

PÓS-GRADUAÇÃO ON

DATATHON

DATA ANALYTICS

Gabriel Silva Ferreira

Gustavo Duran Domingues

Jhonny Amorim Silva

Lucas Alexander dos Santos

Sandro Semmer

ANÁLISE DE DADOS DA ASSOCIAÇÃO PASSOS MÁGICOS



Objetivo

Nosso grande objetivo nesse Datathon é atuar como uma equipe de cientista de dados, criando uma proposta preditiva dos dados da ONG (Organização Não Governamental) Passos Mágicos e também como analistas de dados, realizando uma proposta analítica de seus indicadores e métricas de ensino aos seus alunos e o impacto que a ONG causa na comunidade em que atende.

Com base no dataset disponibilizado de **Pesquisa Extensiva do Desenvolvimento Educacional (PEDE)** elaborado pela própria equipe do Passos mágicos, realizaremos a entrega das duas propostas solicitadas:

Proposta analítica: A ideia é demonstrar os impactos que a ONG “Passos Mágicos” realizou sobre a performance de estudantes e levantar indicadores de performance. Sendo assim, será feita a apresentação de **dashboards** com uma ferramenta de visualização dos dados conhecida pelo mercado - o Power BI e contaremos uma história com os dados através do **storytelling** que será levantado por meio desse documento, tudo isso a fim de auxiliar a Passos Mágicos a tomar as melhores decisões com base nos indicadores e conhecer o perfil dos estudantes.

Proposta preditiva: Ademais, é de responsabilidade de nossa equipe de dados, criar um **modelo preditivo para prever o comportamento do estudante** com base em algumas variáveis que podem ser cruciais para a identificação do desenvolvimento do estudante. Para essa proposta preditiva, utilizamos o modelo de algoritmo **XGBoost**, que é composto por árvore de decisões, que vai nos ajudar a prever o resultado por pedra para um aluno em específico.

Introdução

O acesso à educação de qualidade é um direito fundamental e um pilar para a transformação social, especialmente para jovens em situação de vulnerabilidade. É nisso que a associação Passos Mágicos emerge como exemplo de um projeto social e educacional que busca instrumentalizar o uso da educação como ferramenta para a mudança das condições de vida das crianças e jovens em vulnerabilidade social que residem no Embu Guaçu - Município de São Paulo, através da educação de qualidade, auxílio psicológico/psicopedagógico, ampliação de sua visão de mundo e protagonismo.

O presente trabalho tem como objetivo analisar o impacto da atuação da Passos Mágicos na trajetória educacional de seus estudantes. Por meio da coleta e análise de dados sobre o desempenho acadêmico, frequência e participação nas atividades oferecidas pela organização, busca-se identificar os indicadores de performance que melhor refletem o progresso dos jovens atendidos. Adicionalmente, a análise preditiva visa levantar padrões e tendências que possam subsidiar o aprimoramento das práticas pedagógicas da Passos Mágicos, contribuindo para o fortalecimento de sua missão de transformar vidas por meio da educação.

Base de dados:

A base de dados contém informações educacionais e socioeconômicas dos estudantes da Passos Mágicos no período de 2020, 2021 e 2022. Foram disponibilizadas duas bases de dados com as características de desenvolvimento educacional e questões socioeconômicas dos estudantes e um dicionário de dados com o mapeamento de todas as variáveis. Além da base de dados, alguns relatórios de pesquisa realizada pela Passos Mágicos também foram disponibilizados para auxiliar no conhecimento do negócio.

Análise dos dados de 2020

Impacto do Contexto Socioeconômico no Desenvolvimento Educacional

Em 2020, havia 727 alunos cadastrados na associação. Destes, 598 eram provenientes de escolas públicas de ensino das redes estadual e municipal de Embu Guaçu, representando aproximadamente 82.25% do total de alunos. Os 129 alunos restantes eram bolsistas de instituições privadas - essas instituições fazem parceria com a Associação Passos Mágicos, correspondendo a cerca de 17.74% do corpo discente. Esses dados indicam que a atuação da associação tem se concentrado, predominantemente, em alunos da rede pública. Essa priorização se justifica pela constatação de que estudantes da rede privada, em geral, apresentam níveis mais elevados de qualidade de vida, autopercepção de saúde, motivação para aprender e desempenho escolar, conforme aponta pesquisa realizada na Faculdade de Medicina.

É relevante destacar que fatores como o perfil socioeconômico e a raça também são considerados pelo Censo Escolar como influenciadores do desempenho acadêmico e do acesso ao ensino superior. Nesse contexto, a base de dados analisada contém uma coluna que categoriza os alunos conforme a instituição for pública ou privada. Essa informação será utilizada para verificar se esses fatores realmente refletem nos dados, possibilitando a obtenção de insights relevantes.

