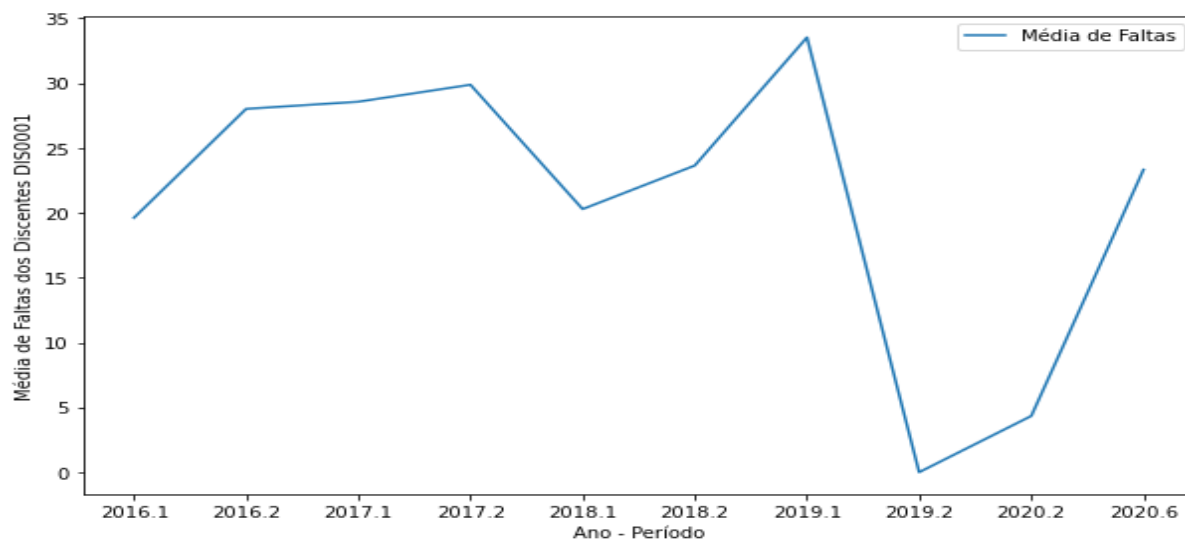


Fonte: Autoria própria

Na figura 17, podemos observar as informações referente a frequência dos estudantes (quantidade média de faltas) onde podemos citar como destaque o semestre 2020.2, onde a foi

bem abaixo dos demais semestres, entretanto no semestre 2020.6 ela nos conformes dos outros semestres, sem apresentar grande variância.

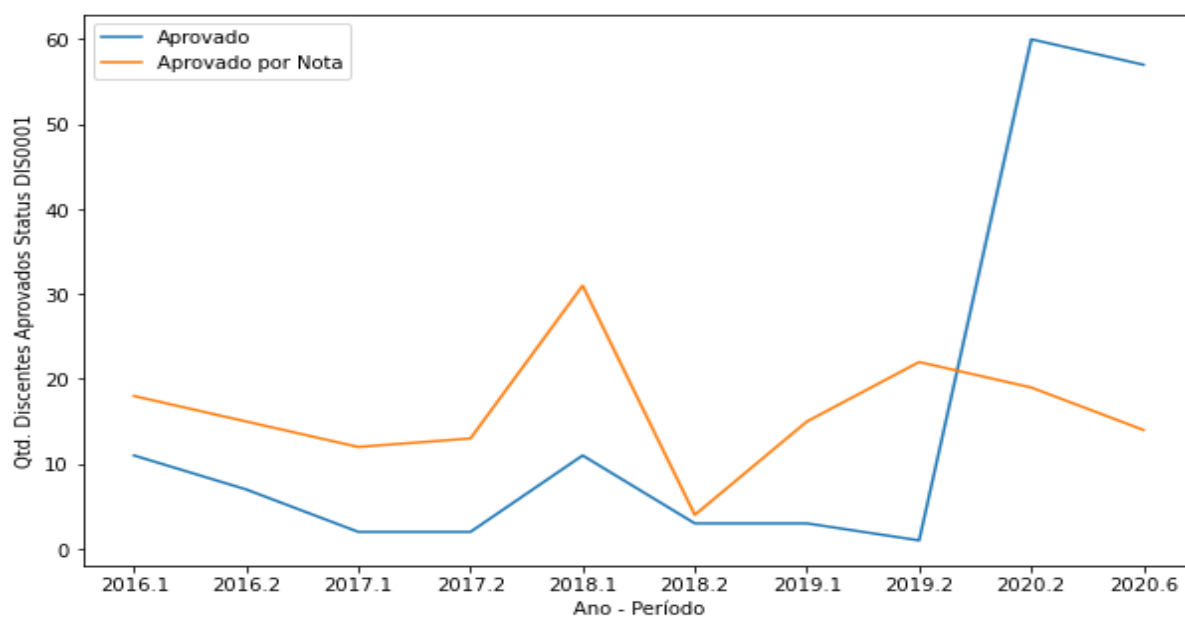
Figura 17 - Média das faltas em DIS0001



Fonte: Autoria própria

Para melhor avaliar os dados de aprovações dos estudantes, podemos observar na figura 18 os dados referentes aos discentes que obtiveram sucesso na disciplina, o qual os semestres em pandemia obtiveram índices elevados de aprovações com status 'aprovado' quando comparados aos demais.

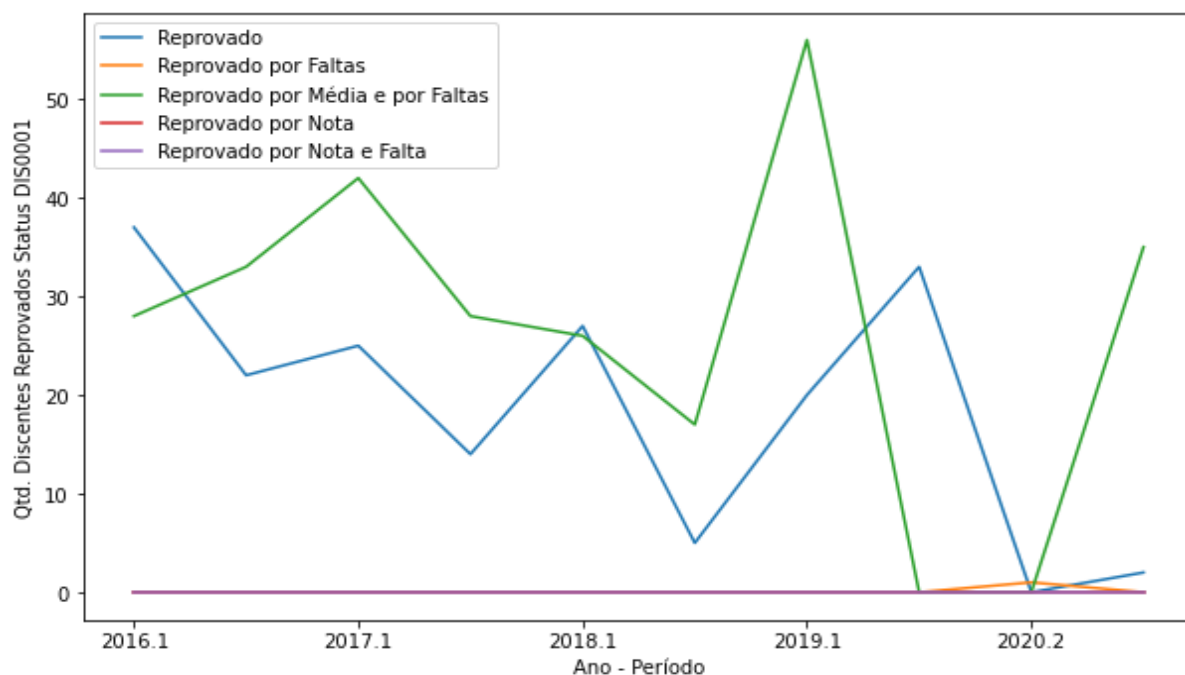
Figura 18 - Quantidade de aprovações por status em DIS0001



Fonte: Autoria própria

Em relação aos dados de reprovações, também houve uma melhora nos períodos em pandemia como pode ser observado na figura 19.

Figura 19 - Quantidade de reprovações por status em DIS0001

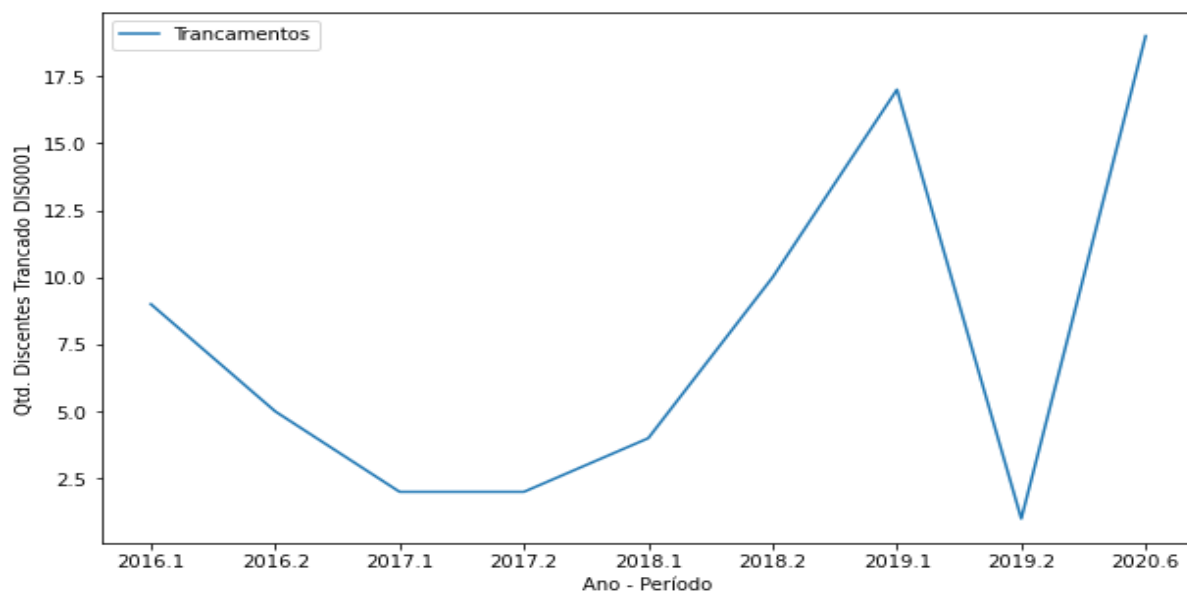


Fonte: Autoria própria

Além dessas informações, é válido apresentar também os dados referentes a discentes com status de trancado e desistência, salientando que no momento que foi iniciado o tratamento dos dados, ainda não se tinha discentes com status de desistência ou trancado.

Na figura 20 é exibido as informações referente aos tratamentos, onde pode-se observar que o semestre 2020.6 apresentou um nível elevado de disciplinas com este status, tendo ocorrência semelhante somente ao semestre 2019.1.

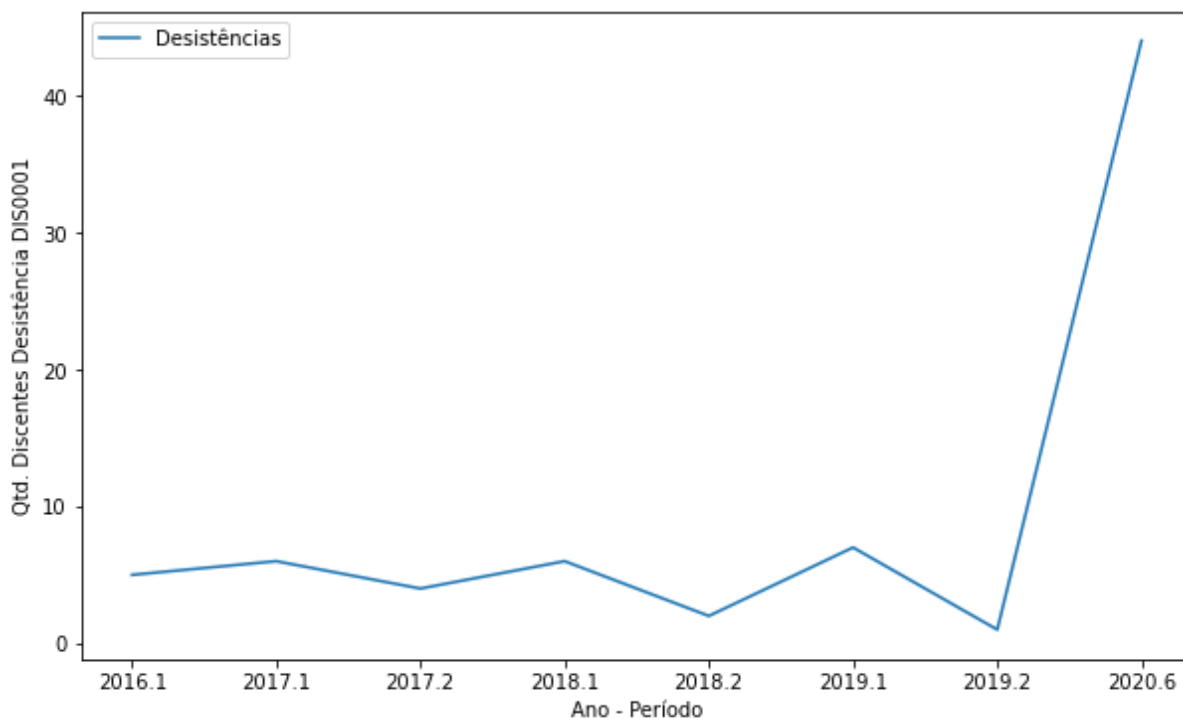
Figura 20 - Quantidade de trancamentos em DIS0001



Fonte: Autoria própria

Na figura 21 é exibido as informações referente as desistências, onde podemos observar que o semestre 2020.6 teve altas quantidade também, assim como de trancamento, estando acima de todos os demais.

