



ALGORITMOS PARA OS INDICADORES DE FLUXO DE ALUNOS PARA AVALIAR A EVASÃO DOS CURSOS

DOI: 10.37702/2175-957X.COBENGE.2022.3955

Gabriel Permonian Bogni - gabpermon@gmail.com
Instituto Federal de São Paulo

Isabella Maria Avesani - isabella.avesani@usp.br
Universidade de Sao Paulo

Esther Eleotério Panequi Ferreira - estherpferreira@usp.br
Universidade de Sao Paulo

Mateus Henrique Platinetti Nanzer - mateusnanzer@gmail.com
Universidade Federal de São Carlos UFSCar

Sarita Mazzini Bruschi - sarita@icmc.usp.br
Universidade de São Paulo

Vilma A Oliveira - voliveira@usp.br
Universidade de São Paulo

Resumo: *Vários estudos são constantemente desenvolvidos com o objetivo de estudar o fenômeno da evasão escolar. No ensino de graduação, é necessário acompanhar a trajetória acadêmica dos alunos para a produção de um diagnóstico da desvinculação dos cursos e definir ações em âmbito institucional para combater as taxas de evasão. A evasão calculada para coorte de ingressantes é importante para entender como as turmas se comportam ao longo dos anos bem como para comparar a evasão nos diferentes cursos de engenharia. Neste trabalho, indicadores de fluxo de alunos são utilizados para estudar a trajetória acadêmica dos alunos. O cálculo dos indicadores de fluxo é automatizado por meio de um programa computacional criado para ler dados de planilhas eletrônicas e realizar o pareamento das bases de dados ano a ano para calcular a soma dos vínculos de alunos aos cursos por ano de ingresso. A automatização da utilização de banco de dados de coorte de ingressantes por programa computacional em Python torna viável o constante acompanhamento da trajetória dos alunos nos cursos bem como auxiliar na escolha de ações a serem implementadas em âmbito institucional a fim de reduzir a evasão.*





Palavras-chave: *Trajectoria acadêmica, coorte de ingressantes, programação computacional, pareamento de dados, automatização, diagnóstico, desvinculação.*



ALGORITMOS PARA CALCULAR OS INDICADORES DE FLUXO DE ALUNOS PARA AVALIAR A EVASÃO DOS CURSOS

Resumo: Vários estudos são constantemente desenvolvidos com o objetivo de estudar o fenômeno da evasão escolar. No ensino de graduação, é necessário acompanhar a trajetória acadêmica dos alunos para a produção de um diagnóstico da desvinculação dos cursos e definir ações em âmbito institucional para combater as taxas de evasão. A evasão calculada para coorte de ingressantes é importante para entender como as turmas se comportam ao longo dos anos bem como para comparar a evasão nos diferentes cursos de engenharia. Neste trabalho, indicadores de fluxo de alunos são utilizados para estudar a trajetória acadêmica dos alunos. O cálculo dos indicadores de fluxo é automatizado por meio de um programa computacional criado para ler dados de planilhas eletrônicas e realizar o pareamento das bases de dados ano a ano para calcular a soma dos vínculos de alunos aos cursos por ano de ingresso. A automatização da utilização de banco de dados de coorte de ingressantes por programa computacional em Python torna viável o constante acompanhamento da trajetória dos alunos nos cursos bem como auxiliar na escolha de ações a serem implementadas em âmbito institucional a fim de reduzir a evasão.

Palavras-chave: Trajetória acadêmica, coorte de ingressantes, programação computacional, pareamento de dados, automatização, diagnóstico, desvinculação.

1 INTRODUÇÃO

A evasão estudantil é usualmente definida como o desligamento do estudante da instituição sem a conclusão do curso, é ocasionada por diversos problemas que podem ser divididos em 4 grandes grupos: (i) problemas institucionais, relacionados à qualidade do curso, atividades desenvolvidas e infraestrutura; (ii) pessoais, relacionados principalmente a sensação de não estar no curso certo; (iii) financeiros e (iv) incertezas quanto à carreira (AMBIEL, 2015). Portanto, diagnosticar e entender os motivos que levam os estudantes a evadirem é fundamental para identificar ações institucionais para reduzir os índices de evasão, o que é benéfico tanto para a universidade quanto para a sociedade como um todo já que tem reflexos negativos no âmbito econômico e social, uma vez que afeta diretamente a formação de profissionais.