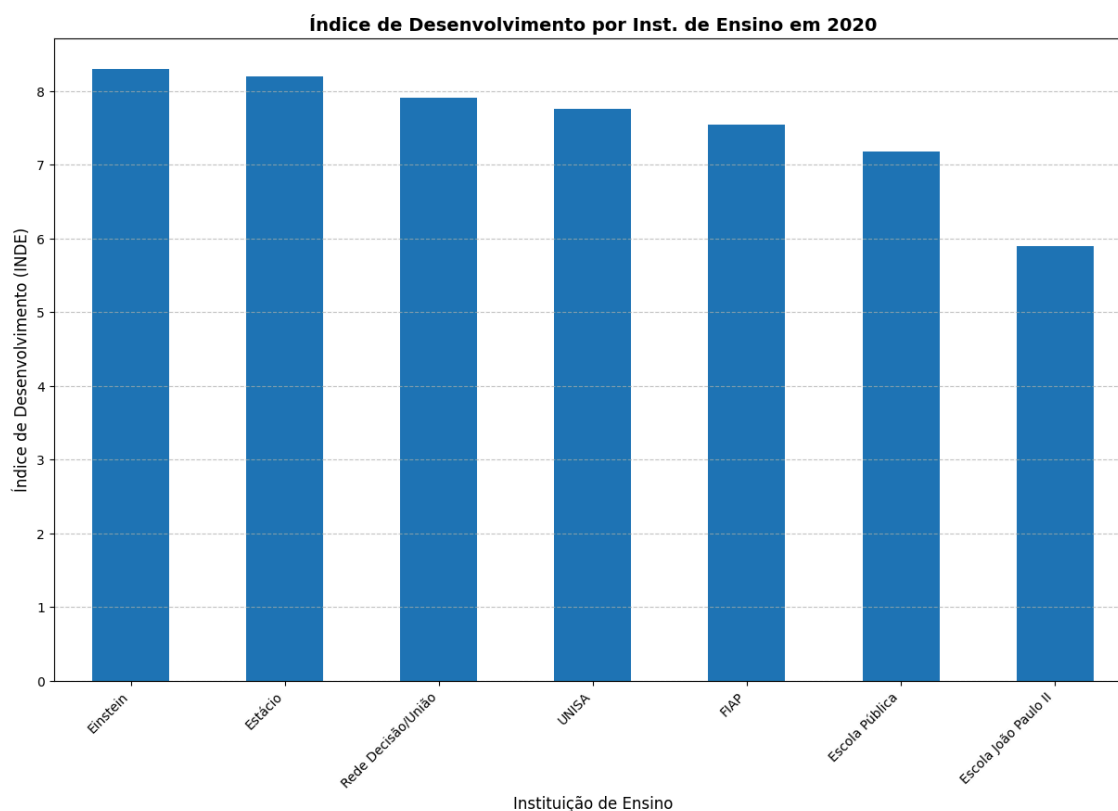
Para avaliar o desempenho dos alunos, a associação utiliza o Índice de Desenvolvimento Educacional (INDE), elaborado a partir da pesquisa PEDE (2020, 2021, 2022). O INDE sintetiza os resultados de indicadores desenvolvidos pela associação para dimensionar o progresso individual de cada aluno. Esses indicadores abrangem três dimensões principais:

(1) Dimensão Acadêmica, que incluem a entrega de lições de casa e notas de provas e o engajamento do aluno nas tarefas e serviços de voluntariado;

(2) Dimensões Psicossociais, que consideram a autoavaliação dos alunos e a avaliação de psicólogos;

(3) Dimensões Psicopedagógicas, com a avaliação de professores e pedagogos.

Inicialmente, analisaremos o Índice de Desenvolvimento da Educação (INDE) por instituição de ensino em 2020, buscando compreender o desempenho dos alunos e identificar possíveis insights para análise:

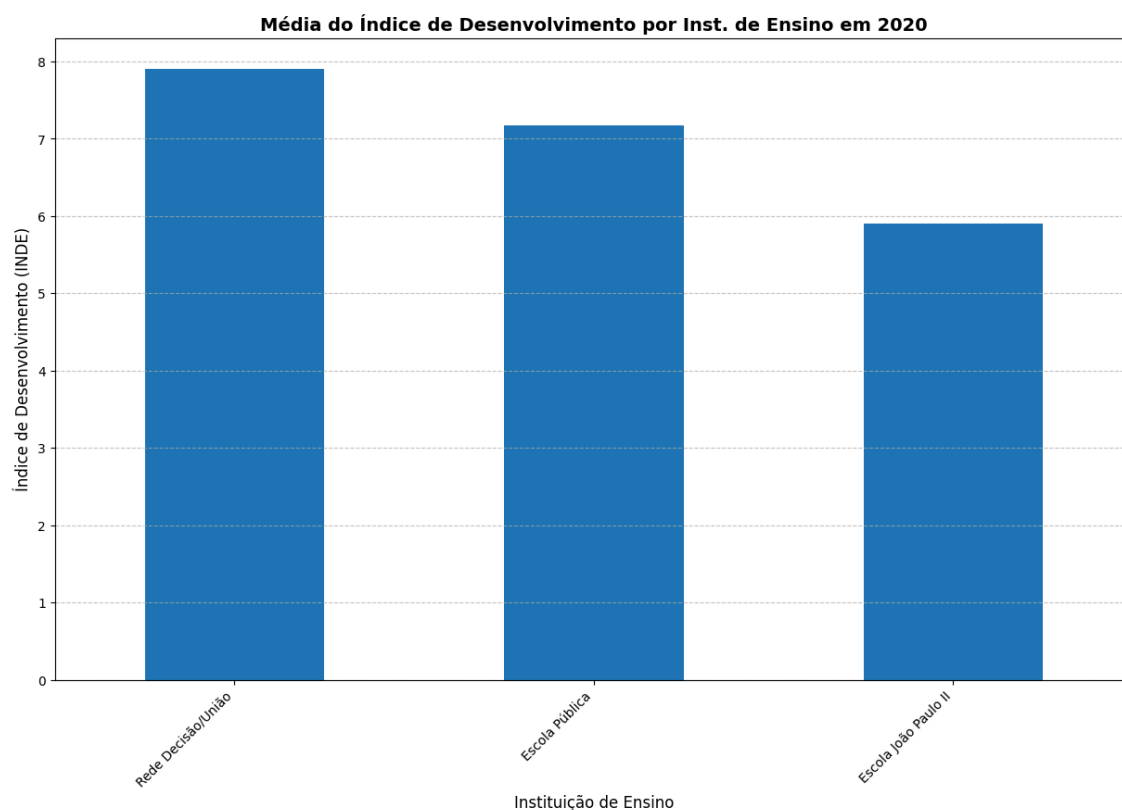


O gráfico apresentado ilustra o padrão mencionado anteriormente, evidenciando que, em geral, alunos de escolas particulares obtêm um desempenho acadêmico superior. É importante ressaltar, no entanto, que a base de dados também diferencia os alunos por turma. A associação Passos

Mágicos adota um critério de organização em fases que abrange desde o ensino fundamental (2º e 3º ano) até o ensino superior (nível 8, conforme nomenclatura da pesquisa PEDE, onde se encontram os universitários).

Entre os universitários, 11 são da FIAP, 10 da UNISA, 3 da Instituição Einstein e 2 da Estácio.

Ao separarmos os universitários dos alunos dos demais níveis, identificamos que 598 são alunos da rede pública, 101 alunos da rede privada Decisão/União e mais 2 alunos da Escola particular João Paulo II. Assim, podemos analisar o desempenho por nível de ensino e por tipo de escola de origem.