Figura 21 - Quantidade de desistências em DIS0001

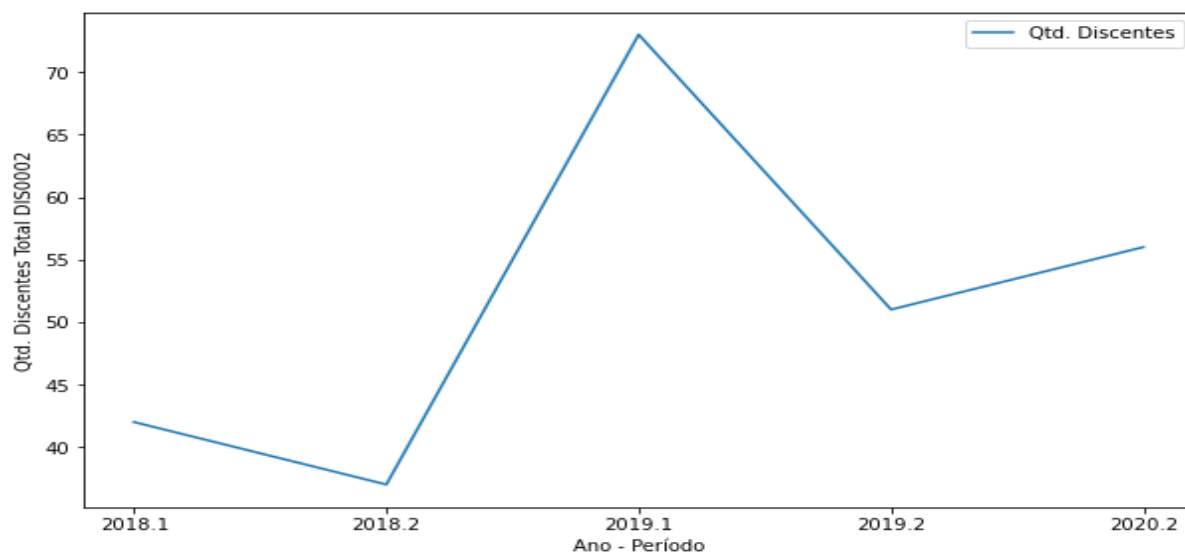


Fonte: Autoria própria

Após estruturamos as informações para a disciplina DIS0001, realizamos a consulta com base na disciplina DIS0002, onde consiste em uma disciplina equivalente, entretanto sendo fornecida por outra unidade.

Como podemos observar na figura 22, a quantidade de matrículas não teve uma diferença considerável referente ao nível de pandemia e os demais como houve com a disciplina DIS0001, a não no semestre não pandêmico de 2019.1, onde apresentou um pico diferente dos demais. Além disso, podemos perceber que a disciplina não foi ofertada no semestre 2020.6, diferente da situação da disciplina equivalente a ela.

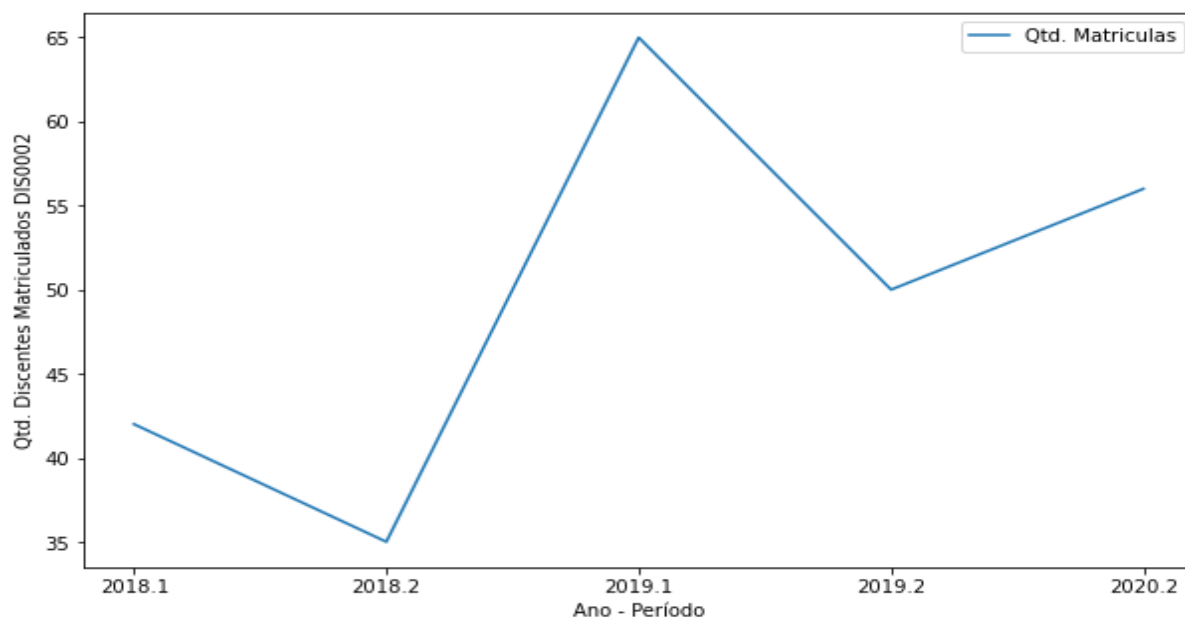
Figura 22 - Quantidade de discentes gerais em DIS0002



Fonte: Autoria própria

O gráfico de matrículas efetivas pode ser observado na figura 23, o qual não apresentou tanta divergência referente ao gráfico de matrículas, assim podemos perceber que o semestre 2020.2 teve um nível de matrículas a mais dos demais, exceto em 2019.1, quando comparado aos demais semestres.

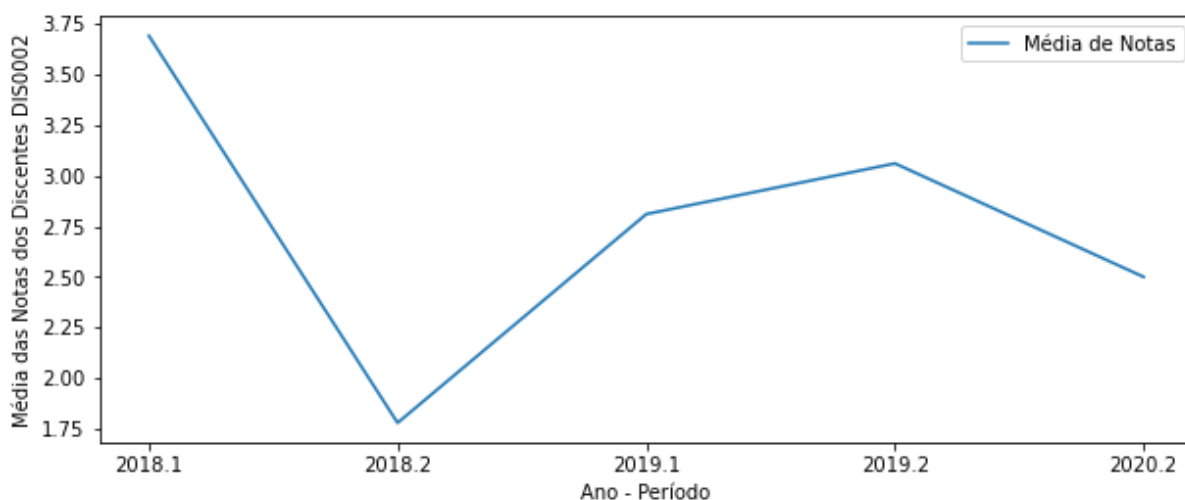
Figura 23 - Quantidade de discentes em DIS0002



Fonte: Autoria própria

Em relação a média das notas, podemos observar na figura 24 que no semestre em pandemia 2020.2 ele apresentou uma queda quando comparado aos outros semestres (exceto 2018.2).

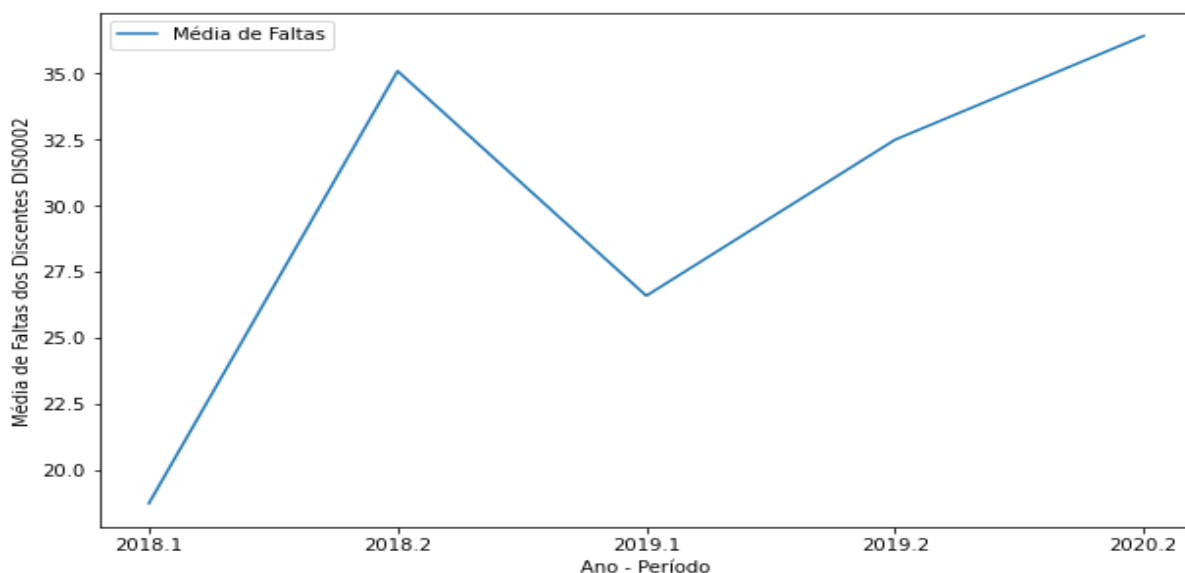
Figura 24 - Média de notas em DIS0002



Fonte: Autoria própria

Em relação à média de faltas, podemos perceber alto nível de faltas, assim tendo uma perda de desempenho também, como pode ser visto na figura 25. Com isso, podemos observar que a média da quantidade de faltas em 2020.2 (semestre em pandemia) foi o mais alto dentre todos os demais.