A automatização do cálculo dos indicadores de fluxo e da construção de ilustração gráfica dos resultados do cálculo pode ser eficientemente realizada através de um programa computacional. Conseqüentemente, o programa computacional pode ser utilizado para gerar mais informações de maneira automática explorando um banco de dados (tabela) com mesma estrutura (ordem das colunas) contendo dados de diversos anos, cursos e habilitações.

Fluxogramas auxiliam no desenvolvimento dos programas computacionais para planejar as funcionalidades desejadas, identificar os requisitos, definir a arquitetura e prevenir erros. Além disso, permitem o entendimento do funcionamento do programa por

peças que não possuem conhecimento na área da computação. A Linguagem Unificada de Modelagem (UML, iniciais em inglês) é uma linguagem gráfica utilizada para facilitar o entendimento das relações entre os critérios e condições do problema tratado em relação à elaboração de software (BOOCH, 2006). Um dos diagramas existentes nessa linguagem é o diagrama de sequência que obedece a ordem temporal em que as mensagens são trocadas entre os objetos envolvidos em um determinado processo (GUEDES, 2019). É usado para enfatizar a comunicação entre os elementos envolvidos no programa que podem ser externos ou não ao programa. Outra característica do diagrama de sequência é que pode ser usado para representar as etapas de um programa em baixo nível de abstração (um nível de detalhamento maior que se aproxima da prática em relação à programação)

Neste trabalho, os indicadores de fluxo de alunos definidos como taxa de permanência (TAP), taxa de conclusão acumulada (TCA) e taxa de desistência acumulada (TDA) são calculados conforme os modelos adotados pela Diretoria de Estatísticas Educacionais (INEP, 2017) a fim de mapear como evoluem ao decorrer dos anos a partir de uma coorte de ingressos em um curso de graduação. Para realizar os cálculos de forma automatizada foram construídos fluxogramas com base nas condições de permanência do curso em termos de desempenho, transferência ou desistência e sucesso usando a linguagem UML e a linguagem de programação Python através da plataforma Google Colab, um ambiente que permite escrever e executar Python no navegador sem necessidade de configurações específicas e de fácil compartilhamento (COLAB, 2022).

2 METODOLOGIA

As análises realizadas no presente trabalho estão voltadas à estatística descritiva dos dados, ou seja, a coleta, entendimento, sumarização e o estabelecimento de correlações entre os dados, sem contudo, se ater a suposições de um modelo probabilístico. Segundo o estudo sobre evasão no ICMC-USP (CURI, 2018), a estatística descritiva permite ao pesquisador criar uma base sólida para diagnóstico e futuras análises.

A partir de fluxogramas criados em linguagem UML para desenvolver a lógica de funcionamento do sistema, planejar a derivação e prevenir erros, foi desenvolvido, em linguagem Python, um programa com o objetivo de automatizar a busca de variáveis para cálculo de TAP, TCA e TDA com o auxílio de modelos matemáticos e uma planilha contendo os dados da instituição de ensino.

2.1 Modelo matemático

Nos processos educacionais formais, em todo o mundo, espera-se que um aluno ingressante siga determinada trajetória alcançando sucesso através da conclusão de certo nível educacional e/ou de um curso. A partir disso, é possível acompanhar cronologicamente sua posição em relação ao itinerário de formação superior por três indicadores: permanência (percurso), desistência (insucesso) e conclusão (sucesso) (INEP, 2017).