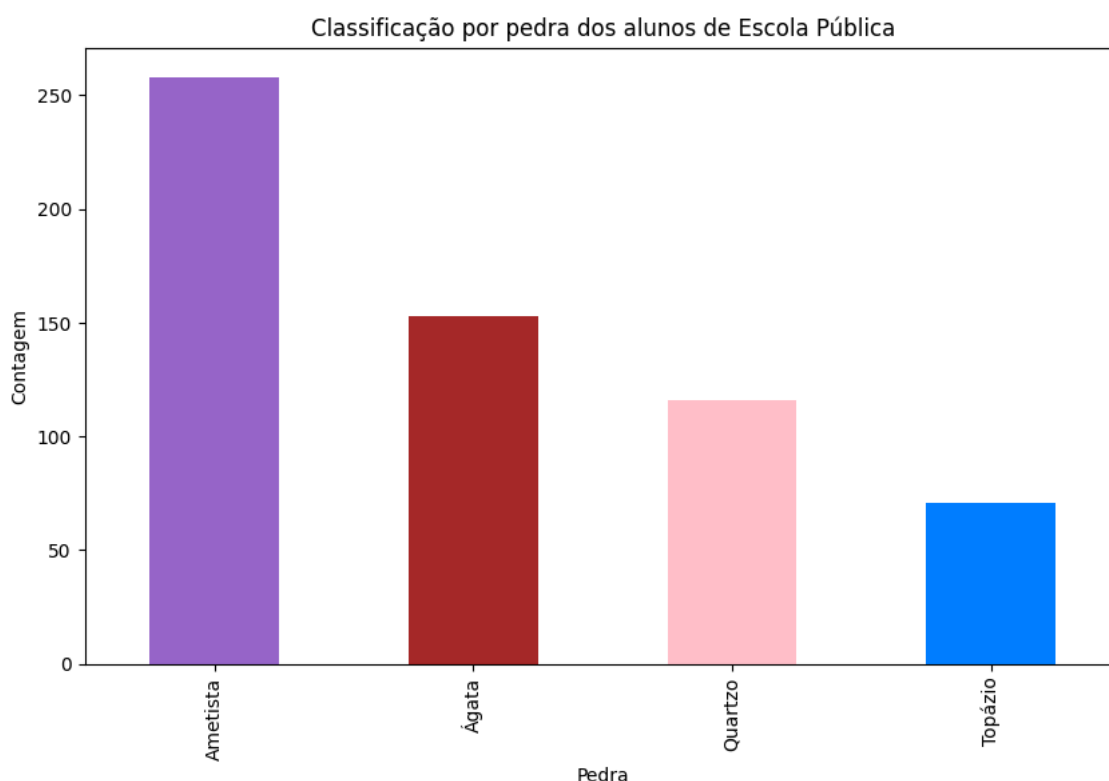


E da mesma forma podemos ver que a rede privada ultrapassa a instituição pública, com exceção da escola João Paulo II, cujos dados são pouco representativos devido ao número reduzido de alunos.

Existe também uma classificação por aluno baseado no número do INDE que caracteriza os resultados através de pedras preciosas, considerando a classificação como a mais valiosa, a acima da média, a média e a menos valiosa. Esse conceito de classificação é dado entre as pedras:

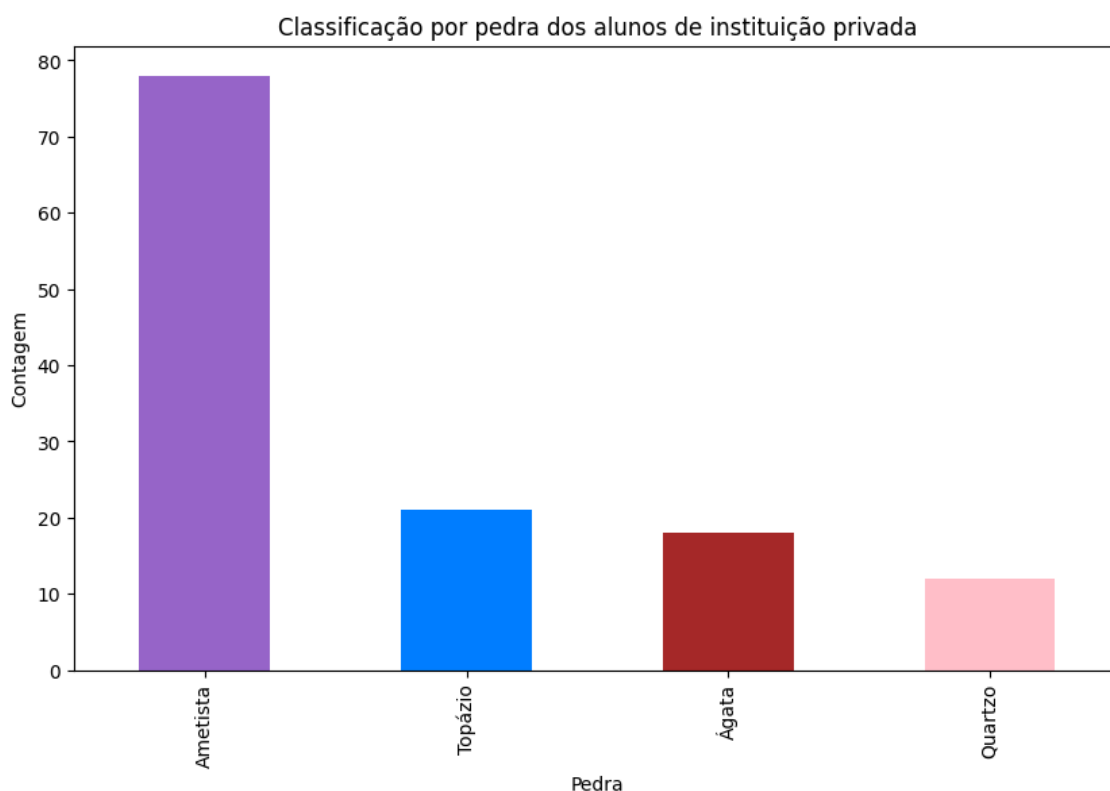
- **Quartzo** – 2,405 a 5,506 (número do INDE);
- **Ágata** – 5,506 a 6,868;
- **Ametista** – 6,868 a 8,230;
- **Topázio** – 8,230 a 9,294.