Figura 25 - Média de faltas em DIS0002

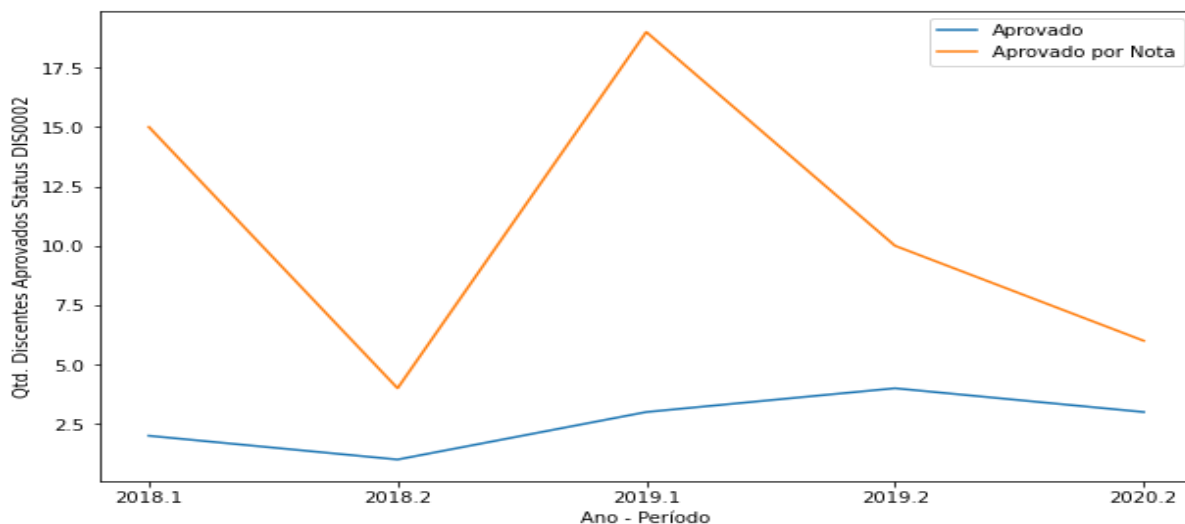


Fonte: Autoria própria

Partindo para verificar os discentes com status referente a aprovações, podemos observar o gráfico da figura 26. Nele consta que a quantidade de discentes aprovados no semestre em pandemia, foi inferior a quase todos os outros semestres. Se olharmos para os discentes com status 'aprovado por nota', podemos perceber a diminuição mais acentuada de

discentes aprovados com este status; e discentes com status aprovado também teve diminuição, mas em nível mais suave comparado ao status anterior.

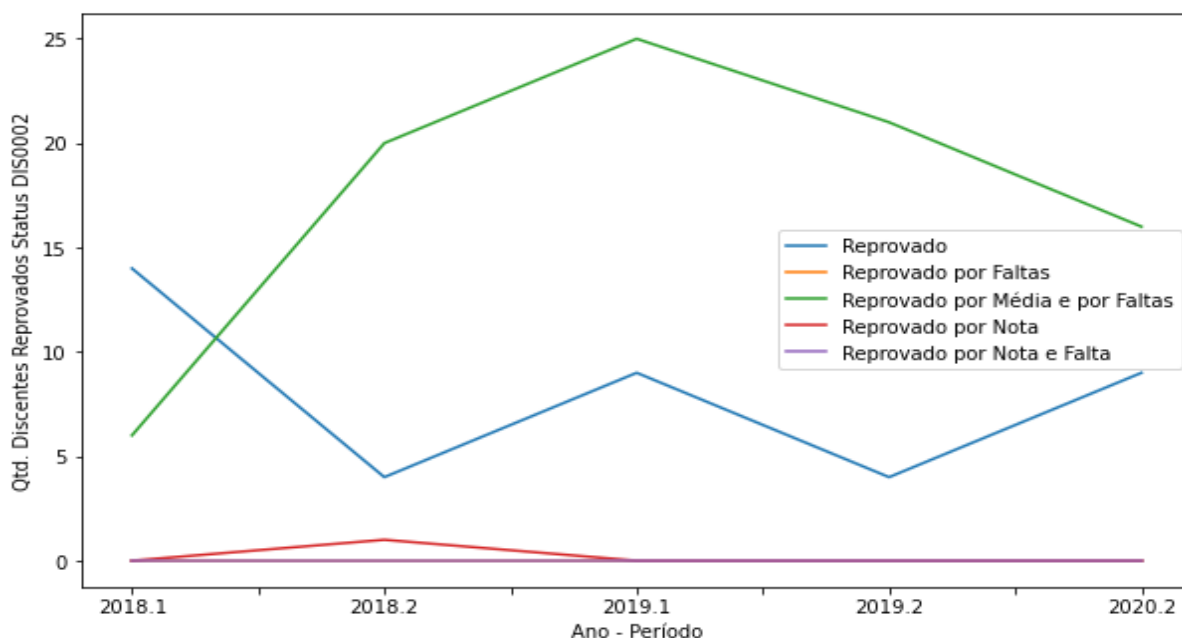
Figura 26 - Quantidade de aprovações por status em DIS0002



Fonte: Autoria própria

Em relação aos discentes com status reprovados, podemos observar a figura 27, onde podemos observar que a quantidade de reprovações, exceto no semestre, 2019.1 se manteve a níveis próximos.

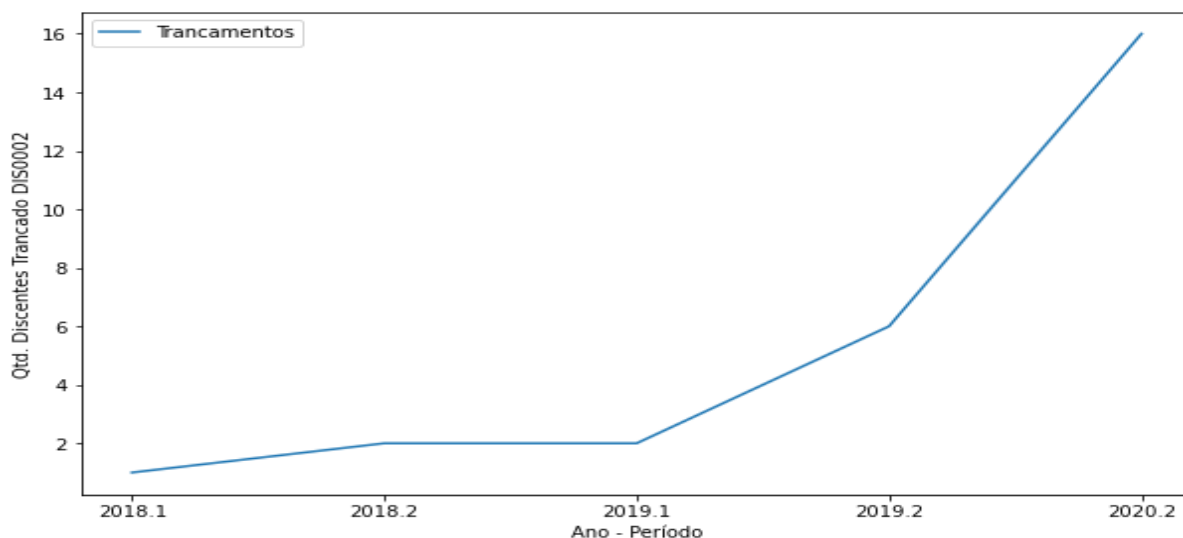
Figura 27 - Quantidade de reprovações por status em DIS0002



Fonte: Autoria própria

Para a disciplina DIS0002, é importante citar o gráfico de disciplinas trancadas na figura 28, onde podemos perceber que no período de pandemia, a quantidade de discentes que solicitaram trancamento foi consideravelmente superior aos demais semestres.

Figura 28 - Quantidade de trancamento em DIS0002

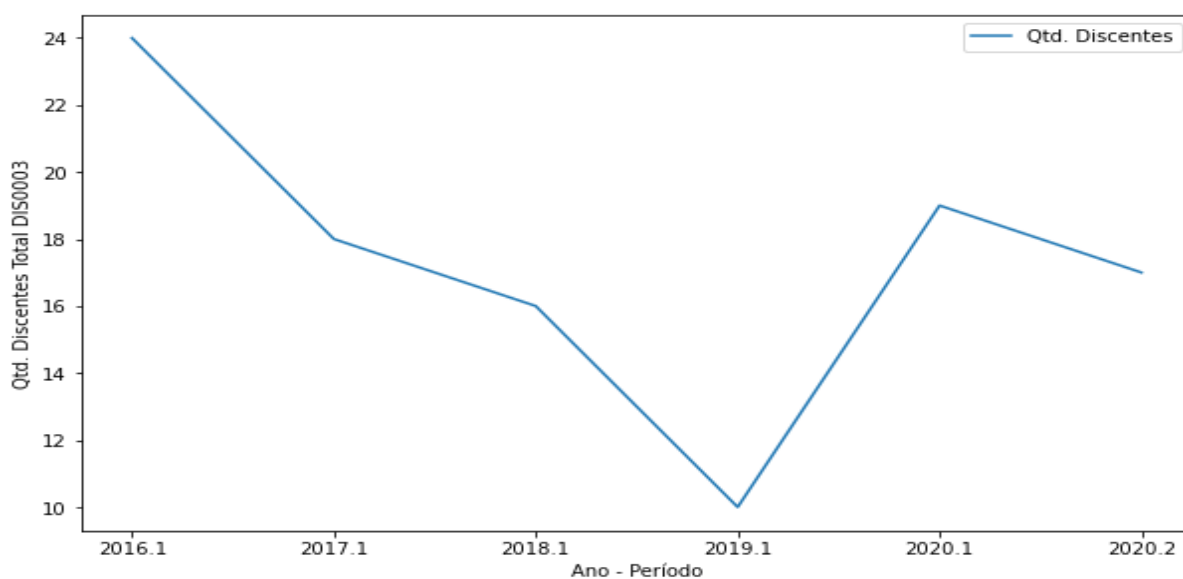


Fonte: Autoria própria

Após averiguarmos ambas disciplinas de graduação selecionadas, vimos como importante trazer uma do nível de Pós-graduação para exibir que tais consultas e filtros poderiam ser feitas na ferramenta, assim foi feito as análises anteriores para a disciplina DIS0003, onde há diferenças desta para as disciplinas de graduação citadas anteriormente.

A quantidade de discentes que fizeram a pré-matrícula na disciplina DIS0003 no semestre em pandemia se manteve na média quando comparados aos outros semestres, exceto em 2016.1 e 2019.1 onde em ambas teve índices mais acentuados, tanto em quantidade acima como abaixo da média, como pode ser observado na figura 29.

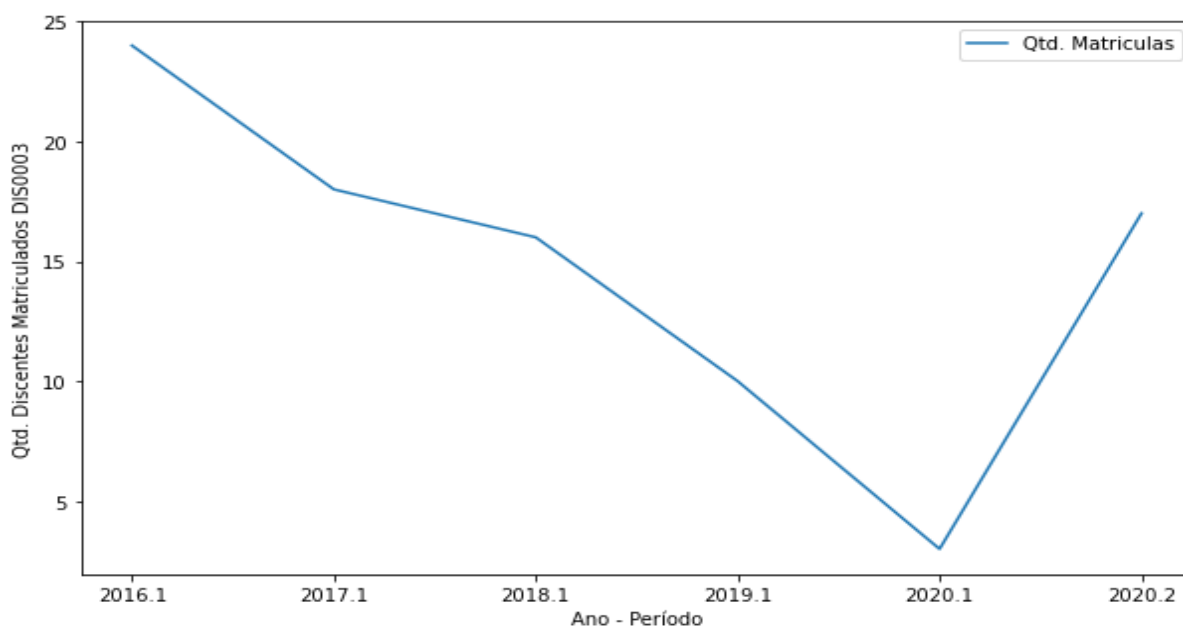
Figura 29 - Quantidade de discentes gerais em DIS0003



Fonte: Autoria própria

Observando o gráfico de matrículas efetivadas (figura 30), podemos perceber que no semestre 2020.2 teve um alto índice de discentes matriculados, mas semelhantes aos semestres 2017.1 e 2018.1, mas vale destacar a diferença entre o semestre 2020.1 e 2020.2, pois no primeiro foi quando houve a alteração de sistema de ensino presencial para a distância.