Dados o ano de ingresso de uma turma, o período de análise e o conjunto de dados, as seguintes taxas descrevem os indicadores de curso baseados nas taxas TAP, TCA e TDA definidas anteriormente. Para o cálculo das taxas acima mencionadas, foi

adotado o método proposto pela Diretoria de Estatísticas Educacionais (DEED) (INEP, 2017). Para cada ano t em um período dado em anos, para os ingressantes no ano T as taxas TAP, TCA e TDA são definidas por T_{ap} , T_{ca} , T_{da} , respectivamente, conforme segue:

$$T_{ap} = \frac{\sum_{i=1}^{n_{1,t}} Cur_{i,t} + \sum_{i=1}^{n_{2,t}} Mtr_{i,t}}{\sum_{i=1}^n IG_i^T - \sum_{w=T}^t \sum_{i=1}^{n_{6,w}} Fal_{t,w,i}} 100 \quad (1)$$

$$T_{ca} = \frac{\sum_{w=T}^t \sum_{i=1}^{n_{5,w}} For_{t,w,i}}{\sum_{i=1}^n IG_i^T - \sum_{w=T}^t \sum_{i=1}^{n_{6,w}} Fal_{t,w,i}} 100 \quad (2)$$

$$T_{da} = \frac{\sum_{w=T}^t \sum_{i=1}^{n_{3,w}} Des_{t,w,i} + \sum_{w=T}^t \sum_{i=1}^{n_{4,w}} Trans_{t,w,i}}{\sum_{i=1}^n IG_i^T - \sum_{w=T}^t \sum_{i=1}^{n_{6,w}} Fal_{t,w,i}} 100 \quad (3)$$

em que:

- Cur = Estudantes com situação de vínculo igual a “Cursando” no ano t .
- Mtr = Estudante com situação de vínculo igual a “Matrícula trancada” no ano t .
- IG = Número total de ingressantes com no ano T (fixado). Dado pela soma dos alunos na situação de cursando, falecidos, formandos, desvinculados e transferidos do ano T inicial até o ano t .
- Fal = Estudantes com tipo de encerramento a “Falecido” no ano t .
- For = Estudantes com tipo de encerramento igual a “Formado” no ano t .
- Des = Estudante com situação de vínculo igual a “Desvinculado do curso” no ano t .
- $Trans$ = Estudantes com tipo de encerramento igual a “Transferido para outro curso da mesma instituição de ensino superior” no ano t .

Em (1)-(3) foi considerado desvinculado do curso, o estudante que foi desligado pelos seguintes motivos: Abandono 2 semestres sem matrícula, Abandono 3 semestres sem matrícula, Cancelamento 0 créditos em 2 semestres, Cancelamento menos 20% dos créditos, Cancelamento menos de 20% créditos 2 semestres, Cancelamento outras IES (Lei Federal), Cancelamento trancamento 4 semestres, Cancelamento ultrapassou prazo máximo, Desistência a pedido, Desligamento aluno convênio, Encerramento novo ingresso, Ingressante sem frequência, Liminar cassada, Não cumpriu normas de retorno ao curso, Não cumpriu normas Decreto (PEC-G).

Vale ressaltar que a soma dos três indicadores é sempre igual a 100%, o que poderá ser observado nos gráficos apresentados. Neste trabalho, foi analisado como a desistência, a conclusão e a permanência evoluem no decorrer dos anos até t para a turma ingressante em T .

No estudo de caso considerado, foi feita a simulação dos indicadores de permanência, conclusão e desistência de uma Instituição de Ensino como um todo e, também, para cada curso da instituição individualmente. Para análise comparativa dos indicadores foram utilizadas as mesmas fórmulas em ambos os casos. É importante ressaltar apenas que, exclusivamente na análise de todo o instituto, o tipo de encerramento igual a “Transferência interna” foi considerado como aluno com situação de

"cursando" (*Cur*) já que, nesse caso, o aluno teria mudado de curso mas ainda seria um aluno ativo.

A implementação dos cálculos e contagens foi feita por rotinas em programa de planilha eletrônica e por códigos em linguagem Python o que trouxe maior confiabilidade e possibilitou alterações dinâmicas tanto nas condições de permanência e ocorrências previstas quanto nas variáveis.