Se considerarmos esse método de classificação de acordo com as pedras teremos a seguinte distribuição:



A distribuição dos 258 alunos da escola pública por classificação de desempenho foi a seguinte: 43,14% no grupo Ametista, 25,58% no grupo Ágata, 19,39% no grupo Quartzo e 11,87% no grupo Topázio. Ao analisarmos

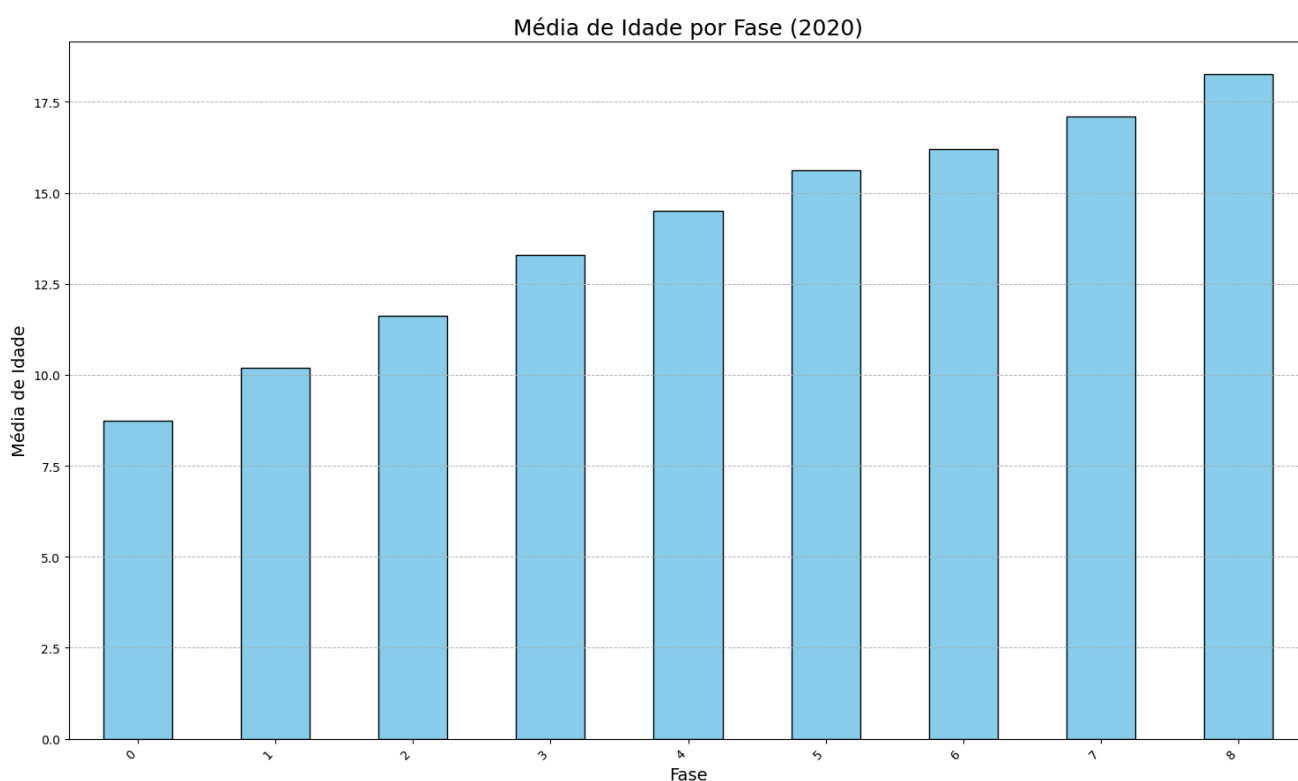
os grupos Ametista e Topázio, que representam os alunos com melhor desempenho e com a melhor pedra, observamos que eles somam cerca de 55% do total. Em contrapartida, aproximadamente 45% dos alunos apresentaram resultados inferiores.



A análise dos 129 alunos bolsistas revela que 78 deles (60,46%) foram categorizados no grupo Ametista, o que representa a maioria expressiva. Em seguida, temos 21 alunos (16,27%) no grupo Topázio, 18 alunos (13,95%) no grupo Ágata e 12 alunos (9,30%) no grupo Quartzo. Um dado relevante é que os grupos de melhor desempenho, Ametista e Topázio, juntos, englobam 99 alunos, ou seja, cerca de 77% do total de bolsistas. Já os 30 alunos restantes, que representam 23%, tiveram um desempenho inferior.