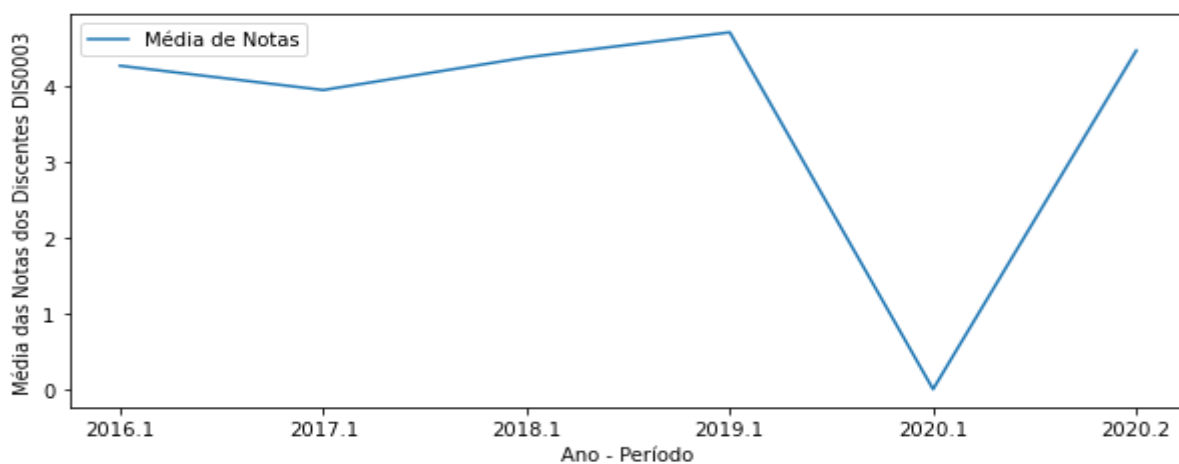
Figura 30 - Quantidade de discentes em DIS0003



Fonte: Autoria própria

Para podermos fazer uma visualização de desempenho referente a média das notas e de faltas, foram gerados os seguintes gráficos nas figuras 31 e 32. Na figura 31, onde consta a média de notas, podemos perceber que no semestre pandêmico 2020.2 a média dos alunos foi próxima aos demais semestres (exceto 2020.1).

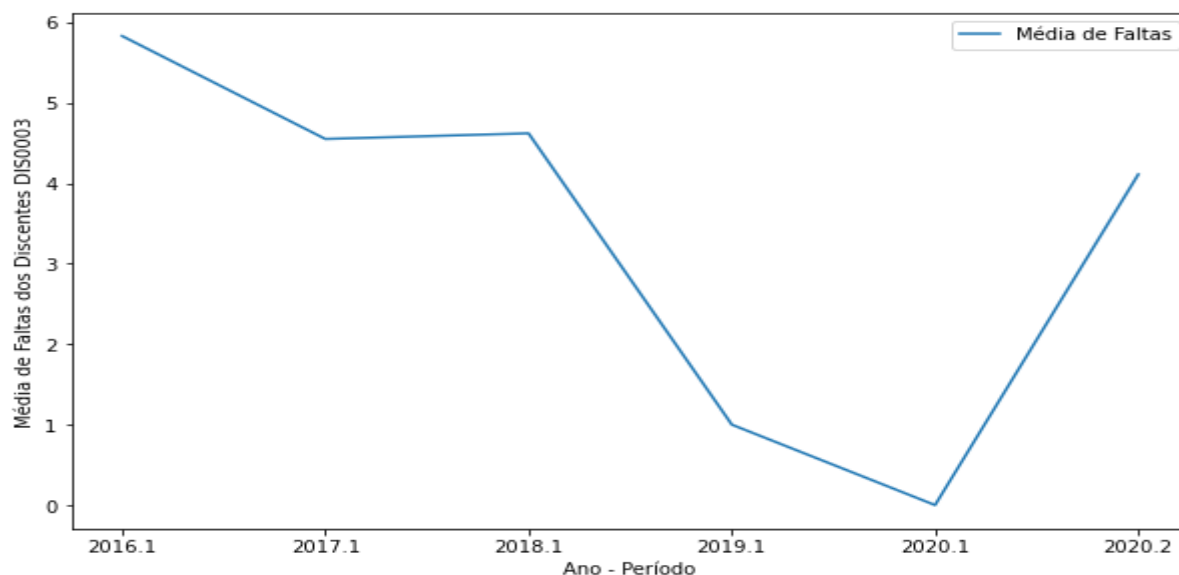
Figura 31 - Média de notas em DIS0003



Fonte: Autoria própria

Na figura 32, onde consta a média de faltas, podemos perceber que no semestre pandêmico 2020.2 a quantidade de faltas apresentou comportamento semelhante à média das notas, exceto no semestre 2019.1.

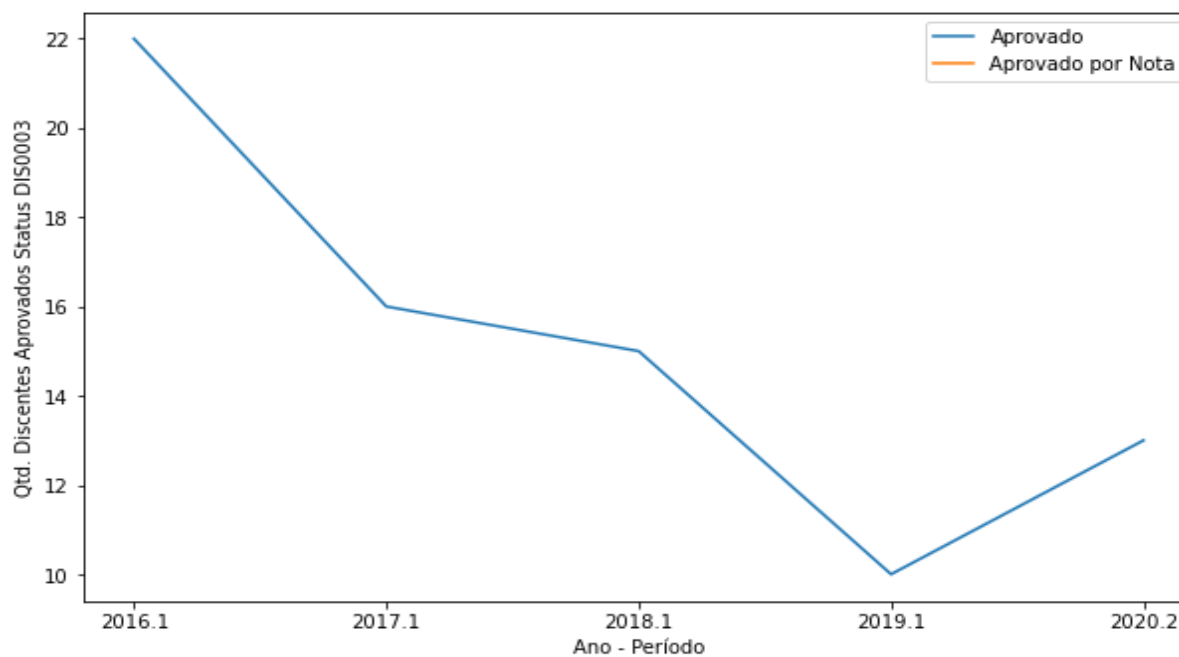
Figura 32 - Média de faltas em DIS0003



Fonte: Autoria própria

Em relação aos status de aprovação, foi observado somente com status 'Aprovado Por Nota' e pode ser observado na figura 33, onde podemos perceber que, exceto 2019.2, o semestre de 2020.2 teve a menor quantidade de discentes aprovados.

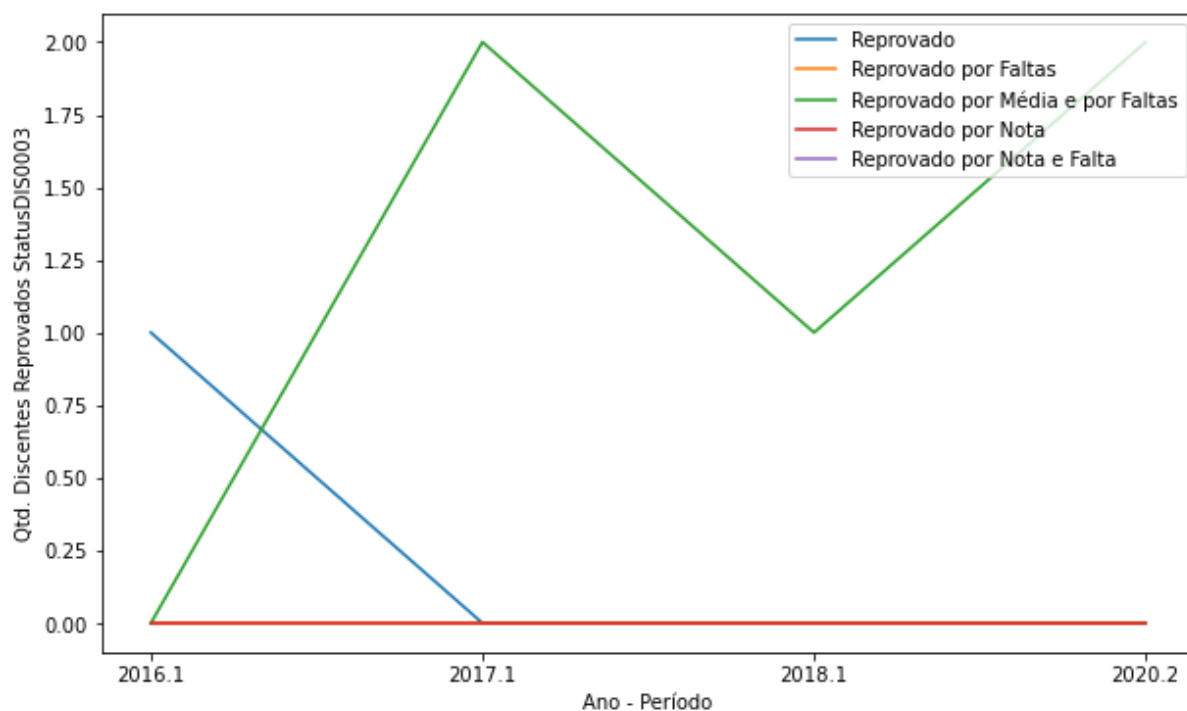
Figura 33 - Quantidade de aprovações por status em DIS0003



Fonte: Autoria própria

Em relação à quantidade de discentes reprovados, exibido na figura 34, foi observado que em 2020.2 obteve a mesma quantidade do semestre 2017.1, sendo elas com quantidade maiores se comparados aos demais semestres.

Figura 34 - Quantidade de reprovações por status em DIS0003

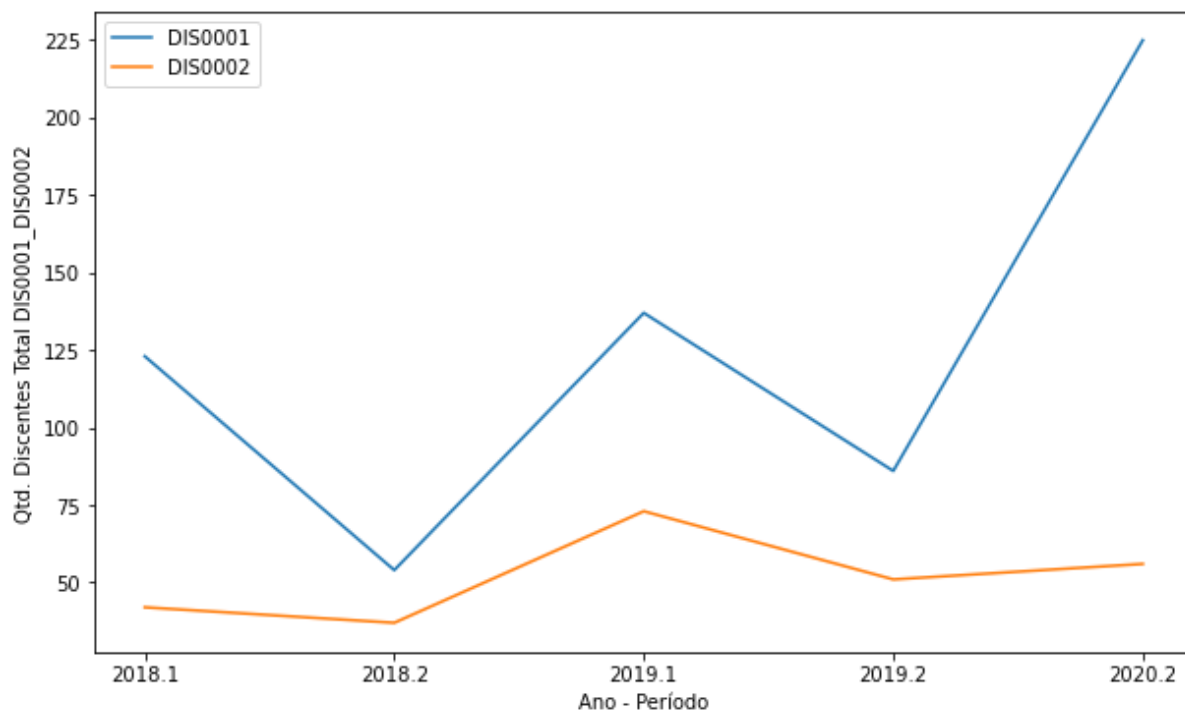


Fonte: Autoria própria

Com essas informações, podemos ver que é possível realizar uma busca e trazer informações que auxiliam ao usuário a realizar melhores decisões além de tornar estas mais ágeis e exibimos resultados de disciplinas específicas, mas pode ser utilizada entre 2 disciplinas, como exemplo, trazei o exemplo da disciplina DIS0001 e DIS0002, ambas exibidas separadamente acima. Elas foram escolhidas por serem disciplinas equivalentes, mas de unidades diferentes.