2.3 Criação dos fluxogramas

Diversos diagramas de sequência em linguagem UML foram escritos e alterados ao longo do desenvolvimento do programa, mas apenas a parte principal do trecho responsável pelo cálculo das taxas de permanência ao longo dos anos para um curso seguido da habilitação específica é apresentada na Figura 1. Chama-se a atenção para a seção do fluxograma que faz o somatório dos alunos desistentes do banco de dados usado no estudo de caso feito, o qual segue o Artigo 75 do Regime Geral da USP¹ que regulariza as formas de cancelamento de matrícula, ou seja, cessação total dos vínculos do aluno com a Universidade. Formas essas que estão contidas no fluxograma da Figura 1. Os diagramas como o mostrado na Figura 1 indicam a lógica existente para produção de trechos de código que são responsáveis por calcular os indicadores de fluxo, formando a base para a construção posterior de outras partes do código, as quais, fazem uso dos resultados dos cálculos dos indicadores de fluxo para geração de gráficos. Assim, o diagrama da Figura 1 é um dos três fluxogramas principais para a criação do programa que calcula os indicadores taxas de permanência, conclusão acumulada e desistência acumulada.

2.4 Desenvolvimento do código Python

A linguagem de programação Python foi escolhida devido a facilidade de programação e também por sua utilização no desenvolvimento de programas anteriores e pertencentes ao projeto de estudo sobre a evasão para criar gráficos a respeito de estatísticas como perfil dos ingressantes, reprovações e formas de ingresso.

Para a realização dos cálculos, primeiramente é necessário que os dados estejam em arquivo no formato .xls (planilha Excel), com informações relevantes para os cálculos, tais como: ano de ingresso, tipo de encerramento e data do encerramento, cada uma apresentada na forma de coluna. No início do programa, esse arquivo é lido e os dados são armazenados em listas, uma estrutura de dados do Python. Através do modo que o arquivo foi construído, o primeiro elemento de uma coluna, que está localizado na primeira linha, se torna o primeiro elemento da lista, o segundo elemento da coluna, situado na segunda linha, se torna o segundo elemento da lista. Seguindo essa lógica, os elementos são dispostos em uma lista no código, entretanto, como o primeiro valor de uma coluna se refere ao cabeçalho, este não representa dado de um aluno, precisando ser ignorado no processo de cálculo da taxa de permanência, taxa de conclusão acumulada e taxa de permanência, pois não apresenta uma informação útil ao sistema.

Com as listas formadas, cada elemento de listas diferentes são unidos a fim de juntar informações sobre um mesmo aluno, logo é necessário cuidado no processo de formação de cada lista uma vez que um erro pode gerar problema de emparelhamento de dados e comprometer o resultado dos cálculos.

¹ <http://leginf.usp.br/?cat=17>

Conforme o algoritmo produzido, a taxa sobre algum indicador de fluxo é computada para cada ano desde o ano de ingresso do coorte de alunos analisado até um ano configurado na execução do algoritmo. Para todo ano calculado, o resultado é inserido em uma lista. Posteriormente, com dados de todos os anos, o gráfico é gerado para um dos indicadores de fluxo que pode ser a taxa de permanência, conclusão acumulada ou desistência acumulada.

Apesar do desenvolvimento do código seguir a lógica do fluxograma criado anteriormente, modificações foram necessárias e realizadas para a criação do programa computacional. Devido a isso, com o desenvolvimento do programa, o fluxograma precisou ser atualizado para estar de acordo com o programa final. A Figura 2, exibida a seguir, contém o trecho do código correspondente ao fluxograma da Figura 1. O código completo para a taxa de permanência bem como das taxas de conclusão acumulada e desistência acumulada podem ser obtidos via plataforma GitHub: https://github.com/Estudo-Evasao/Projeto_Programa_Evasao/blob/main/Programa_de_Representacao_Grafica_dos_Indicadores_de_Fluxo_de_Alunos_dos_Cursos_da_EESC_e_ICMC.ipynb