Separação por fase

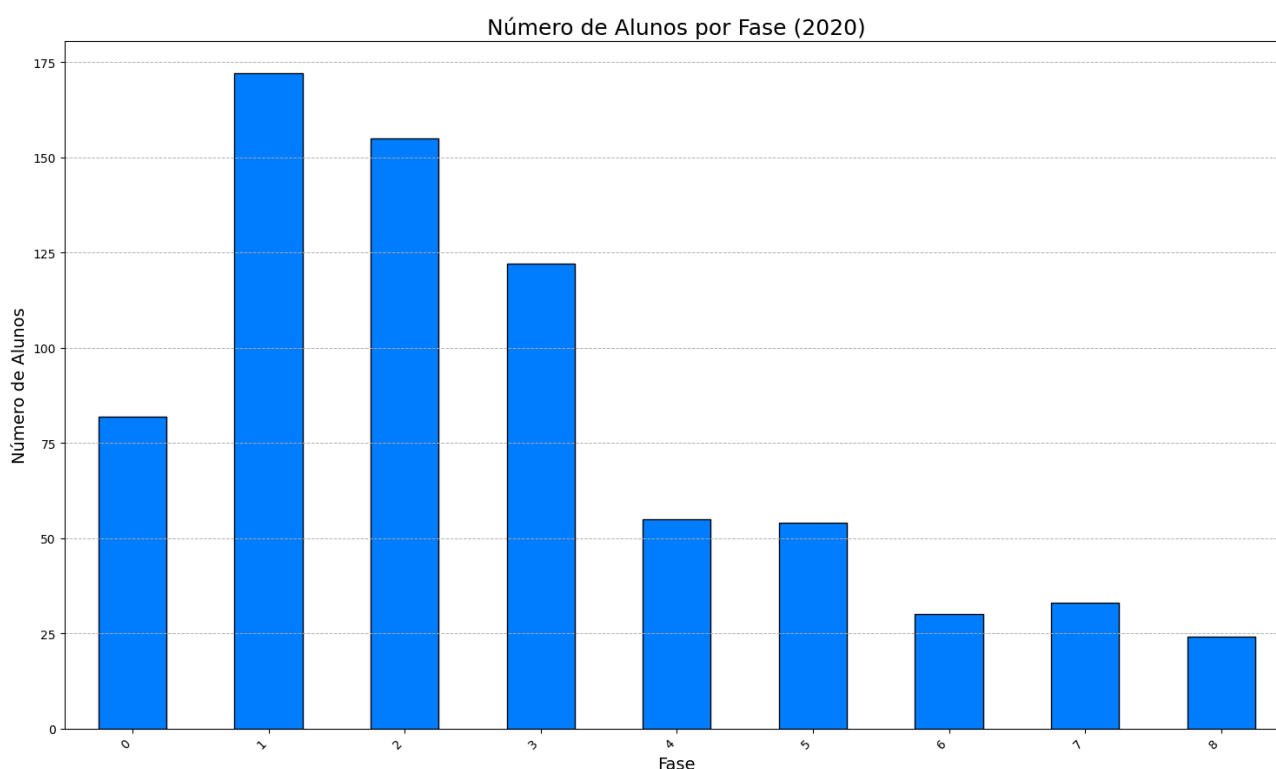
A separação dos alunos por fase, que abrange do nível 0 ao 8, revela diferentes estágios de desenvolvimento. O nível 0, denominado alfa, corresponde aos alunos do 2º ou 3º ano do ensino fundamental, com uma faixa etária média de 8 anos. No nível 1, encontram-se os alunos do 4º ano, com idade média de 10 anos. O nível 2 abrange os estudantes do 5º e 6º ano, com média de 11 anos. O nível 3 engloba os alunos do 7º e 8º ano, com uma média de 13 anos. Já o nível 4 é composto pelos alunos do 9º ano, com idade média de 14 anos. O nível 5 corresponde ao 1º ano do ensino médio, com alunos em torno de 15 anos. Os níveis 6 e 7 abrangem o 2º e 3º ano do ensino médio, com idades médias de 16 e 17 anos, respectivamente. Por fim, o nível 8 é destinado aos universitários, com 18 anos ou mais.



Cada aluno é alocado em uma fase e turma, e espera-se que sua faixa etária corresponda ao nível ideal de ensino. Caso contrário, o aluno será considerado em defasagem. No entanto, se a idade do aluno for inferior à esperada para sua fase, ele apresentará um avanço em relação à progressão padrão.

Por exemplo, em 2020, foram identificados 8 alunos com 16 anos alocados na fase 4 (9º ano, cuja faixa etária média é de 14 anos), e 6 alunos com 11 anos na fase alfa (2º e 3º ano, com média de 8 anos). Estes são casos de alunos com defasagem em relação ao nível ideal para suas idades. Em contrapartida, um aluno de 15 anos, que deveria estar no ensino médio (nível 5), foi alocado no nível 8, destinado a universitários, caracterizando um avanço no seu ensino acadêmico.

Em 2020, a distribuição de alunos por fase apresentou a seguinte configuração: a fase 0 teve 82 alunos; a fase 1 concentrou 172 alunos; a fase 2, 155; fase 3, 122; a fase 4, 55; a fase 5, 54; a fase 6, 30; a fase 7, 33 e a fase 8, destinada a universitários, 24 alunos. Essa distribuição evidencia a concentração de alunos nas fases iniciais do ensino fundamental e um decréscimo gradual nas fases subsequentes, culminando na menor quantidade de alunos no nível universitário.



Isso oferece uma organização dos estudantes na instituição Passos Mágicos no ano de 2020. No entanto, a compreensão plena da dinâmica de ensino-aprendizagem requer uma análise longitudinal, abarcando os anos subsequentes. Dessa forma, pretendemos discorrer sobre os dados dos anos posteriores, examinando também as performances dos alunos e outros indicadores valiosos. Essa análise comparativa possibilitará a identificação de tendências, a avaliação da eficácia das práticas pedagógicas e, crucialmente, a tomada de decisões, visando o aprimoramento contínuo da qualidade do ensino e a progressão adequada de seus alunos.