Em relação a quantidade de discentes que realizaram a pré-matrícula (figura 35), podemos observar que a disciplina DIS0001 teve uma busca maior quando comparada a DIS0002.

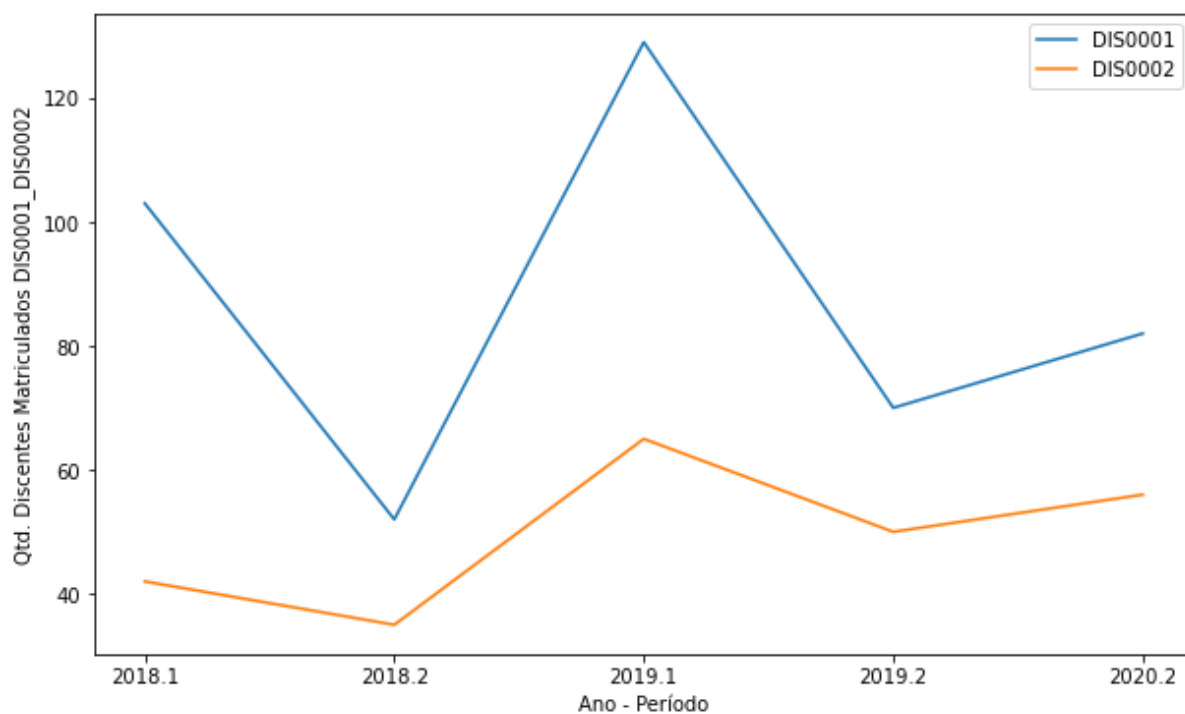
Figura 35 - Quantidade de discentes gerais em DIS0001 e DIS0002



Fonte: Autoria própria

Referente aos discentes que conseguiram ser matriculados (figuras 36), podemos visualizar que na quantidade de discentes com matrícula efetivadas, a DIS0001 também obteve quantidades maiores.

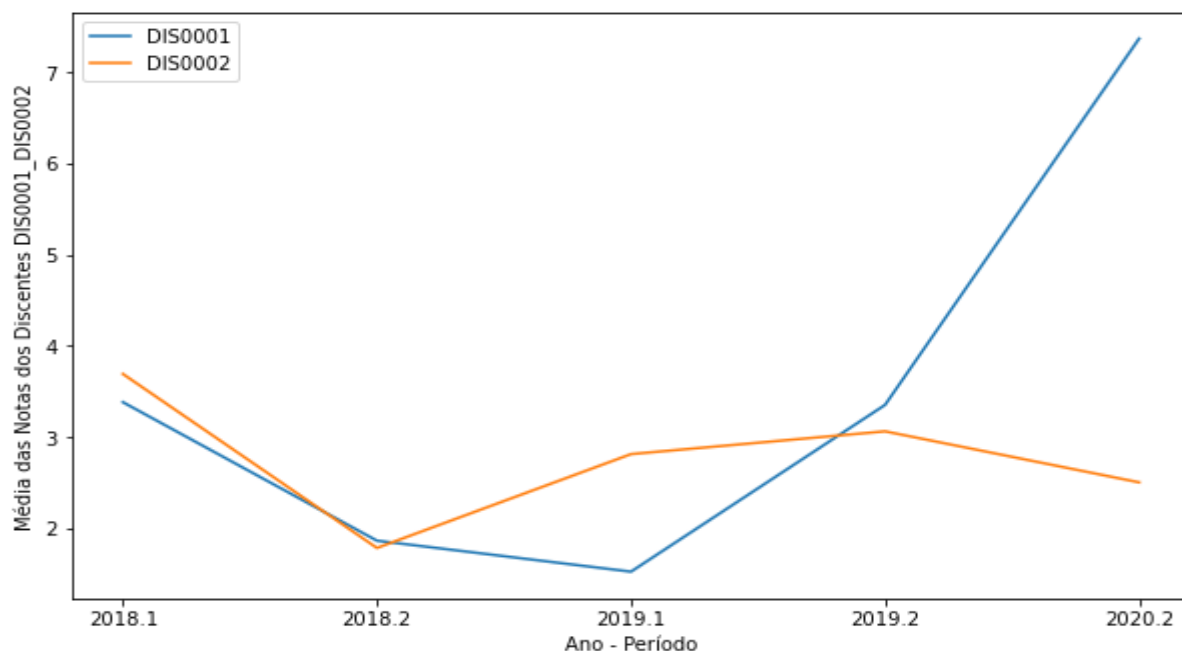
Figura 36 - Quantidade de discentes em DIS0001 e DIS0002



Fonte: Autoria própria

Para verificar o desempenho de forma geral, foi verificada a média das notas e faltas das duas disciplinas. Como podemos perceber na figura 37, a média dos estudantes em período de pandemia da disciplina DIS0001 foi superior a disciplina DIS0002.

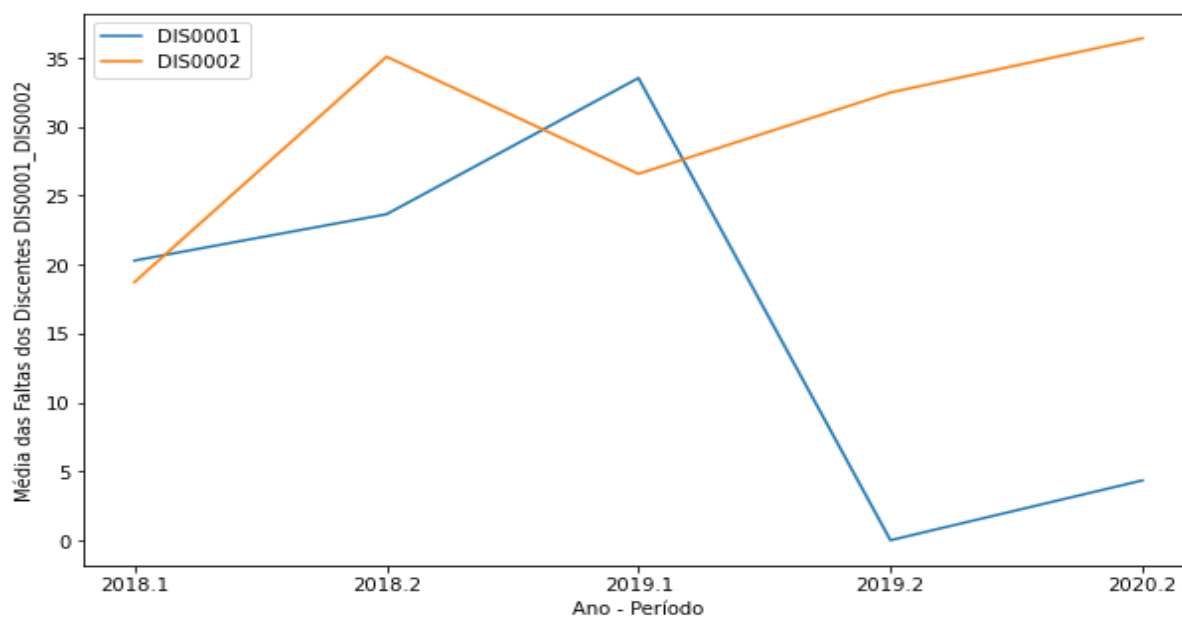
Figura 37 - Média de notas em DIS0001 e DIS0002



Fonte: Autoria própria

Outro ponto positivo para a disciplina DIS0001, como podemos perceber no gráfico da figura 38 é referente a média de faltas, onde a quantidade de faltas média foi bem abaixo quando comparada a disciplina DIS0002 nos períodos pandêmicos.

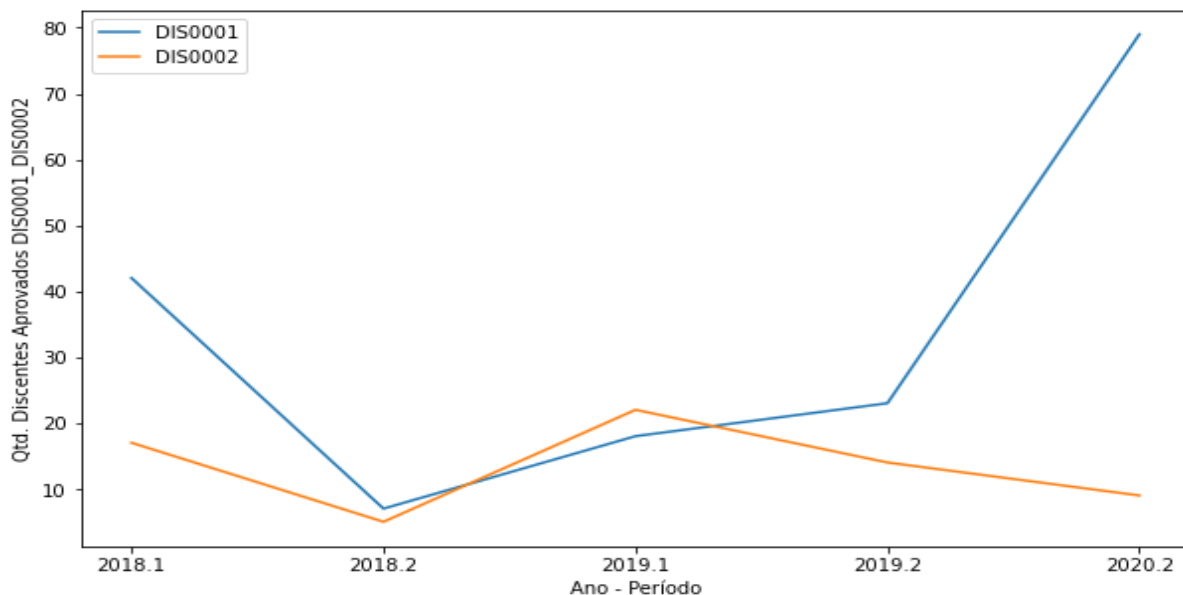
Figura 38 - Média de faltas em DIS0001 e DIS0002



Fonte: Autoria própria

Para uma melhor visualização do desempenho entre ambas as disciplinas, pode-se visualizar os gráficos de aprovações e reprovações de ambas as disciplinas. Observando a figura 39, onde exibe a quantidade de aprovações, pode-se observar que a disciplina DIS0001 possui uma quantidade de aprovações superior quando comparada a disciplina DIS0002.