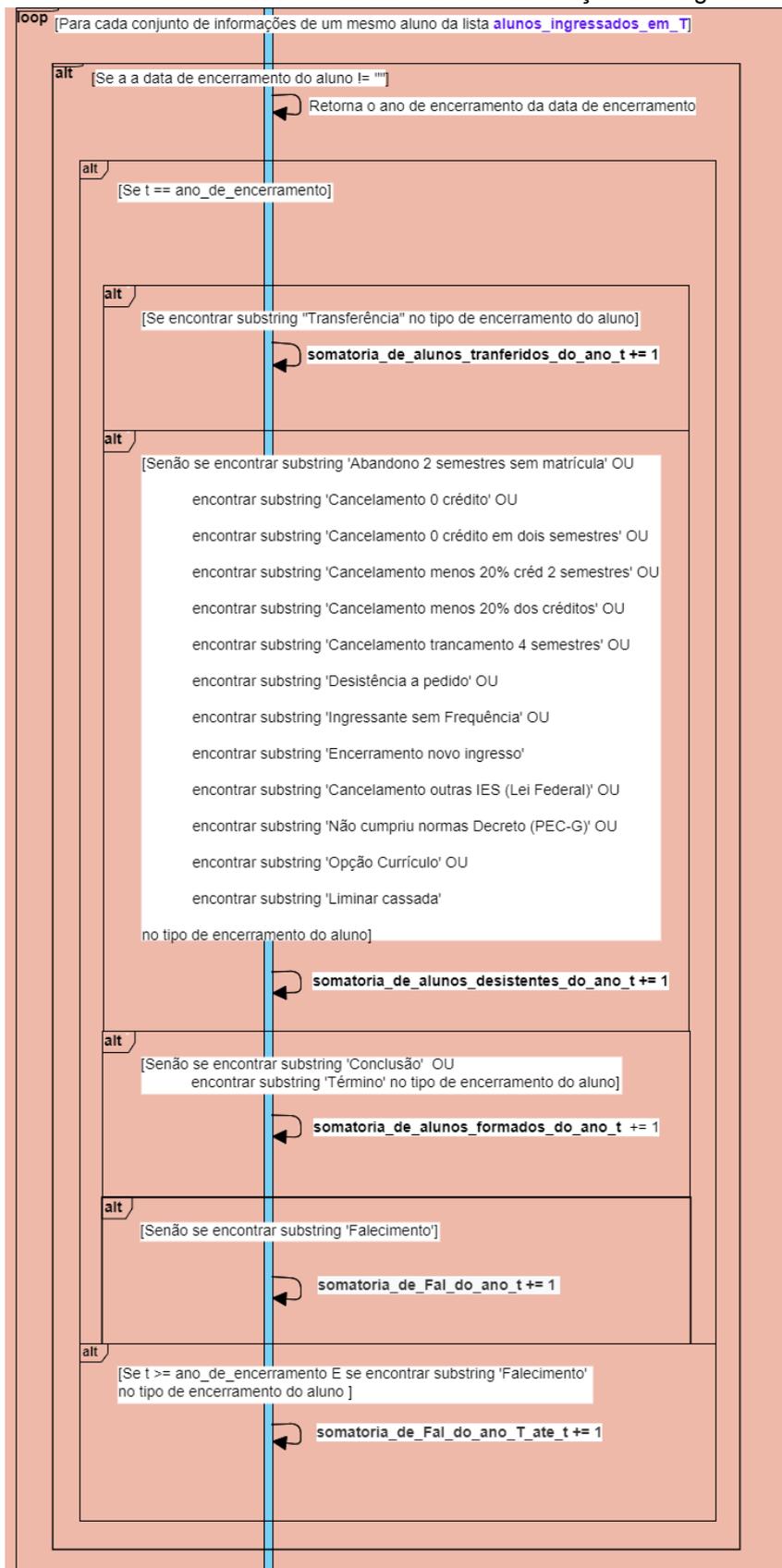
O código pode, com as adaptações devidas, como por exemplo duração do curso, períodos da formatura no ano seguinte, ser utilizado para avaliar a trajetória de alunos de outros cursos. Os dados usados e disponibilizados no GitHub não correspondem a nenhum curso da EESC.

2.5 Validação dos resultados

O resultado obtido pelo programa Python é mostrado na Figura 3 acrescida das taxas de conclusão acumulada e desistência acumulada. Para efeito de validação, o programa foi executado usando o mesmo banco de dados, empregado anteriormente para calcular os valores dos indicadores de fluxo a partir das planilhas eletrônicas. Esse procedimento foi usado para conferir os resultados com os gráficos gerados pela planilha eletrônica e validar os resultados da automatização e assim foi possível observar dados ignorados pelo programa durante a sua execução e conseqüentemente adequar a programação.