Análise dos dados de 2021

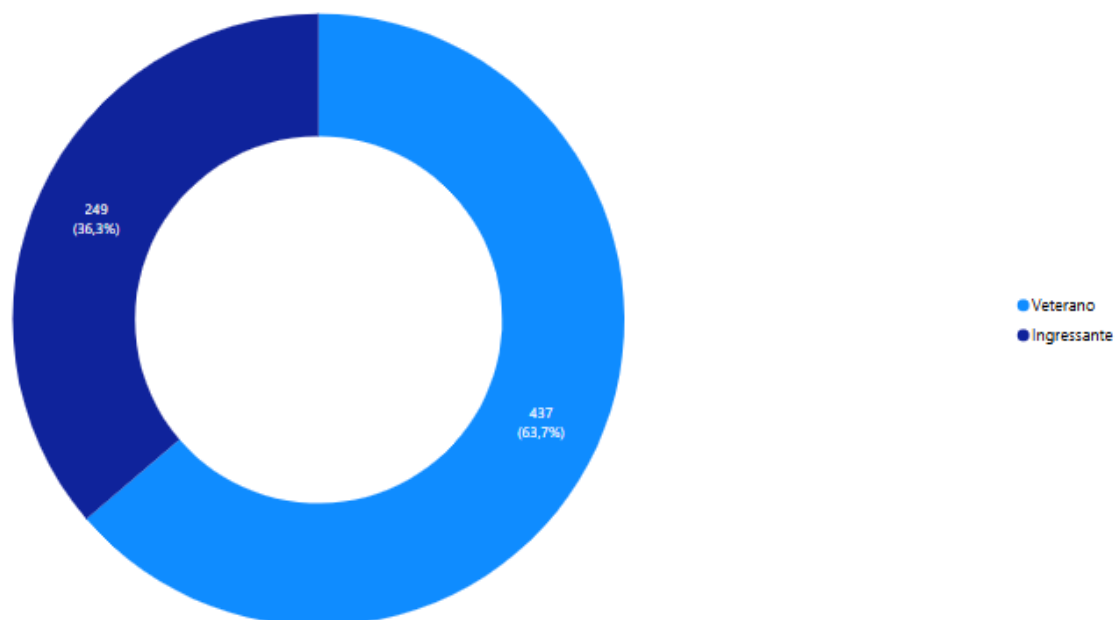
Comparativa entre os anos 2020 e 2021

Para o ano de 2021, são 684 alunos matriculados. Destes, 557 são de escola pública e 129 de particular, mesmo número do ano anterior de 2020. Mas para aqueles matriculados da rede estadual ou municipal houve uma queda nos números.

De acordo com o relatório PEDE de 2021, gerado e disponibilizado pela Passos Mágicos, uma das principais causas da evasão dos alunos nesse período pode ter sido altamente influenciado devido a pandemia que aconteceu no ano vigente:

"Os resultados de evasão podem ter sido altamente influenciados pela pandemia COVID-19, em 2020 e 2021, que forçou a Associação a realizar todas as suas atividades educacionais e culturais de forma não-presencial." - PEDE 2021 (pg. 10).

Apesar da baixa, no ano de 2021 houve 249 alunos ingressantes e 437 eram considerados veteranos.



Ingressantes e Veteranos

Entre os 249 alunos ingressantes, 237 eram provenientes de escola pública, enquanto 12 eram bolsistas de escolas privadas.

No que se refere aos alunos veteranos (matriculados em 2020 ou em anos anteriores), 320 são de escola pública, ao passo que 117 eram bolsistas de escolas privadas.

Fases de ensino

A maior parte dos ingressantes foi na fase 0, totalizando 93 alunos, o que corresponde a 37,35% do total. A fase 1 recebeu 48 alunos (19,28%); a fase 3, 41 alunos (16,74%); a fase 2, 23 alunos (9,24%); a fase 5, 19 alunos (7,63%); a fase 4, 18 alunos (7,23%); a fase 6, 5 alunos (2,01%); e a fase 7, 2 alunos. Essa

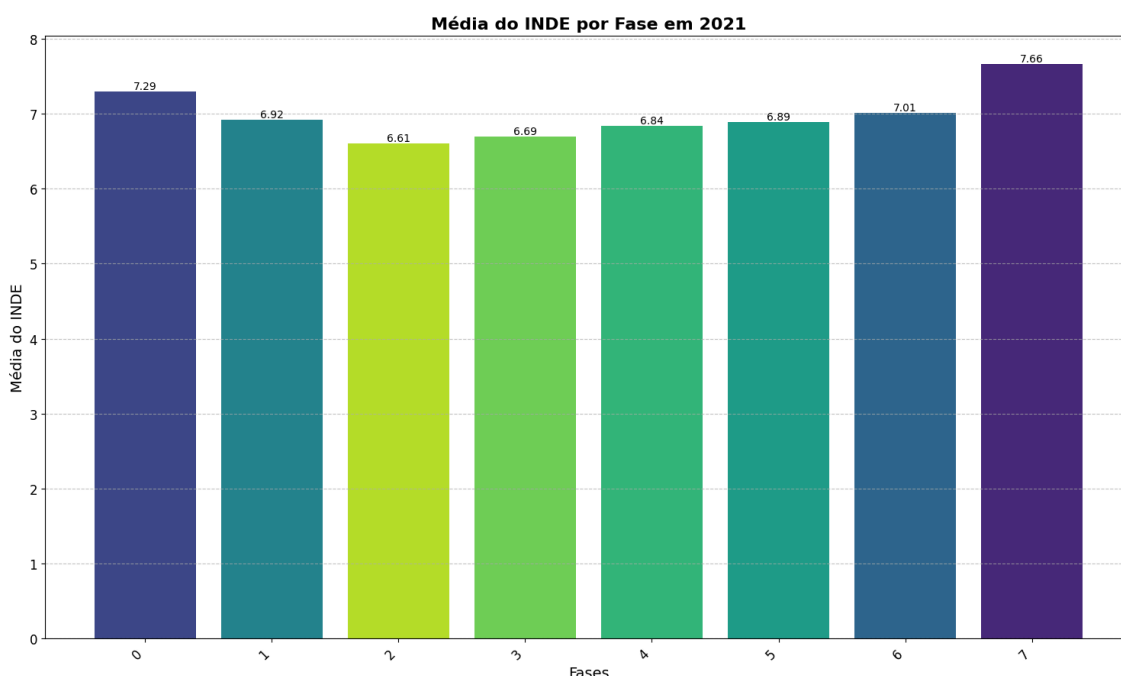
distribuição evidencia que a maior demanda se concentra nas fases mais básicas do ensino, enquanto as últimas fases apresentam uma demanda significativamente menor.

Entre os veteranos, a maior concentração estava na fase 2, com 139 alunos (31,81%), seguida pela fase 1, com 88 alunos (20,14%). A fase 3 contava com 74 alunos (16,93%); a fase 4, com 41 alunos (9,38%); a fase 5, com 31 alunos (7,09%); a fase 0, com 29 alunos (6,64%); a fase 6, com 18 alunos (4,12%); e a fase 7, com 17 alunos (3,89%).

INDE em 2021

No quesito de performance dos alunos em 2021, podemos ver o Índice de Desenvolvimento sobre diferentes visões dos dados:

INDE por fase



É possível notar que algumas fases têm médias significativamente maiores que outras. Sendo a fase 7 a maior média do INDE (7.66) com os alunos do último

ano do ensino médio com maior destaque. Nessa fase a maioria dos alunos são veteranos, ou seja, já estavam na associação Passos Mágicos, e conta-se apenas com 2 alunos novos para essa fase, do total de 19 alunos no ano de 2021. Esses alunos geralmente possuem a maior carga horária de estudos devido a fase de pré-vestibulares que se encontram, ocasionando melhores resultados em performances estudantis.