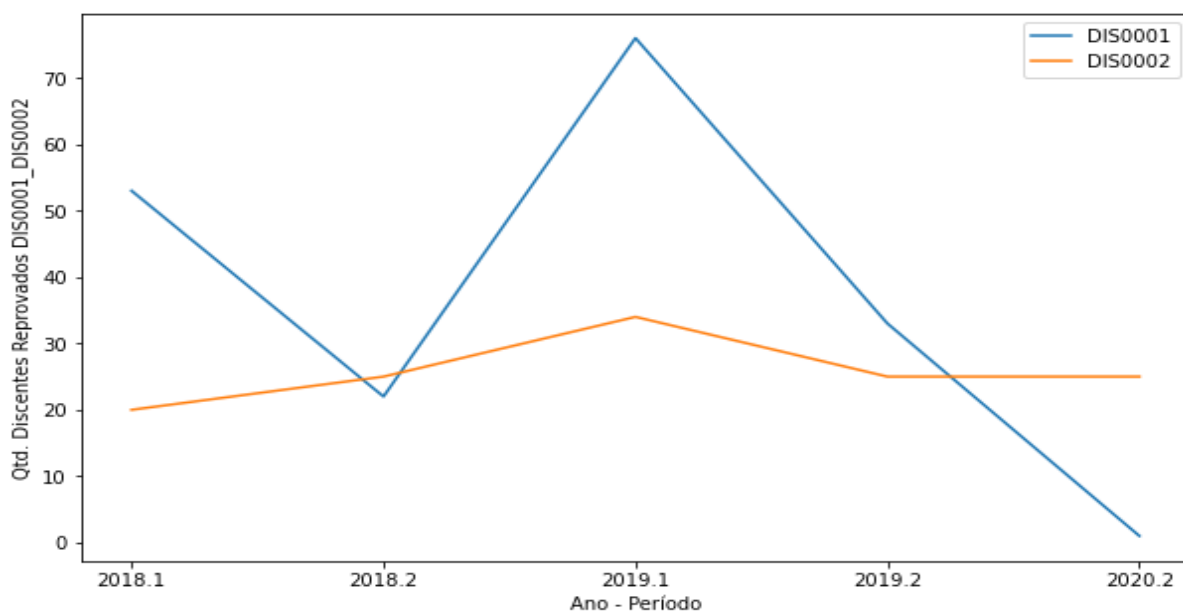
Figura 39 - Quantidade de aprovações por status em DIS0001 e DIS0002



Fonte: Autoria própria

Com base na figura 40, onde exibe a quantidade de reprovações, se avaliar somente o período pandêmico, a disciplina DIS0001 também apresenta uma quantidade de reprovação abaixo da outra disciplina.

Figura 40 - Quantidade de reprovações por status em DIS0001 e DIS0002

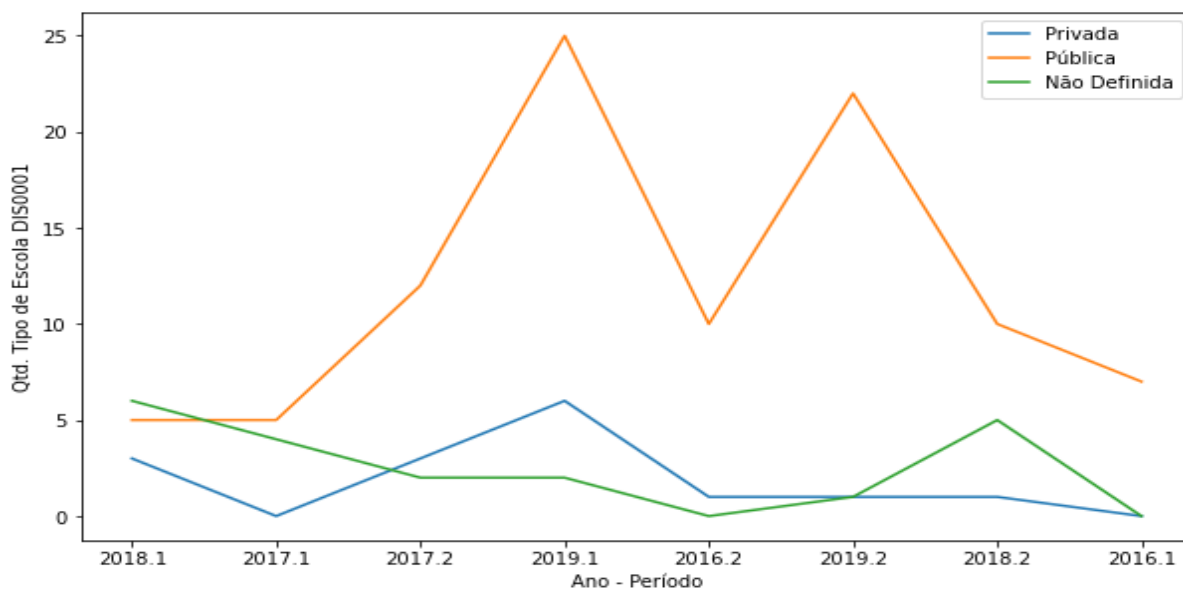


Fonte: Autoria própria

Além dos dados referente a notas e status das matrículas, pela ferramenta podem ter acesso aos dados socioeconômicos, como por exemplo a escola de segundo grau, a quantidade de bolsas e a faixa de renda dos discentes.

Verificando a figura 41, podemos observar que a maioria dos discentes que cursaram a disciplina DIS0001, fizeram o ensino médio em escolas públicas.

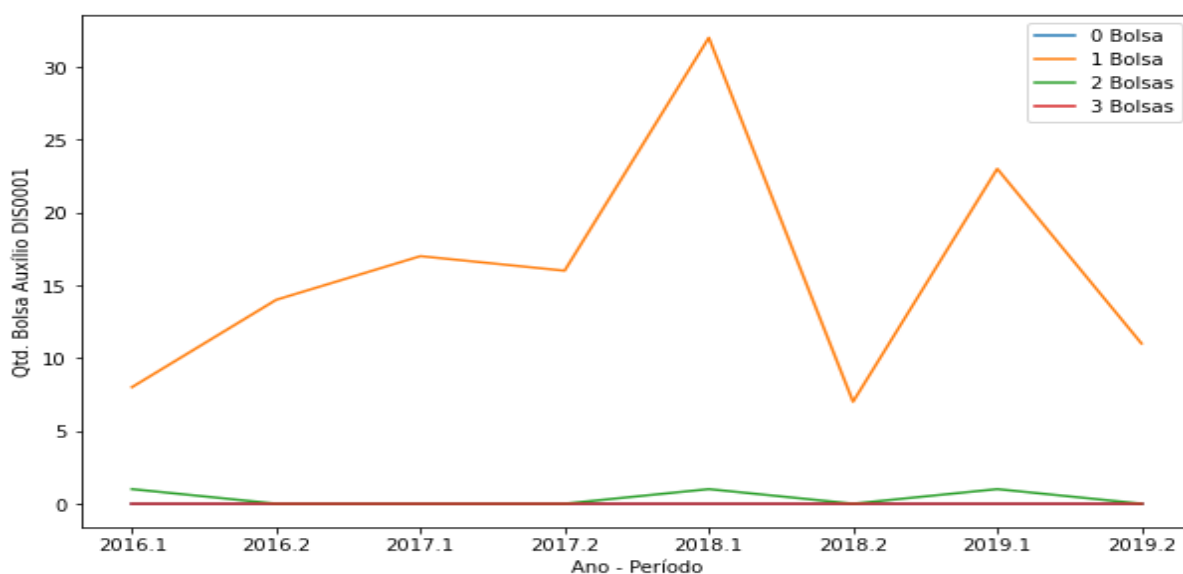
Figura 41 - Quantidade por escola ensino médio em DIS0001



Fonte: Autoria própria

Com base na quantidade de bolsas, podemos verificar a figura 42, onde exibe que a maioria dos estudantes da disciplina DIS0001 contém somente uma bolsa, mas tendo semestre da existência de discentes com 2 bolsas.

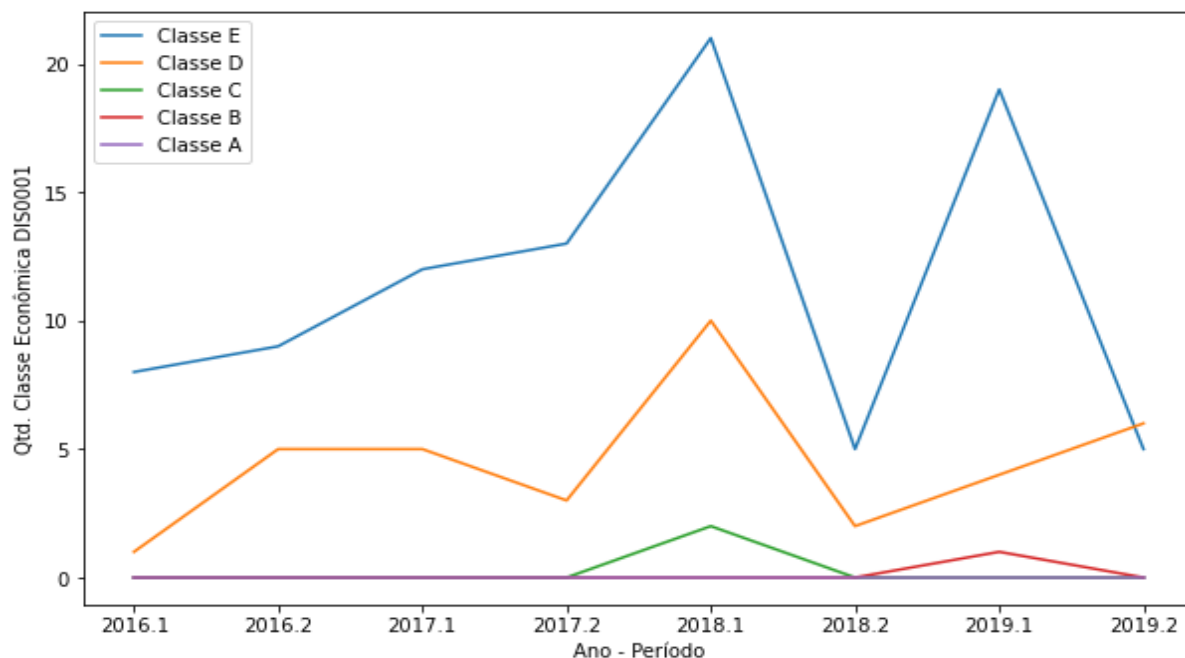
Figura 42 - Quantidade de bolsas em DIS0001



Fonte: Autoria própria

Para verificar as classes socioeconômicas dos estudantes da disciplina DIS0001, podemos verificar na figura 43 que a maioria dos estudantes são de classe E e classe D.

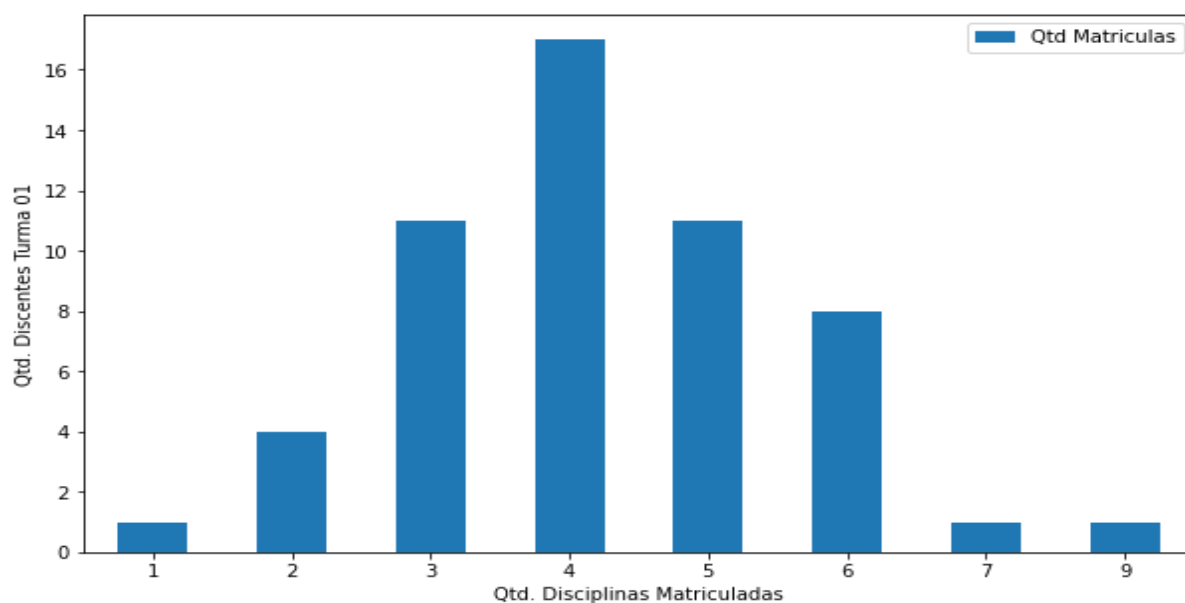
Figura 43 - Classe econômica em DIS0001



Fonte: Autoria própria

Para complementar a informações, pode-se fazer uma busca com base em uma disciplina para verificar quantas matérias cada discente está matriculado, para que seja possível mensurar se a alta densidade de disciplina pode ter impactado no desempenho deste aluno, como podemos observar na figura 44.