Assim, a Figura 4 apresenta o gráfico gerado a partir da planilha eletrônica dos dados sobre as taxas de permanência, taxas de conclusão acumulada e taxas de desistência acumulada sobre um período de tempo e para um curso, a qual pode ser comparada com a Figura 3 que foi gerada pelo programa Python e apresenta as mesmas taxas a respeito do mesmo período e curso.

Figura 1 - Fragmento principal do diagrama de sequência referente ao cálculo das taxas de permanência para o estudo de caso realizado em um curso da instituição ao longo dos anos.



Fonte: Elaborado pelos autores (2022).

Figura 2 - Código para a taxa de permanência.

```

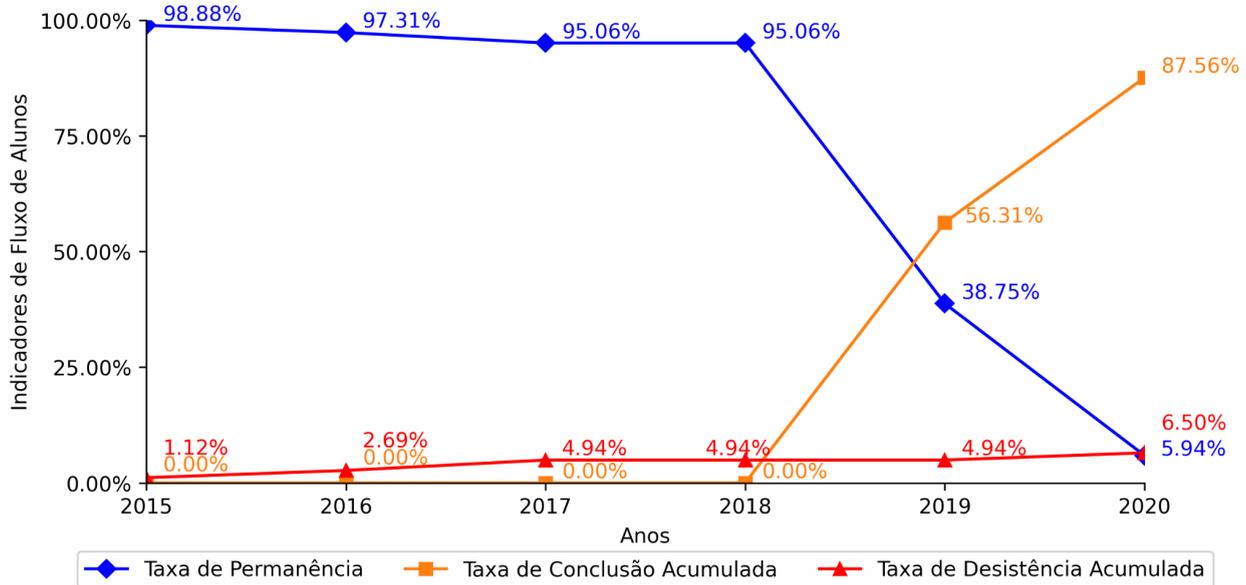
14  for ano_de_ingresso, data_de_encerramento, tipo_de_encerramento in alunos_ingressados_em_T:
15
16      if data_de_encerramento != '':
17
18          if t == data_de_encerramento.year:
19
20              if tipo_de_encerramento.find('Transferência')!= -1:
21                  somatoria_de_alunos_tranferidos_do_ano_t +=1
22
23              elif (tipo_de_encerramento.find('Abandono 2 semestres sem matrícula')!= -1 or
24                   tipo_de_encerramento.find('Cancelamento 0 crédito')!= -1 or
25                   tipo_de_encerramento.find('Cancelamento 0 crédito em dois semestres')!= -1 or
26                   tipo_de_encerramento.find('Cancelamento menos 20% créd 2 semestres')!= -1 or
27                   tipo_de_encerramento.find('Cancelamento menos 20% dos créditos')!= -1 or
28                   tipo_de_encerramento.find('Cancelamento trancamento 4 semestres')!= -1 or
29                   tipo_de_encerramento.find('Desistência a pedido')!= -1 or
30                   tipo_de_encerramento.find('Ingressante sem Frequência')!= -1 or
31                   tipo_de_encerramento.find('Encerramento novo ingresso')!= -1 or
32                   tipo_de_encerramento.find('Cancelamento outras IES (Lei Federal)')!= -1 or
33                   tipo_de_encerramento.find('Não cumpriu normas Decreto (PEC-G)')!= -1 or
34                   tipo_de_encerramento.find('Opção Currículo')!= -1 or
35                   tipo_de_encerramento.find('Liminar cassada')!= -1):
36
37                  somatoria_de_alunos_desistentes_do_ano_t +=1
38
39              elif (tipo_de_encerramento.find('Conclusão')!= -1 or
40                   tipo_de_encerramento.find('Término')!= -1):
41
42                  somatoria_de_alunos_formados_do_ano_t +=1
43
44              elif tipo_de_encerramento.find('Falecimento') != -1:
45                  somatoria_de_Fal_do_ano_t +=1
46
47              if t >= data_de_encerramento.year and tipo_de_encerramento.find('Falecimento') != -1:
48                  somatoria_de_Fal_do_ano_T_ate_t +=1