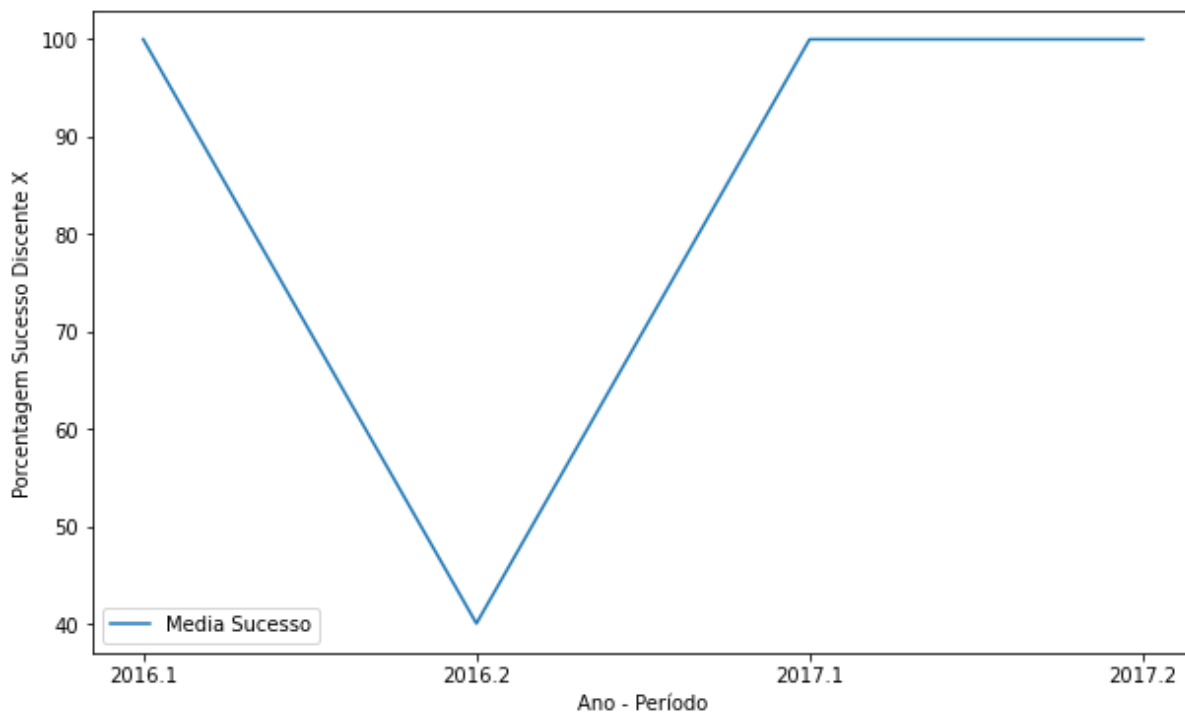
Figura 44 - Quantidade de alunos por quantidade de matrículas



Fonte: Autoria própria

E para ser mais específico, será possível o usuário do sistema verificar com base em um discente, o seu desempenho ao longo de semestres, como por exemplo na figura 45.

Figura 45 - Sucesso por um discente ao longo dos semestres



Fonte: Autoria própria

Podemos também analisar as informações mais detalhadas por turmas, conforme exibido na tabela 1. Nesse caso, pensamos em um modelo de tabela onde exhibe os principais dados de cada um de forma mais direta.

Tabela 1 - Comparação entre turmas

	Turma 1	Turma 2	Turma 3
Ano-Período	2020.6	2020.6	2020.6
Disciplina	DIS0001	DIS0001	DIS0001
Horário	246N12	246M56	246T34
Nível	GRADUAÇÃO	GRADUAÇÃO	GRADUAÇÃO
Número de unidades	3	3	3
Professor	Siape: 0000001 - PAULO DA CUNHA	Siape: 0000002 - JOÃO PEREIRA	Siape: 0000001 - PAULO DA CUNHA
Aulas lançadas	84	111	84
Alunos Totais	56	83	29
Alunos Aprovados	0	68	3

Fonte: Autoria própria

Pensamos também em uma funcionalidade somente para os gestores/coordenadores do curso, que é a análise pelo docente, com isso, caso a pessoa que esteja utilizando o sistema seja coordenadora de um curso, ela teria acesso a todos os filtros já disponibilizados e ao filtro de professor, podendo assim realizar uma análise das disciplinas lecionadas por docente ao longo dos semestres.

Com isso, podemos observar algumas das possíveis informações que poderiam ser apresentadas na ferramenta e possíveis análises que poderiam ser realizadas, interpretando-as.

5. PROPOSTA DA FERRAMENTA

A ferramenta foi pensada para auxiliar os docentes e gestores de ensino a analisar melhor o desempenho dos estudantes, identificar possíveis causas de desempenho abaixo do esperado e com isso melhorar o seu rendimento escolar. Com isso, foi pensando as seguintes funcionalidades:

1. Realizar a comparação do desempenho de estudantes entre disciplinas com base na média de notas e faltas, status de aprovações e reprovações, além de ser possível verificar disciplinas semelhantes mas que possuem departamentos diferentes;
2. Verificar graficamente as informações referente a dados socioeconômicos (Faixa de renda, existência de bolsas/auxílios, tipo de escola no ensino médio) relacionados aos discente da disciplina selecionada;
3. Realizar comparação entre turmas, podendo analisar além dos itens anteriores, as informações mais detalhadas de cada turma, como por exemplo o horário, o professor atribuído a ela e a nota por unidade;
4. Verificar para cada discente, a quantidade de disciplinas que o mesmo está matriculado no semestre. Assim, seria possível verificar se os alunos daquele período estão matriculados em disciplinas em excesso (podendo prejudicar o desempenho) ou até mesmo, o contrário, em poucas disciplinas;
5. Comparar desempenho por estudante ao longo dos semestres, podendo assim verificar a taxa de sucesso e insucesso de estudantes ao longo dos semestres passados;
6. Opção de realizar diversos filtros como por exemplos período, curso, disciplina e etc, para poder ter as informações que o usuário esteja buscando;
7. Salvar consultas feitas para monitoramento posterior, assim não precisando colocar todos os filtro em um segundo momento;
8. Possibilidade de exportar as informações para um modelo de planilha, caso assim o usuário desejar;
9. Importação de dados por JSON, para uma análise mais rápida. Dependendo da fonte de dados, essa opção pode auxiliar o docente para que o mesmo insira dados que ainda não estejam disponíveis (ex: comparar as informações após a prova de sua unidade 1);

5.1. Mockup

Com isso, iniciamos o desenvolvimento para que seja possível exibir o layout com base em um mockup e algumas das possíveis informações que teriam no sistema.

Em relação ao mockup, foi pensado uma forma de o usuário poder realizar a consulta de forma prática, mas completa, ficando assim a cargo do mesmo definir quais parâmetros utilizar para realizar a consulta, além de ter acesso a abas, para assim utilizar caso queira comprar disciplinas/turmas distintas. Com base nisso é feito a consulta levando em consideração os parâmetros escolhidos.

Os resultados foram pensados para serem exibidos por meio de tabela ou em formato de gráfico e foi produzido através do site *moqups*¹⁷. Como exemplo, podem ser observadas as figuras 46 e 47, com dados hipotéticos.

A figura 46 exibe o menu na lateral esquerda, os filtros e uma tabela informativa contendo as informações mais específicas de uma turma.

Na lateral esquerda é onde se localiza o menu de acesso às funcionalidades do sistema.

Ao realizar a consulta, poderá realizar com os seguintes filtros: ano, período, nível de ensino, nome de departamento, nome do curso, nome da disciplina, turma, unidade, status/situação da matrícula, professor, modelo de ensino, discente, quantidade de bolsa auxílio e escolaridade do ensino médio.

Além disso, inicia-se a exibição das informações com base nos parâmetros selecionados, a primeira parte é em forma de tabela, apresentado algumas informações mais específicas da turma, como por exemplo a disciplina, a turma, o docente, o horário e a taxa de aprovação com base nos alunos que obtiveram sucesso e a quantidade total de discentes matriculados.

¹⁷ <https://moqups.com/>

Figura 46 - Filtro e tabela de dados no Mockup

The mockup displays a user interface for a professor named CARLA. It features a sidebar with navigation icons and a main content area. The filter panel includes dropdown menus for various attributes, and the data table below it shows details for three different course groups (Turmas).

Filtro de Dados:

- Selecção 1, Seleção 2, Seleção 3, +
- Ano, Período, Nível de ensino, Departamento
- Curso, Disciplina (nome/código), Turma, Unidade
- Status da matrícula, Professor*, Modelo de ensino, Discente
- Bolsa auxílio, Escolaridade ensino me
- Pesquisar, Voltar

DADOS DAS TURMAS (DIS0001)

Turma	Docente	Ano - Período	Aprovações	Horário
Turma: 01	Docente: Carla da Silva	Ano - Período: 2019.1	Aprovações: 30/70	Horário: 246N12
Turma: 02	Docente: Carla da Silva	Ano - Período: 2019.2	Aprovações: 52/70	Horário: 246N34
Turma: 03	Docente: Carla da Silva	Ano - Período: 2020.2	Aprovações: 43/70	Horário: 246T34

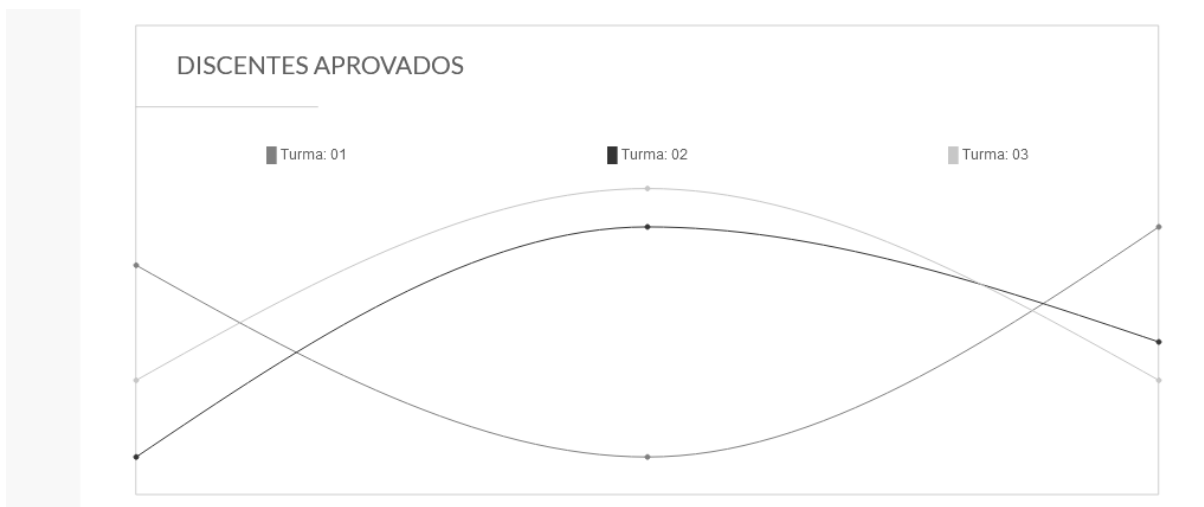
Mais detalhes

Fonte: Autoria própria

Na figura 47 podemos observar como seria possível exibir as informações de desempenho por meio de gráficos, onde a legenda auxilia no entendimento de qual informação está sendo exibida com auxílio dos títulos dos eixos x e y, sendo eles assim passíveis de um melhor entendimento para as informações.

Foi realizada a escolha do tipo gráfico de linha, também conhecido como série ou pontos conectados, por se tratar de um gráfico que exibe informações ao longo de um período de tempo, assim sendo uma das principais formas do sistema exibir as informações para serem utilizadas para as análises.

Figura 47 - Gráficos no Mockup



Fonte: Autoria própria

Para uma melhor visualização de como seria o funcionamento da ferramenta após construção, podemos observar a figura 48 e 49, onde ilustra um mockup contendo os dados semelhantes ao utilizados no capítulo de desenvolvimento deste projeto, assim apresentando na figura 48 possíveis filtros já selecionados pelo usuário, além de apresentar na figura 49 os gráficos de discentes aprovados e reprovados com base no parâmetros de busca. Importante citar que a ferramenta não se limitaria a somente estas informações, tendo assim outras informações a serem exibidas. Com isso, pode-se observar os parâmetros selecionados na figura 48: Graduação como nível de ensino, Departamento de Computação como departamento responsável pela disciplina, Curso de engenharia de software e a disciplina DIS0002.