```

Fonte: Elaborado pelos autores (2022).

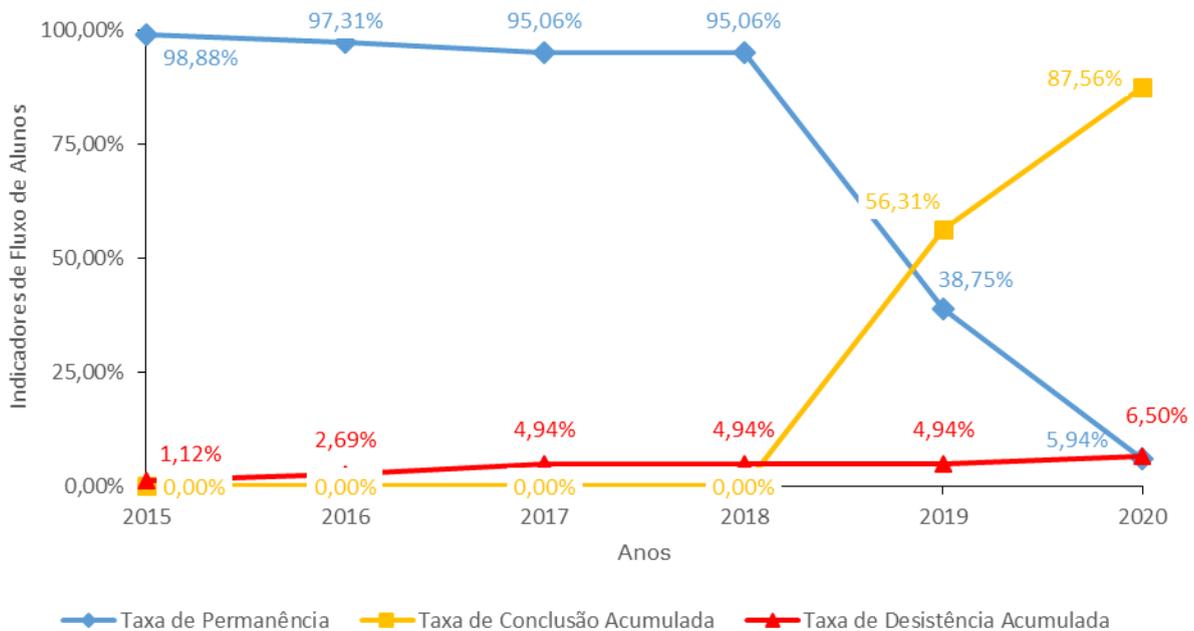
Como é possível observar pelas Figuras 3 e 4, o resultado do programa computacional configurado para realizar os cálculos de taxas de permanência pode ser usado com sucesso para substituir os resultados calculados via planilha eletrônica. Podem ocorrer pequenas discrepâncias dependendo da data das ocorrências de colação de grau de um ano, bem como de ocorrências ainda não previstas no programa computacional e erros no uso de planilhas nos gráficos gerados pela planilha eletrônica.