Figura 48 - Exemplo de filtros com dados da disciplina DIS0002

O mockup da interface de usuário mostra uma tela de filtros para a disciplina DIS0002. A interface inclui um menu lateral com ícones para home, busca, documentos, gráficos, compartilhamento e perfil. No topo direito, há uma barra de usuário com o nome "CARLA" e o cargo "Professor(a)". A seção de filtros contém uma barra de seleção com "Seleção 1", "Seleção 2", "Seleção 3" e um botão "+". Abaixo, há uma grade de campos de seleção:

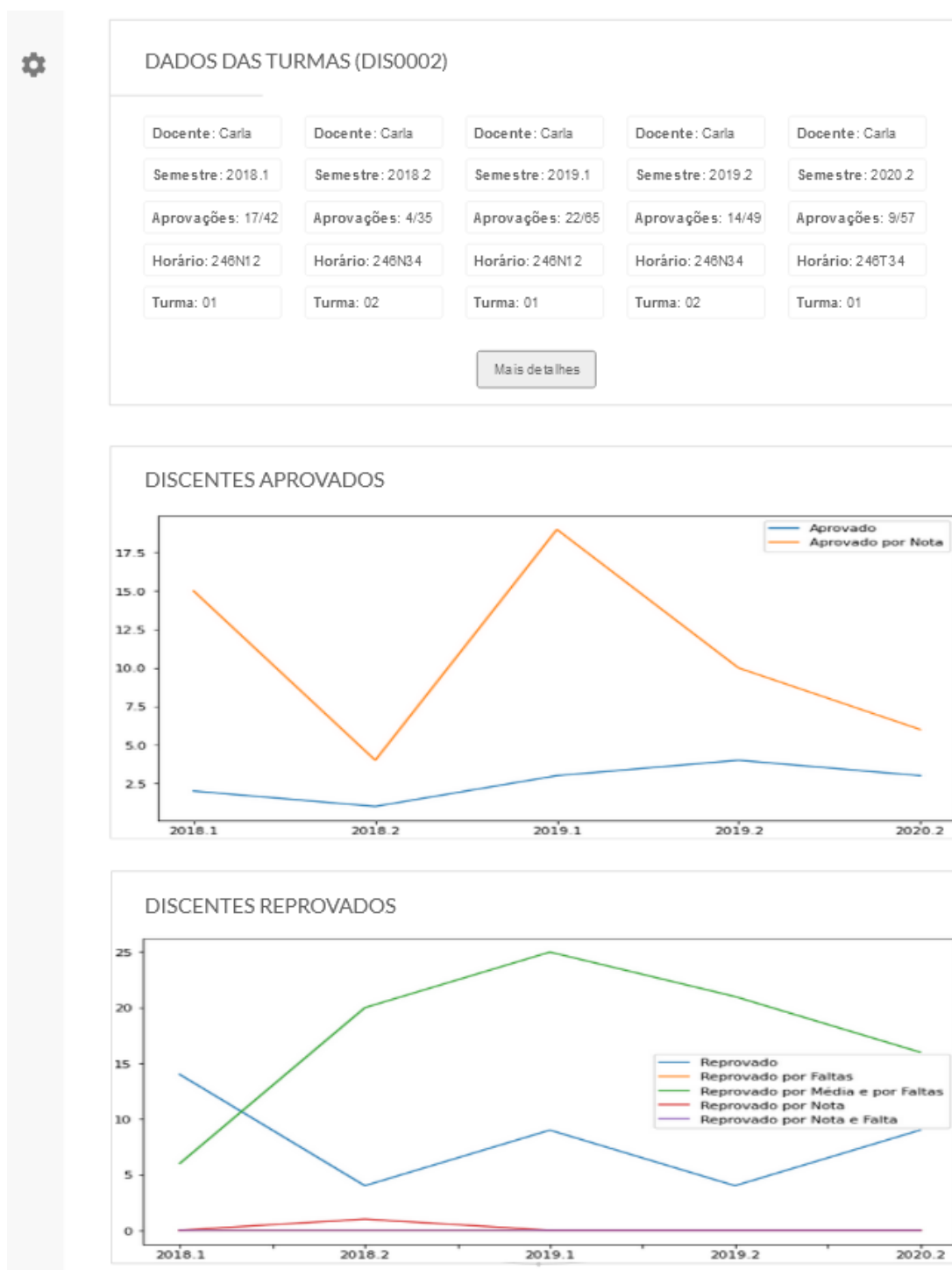
- Ano
- Período
- Graduação
- Departamento de Computação
- Engenharia de Software
- DIS0002
- Turma
- Unidade
- Status da matrícula
- Professor*
- Modelo de ensino
- Discente
- Bolsa auxílio
- Escolaridade ensino médio

Na base da seção de filtros, há dois botões: "Pesquisar" e "Voltar".

Fonte: Autoria própria

Na figura 49 pode-se observar algumas das possíveis informações a serem obtidas a partir dos parâmetros selecionados como filtro.

Figura 49 - Exemplo de informações obtidas referente a disciplina DIS0002



Fonte: Autoria própria

Com isso, é possível verificar alguns dos recursos da ferramenta e um mockup como exemplo, podendo assim auxiliar na análise do desempenho acadêmico dos estudantes.

6. CONCLUSÃO

O desempenho dos estudantes em uma instituição de ensino é de suma importância, não só para os discentes, mas como para os docentes e toda a instituição. Com o surgimento do coronavírus, a metodologia de ensino teve de ser adaptada em pró da saúde e da educação para um formato novo onde todos os envolvidos não estavam acusados, sendo assim um momento de aprendizado de professor e aluno.

A utilização de uma ferramenta como a citada nesse trabalho, que pudesse exibir informações importantes de disciplinas, turmas e discentes, poderia contribuir para melhorar o compartilhamento de conhecimento entre docente e discente, além de otimizar o tempo e melhorar as análises feita com base nos semestre anteriores, auxiliando os gestores e professores para melhorar o desempenho do curso e inclusive melhorar a metodologia de ensino para que nos demais semestres possa se ter um desempenho melhor dos discentes.

Com base nas informações exibidas das disciplinas de graduação DIS0001 e DIS0002 e considerando como caso de teste a época em pandemia, podemos verificar que os discentes da disciplina DIS0001 obtiveram uma quantidade de aprovações bem superior aos demais semestres além de uma média de notas alta e de faltas baixa. O qual o mesmo não aconteceu com a disciplina DIS0002, onde pode-se observar que o nível de sucesso dos estudantes foi inferior a alguns semestres anteriores. Se analisarmos a disciplina de Stricto DIS0003, podemos observar que apesar da média das notas continuarem altas, a quantidade de faltas aumentaram no período em pandemia, além disso, a quantidade de aprovações foi relativamente menor a alguns semestre, entretanto as reprovação se manteve na média dos demais. Além dessas análises, podemos ver que é possível informar dados socioeconômicos de uma turma, sendo eles o tipo de escolaridade do ensino médio, a classe econômica e a quantidade de bolsas.

Com estas informações é visível que uma ferramenta que forneça essas informações em larga escala, ou seja, sendo utilizado pelos docentes e coordenadores de cursos, irá fornecer uma visão mais ampla sobre o rendimento acadêmico e auxiliar na descoberta de possíveis fatores que estejam prejudicando o desempenho estudantil. Além de otimizar o tempo no momento de uma análise em uma disciplina ou até mesmo na elaboração de um projeto pedagógico. Não se limitando a uma análise geral, ela pode ser feita referente a uma turma ou discente, podendo assim fornecer uma análise mais direta para tentar corrigir possíveis inconsistências que estejam causando mal desempenho da turma ou de um discente.

Como trabalho futuro, podemos citar o desenvolvimento da ferramenta para que os gestores e docentes da rede de ensino possam utilizar esta ferramenta. Além de ser válido ser um projeto colaborativo com a STI para que se possa desenvolver em conjunto a ferramenta, considerando que ela é portadora das informações necessárias para o desenvolvimento, mas além dessas formas, há a possibilidade de capturar as informações dos dados abertos da UFRN.

Além disso, pode ser realizado outras abordagens além das propostas nesse trabalho, pensando assim em outras informações importantes para que auxilie ainda mais os usuários da ferramenta, inclusive sendo possível expandir essa ferramenta para todos os níveis de ensino e a outras instituições para que também auxiliem as mesmas.

7. REFERÊNCIAS

Boruchovitch, Evely (1999): Estratégias de aprendizagem e desempenho escolar: considerações para a prática educacional.

Bulhões, Rodrigo. Estatística Descritiva e Inferencial, 2011. Disponível em: <https://paca.ime.usp.br/pluginfile.php/45212/mod_resource/content/2/Aula_Rodrigo.pdf>. Acesso em: 30/11/2021.

Capellini, S. A.; Tonelotto, J. M. F. e Ciasca, S. M. (2004): Medidas de desempenho escolar: Avaliação formal e opinião de professores.

Colpani, Rogério (2018): Mineração de Dados Educacionais: um Estudo da Evasão no Ensino Médio com Base nos Indicadores do Censo Escolar.

Costa, E.; Baker, Ryan S.J.d.; Amorin, L.; Magalhães, J. e Marinho, T. (2012): Mineração de Dados Educacionais: Conceitos, Técnicas, Ferramentas e Aplicações.

Creswell, J. W. (2007): Projeto de pesquisa: métodos qualitativo, quantitativo e misto (2a ed., L. de O. Rocha, Trad.). Porto Alegre: Artmed. (Obra original publicada em 2003).

Dados Abertos (2020). Disponível em: <<https://www.bcb.gov.br/acesoinformacao/dadosabertos/>>. Acesso em: 11/10/2021.

Drumond, Kelly (2020). Projeto Político-Pedagógico: o que é e como elaborar o PPP. Disponível em: <<https://www.sistemamaxi.com.br/projeto-politico-pedagogico/>>. Acesso em: 20/11/2021.

Echeveste, S.; Bittencourt, H.; Bayer, A. e Rocha, J. (2005): Educação estatística: perspectivas e desafios.

Ferreira, Pedro (2015): Estatística descritiva e inferencial.

Gomes, M. Manoel (2018). Fatores que facilitam e dificultam a aprendizagem.

Ghidini, Eduardo (2017): Representação de narrativas interativas por meio de visualização com dados extraídos de redes sociais.

Ignácio, S. Aparecido (2010). Importância da estatística para o processo de conhecimento e tomada de decisão.

Kawakami, Guto; Rivero, Luis e Conte, Tayana (2014): Identificando Oportunidades de Melhoria em Tecnologias de Inspeção de Usabilidade para o Contexto Web Móvel através de um Experimento Controlado.

Lacerda, Ivan M. F. (2018): Metodologia de Análise de Programas de Educação a Distância Baseada em Learning Analytics.

Lopes, Celi E. (2008): O ensino da estatística e da probabilidade na educação básica e a formação de professores.

Lourenço, Adriana M. (2020): Trilhas de aprendizagem para formação de líderes.

Mozzato, Anelise R. e Grzybovski, Denize (2011): Análise de Conteúdo como Técnica de Análise de Dados Qualitativos no Campo da Administração: Potencial e Desafios.

Nascimento, Tiago (2019): Os dados são o novo “petróleo” do século XXI.

Pearson (2021). Trilhas de aprendizagem. Disponível em: <<https://hed.pearson.com.br/blog/plataformas-de-aprendizagem/trilhas-de-aprendizagem-voce-ja-conhece-esse-conceito>>. Acesso em: 25/01/2022.

Reis, E.A., Reis I.A. (2002): Análise Descritiva de Dados. Relatório Técnico do Departamento de Estatística da UFMG.

Rivero, L.; Viana, D. e Conte, T. (2013): Mockup DUE: Uma Ferramenta de Apoio ao Processo de Inspeção de Usabilidade de Mockups de Aplicações Web.

Silva, José M. G. (2015): Desenvolvimento de módulo de aprendizagem em UX design – A valorização de metodologias de diagnóstico no desenvolvimento de projeto.

Teixeira, Enise B. (2003): A Análise de Dados na Pesquisa Científica: importância e desafios em estudos organizacionais.

Zeviani, Walmes. Visualização de Dados, 2019. 161 slides. Disponível em: <<http://leg.ufpr.br/~walmes/cursoR/data-vis/slides/00-datavis.pdf>>. Acesso em: 18/11/2021.

Zeviani, Walmes. Almanaque de gráficos para visualização de dados, 2020. 161 slides. Disponível em: <<http://leg.ufpr.br/~walmes/cursoR/data-vis/slides/00-graphics-gallery.pdf>>. Acesso em: 19/11/2021.