Figura 3 - Exemplo de gráfico obtido pelo programa Python para os indicadores de permanência, conclusão e desistência de um curso.



Fonte: Elaborado pelos autores (2022), a partir de dados da pesquisa.

Figura 4 - Exemplo de gráfico obtido via planilha eletrônica para os indicadores de permanência, conclusão e desistência de um curso.



Fonte: Elaborado pelos autores (2022), a partir de dados da pesquisa.

3 CONSIDERAÇÕES FINAIS

O trabalho realizado torna possível acompanhar de forma sistematizada a trajetória dos alunos desde o seu ingresso e assim entender as possíveis causas de evasão, servindo de ferramenta muito importante para desenvolver ações institucionais para garantir o sucesso do aluno. O fluxograma pode ser adaptado à legislação de cada escola permitindo que o programa possa ser utilizado por coordenadores de cursos de outras instituições. A seguir está previsto como continuidade do trabalho mapear o impacto de retenções ou reprovações em disciplinas de matemática na desistência e elaborar pesquisas para obter dados de dificuldades em conteúdos ministrados.

AGRADECIMENTOS

Os autores agradecem ao analista de sistemas, Masaki Kawabata Neto e ao aluno de graduação em Ciências de Computação, Aulos Plautius M. Marino, pelo auxílio no fornecimento dos dados para o estudo de casos e pelo início do desenvolvimento do programa computacional, respectivamente.

REFERÊNCIAS BIBLIOGRÁFICAS

AMBIEL, Rodolfo Augusto Matteo. Construção da escala de motivos para evasão do ensino superior. **Avaliação Psicológica**, (1):44–52, 2015.

ANDIFES, A.; ABRUEM, A.; SESU/MEC, S. Diplomação, retenção e evasão nos cursos de graduação em instituições de ensino superior públicas: resumo do relatório apresentado a ANDIFES, ABRUEM e SESu/MEC pela Comissão Especial. **Revista da Avaliação da Educação Superior**, v. 1, n. 2, 1996. Disponível em: <<http://periodicos.uniso.br/ojs/index.php/avaliacao/article/view/739>>. Acesso em: 24 abr. 2022.

BOOCH, Grady.; RUMBAUGH, James.; JACOBSON, Ivar. **UML: guia do usuário**. Rio de Janeiro: Elsevier Brasil, 2006.

COLAB. **Colaboratory: Conheça o Colab**, 2022. Página google drive. Disponível em: <https://colab.research.google.com/?utm_source=scs-index#scrollTo=Nma_JWh-W-IF>. Acesso em: 27 de abril de 2022.

CURI, Mariana *et al.*. **Estudo sobre evasão no ICMC-USP**. 2018. Relatório Técnico - Mineração Estatística de Dados aplicada à evasão de alunos dos cursos de graduação do Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo.

GUEDES, Gilleanes Thorwald Araujo. **UML 2 Uma Abordagem Prática**, 2 ed., São Paulo: Novatec editora, 2011.

INEP. **Metodologia de cálculo dos indicadores de fluxo da educação superior**. Disponível em:

https://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2017/metodologia_indicadores_trajetoria_curso.pdf. Acesso em: 21 de abril de 2022.

ALGORITHMS FOR CALCULATING STUDENT FLOW INDICATORS TO ACCESS UNIVERSITY DROPPING-OUT

Abstract: Several studies are constantly developed with the objective of studying the phenomenon of school dropout. In undergraduate courses, it is necessary to follow the academic trajectory of students to produce a diagnosis of the disconnection of courses and define actions at the institutional level to reduce dropout rates. The dropout rate calculated for the cohort of admissions is important to understand how the incoming classes behave over the years as well as to compare dropout rates in different engineering courses. In this work, student flow indicators are used to study their academic trajectories. The calculation of flow indicators is automated through a computer program created to read data from electronic spreadsheets and perform the pairing of databases year by year to calculate the sum of student links to courses by year of admission. The automation of the use of an admission cohort database by a computer program in Python makes viable the constant monitoring and embracing policies offered to students.

Keywords: Academic trajectory, admission cohort, flowchart from database, computer programming, data pairing, automation, diagnosis.