Em contrapartida, os estudantes da fase 2 e 3 tiveram médias menores e possuem o maior número de estudantes nessa fase, sendo 162 alunos e 115 alunos respectivamente. Isso levanta um ponto de atenção e análise para esse grupo.

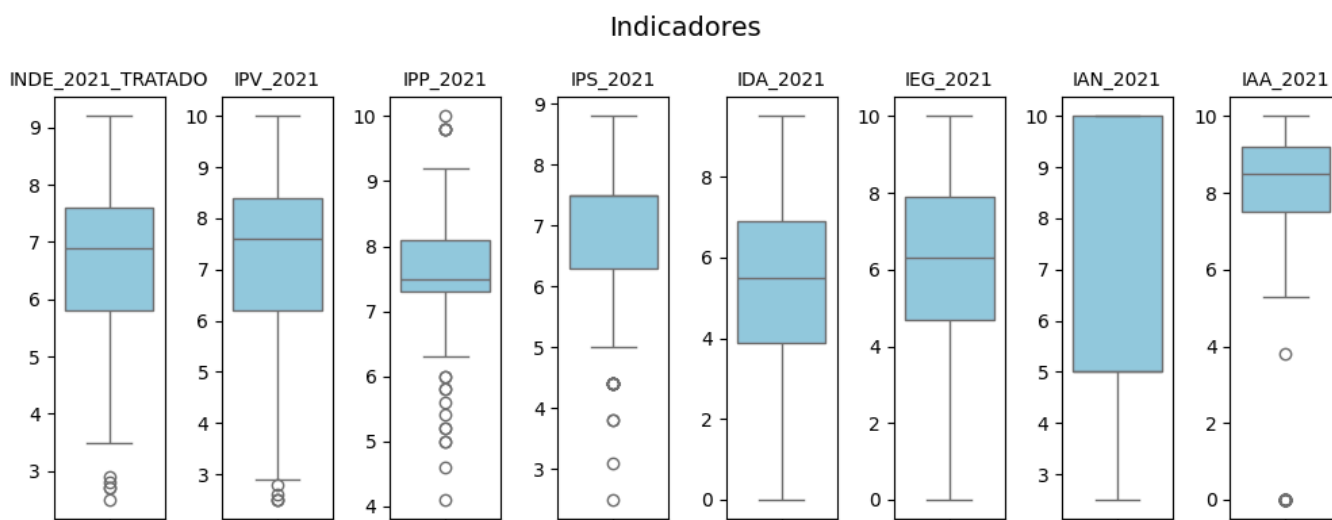
Nível 2 e 3 são os alunos do 5º ano, 6º, 7º e 8º ano. Esses alunos podem ter uma maior dificuldades e carências no que se refere ao aprendizado, entrega de lições e participações de atividades sociais.

Para entender o que influenciou nessa queda das médias do INDE dessas fases, precisa-se entender sobre as métricas utilizadas na Associação que resulta e define o desempenho dos alunos no final.

Indicadores de dimensões acadêmicas, psicossociais e psicopedagógicas:

- **IAA** - Indicador de Autoavaliação
- **IEG** - Indicador de Engajamento
- **IAN** - Indicador de Adequação ao Nível
- **IDA** - Indicador de Desempenho Acadêmico
- **IPS** - Indicador Psicossocial
- **IPP** - Indicador Psicopedagógico
- **IPV** - Indicador de Ponto de Virada

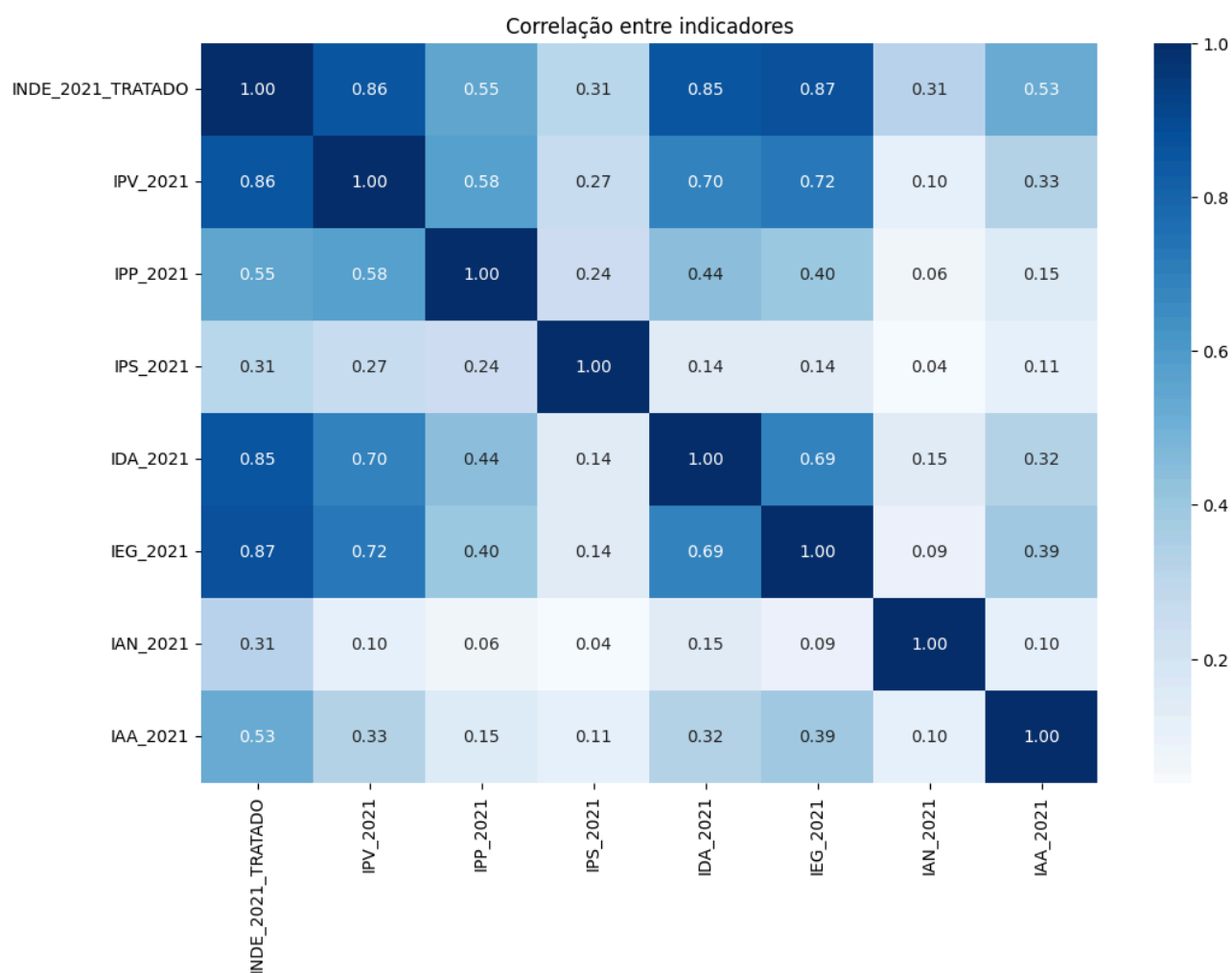
Índices de desempenhos acadêmicos e psicossociais dos alunos das fases 2 e 3 em 2021:



A média das duas fases foi pior entre os indicadores IDA (5,2), que de acordo com o Relatório PEDE, para as turmas de fases 0 a 7, trata-se de notas de provas ou exames padronizados de avaliação interna das disciplinas oferecidas pela Associação. E também o INDE aparece com uma média menor (6.6), como já visto anteriormente, e como analisado, a média entre os demais indicadores fica em torno da nota 7 ou 6 (IAA - 7.6; IAN - 7; IEG - 6.1; IPS - 6.9; IPP - 6.9; IPV - 7.2). Em destaque ao IEG, que é o índice de engajamento, calculado conforme como o aluno se engaja nas tarefas curriculares requeridas em cada uma das disciplinas e/ou seu engajamento em atividades de voluntariado desenvolvidas na Associação.

Entre esses dois indicadores - IDA e IEG, podemos entender que existe uma correlação forte entre alguns dos indicadores que influenciam no resultado final do INDE - Desempenho Educacional dos alunos.

Correlação entre as variáveis de indicadores com o INDE:



Podemos ver através da matriz de correlação acima onde mostra quão fortemente os indicadores se relacionam. Valores próximos de 1 (ou -1) indicam uma forte correlação positiva (ou negativa), valores próximos de 0 indicam uma correlação fraca.

Forte Correlação Positiva: O INDE apresenta fortes correlações positivas com:

- **IPV** (0.86): indicador Ponto de Virada, que se trata de avaliação dos pedagogos e professores;
- **IDA** (0.85): Isso mostra porque os alunos da fase 2 e 3 que tiveram um baixo índice de desenvolvimento acadêmico tiveram uma nota baixa no INDE;
- **IEG** (0.87): Mesmo motivo do IDA, o IEG tem forte correlação com o resultado do INDE;
- **IAA** (0.53): Indicador de Autoavaliação, feito através de um questionário. Tem correlação positiva, mas mais moderada do que as anteriores.

Correlação Moderada:

IPP (0.55): Indicador psicopedagógico que também é uma avaliação dos psicopedagogos, mas dessa vez mais centrada no comportamento e desenvolvimento do aluno dentro do aprendizado estipulado;

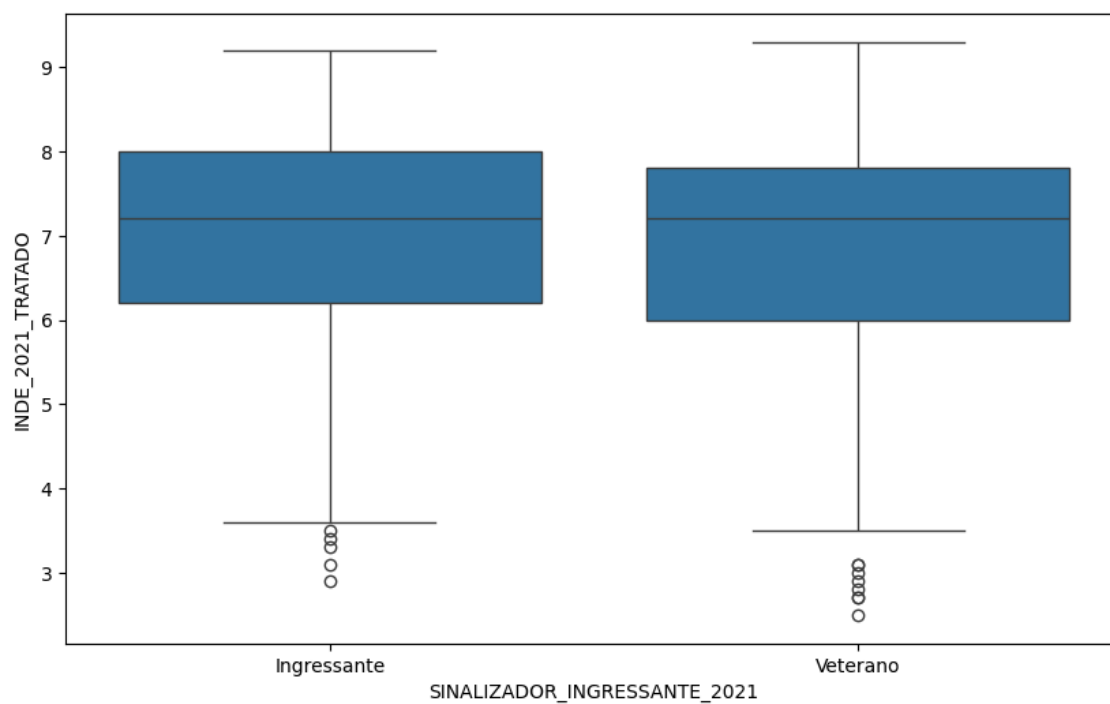
IAN (0.31): Indicador de Adequação ao Nível, que equivale a aprendizagem de acordo com a fase do aluno.

Correlação Fraca:

IPS (0.31): o Indicador Psicossocial é o mais fraco de todos. Que é a avaliação da equipe de psicólogas para caracterizar se o aluno tem tido um bom desempenho pessoal e comportamental.

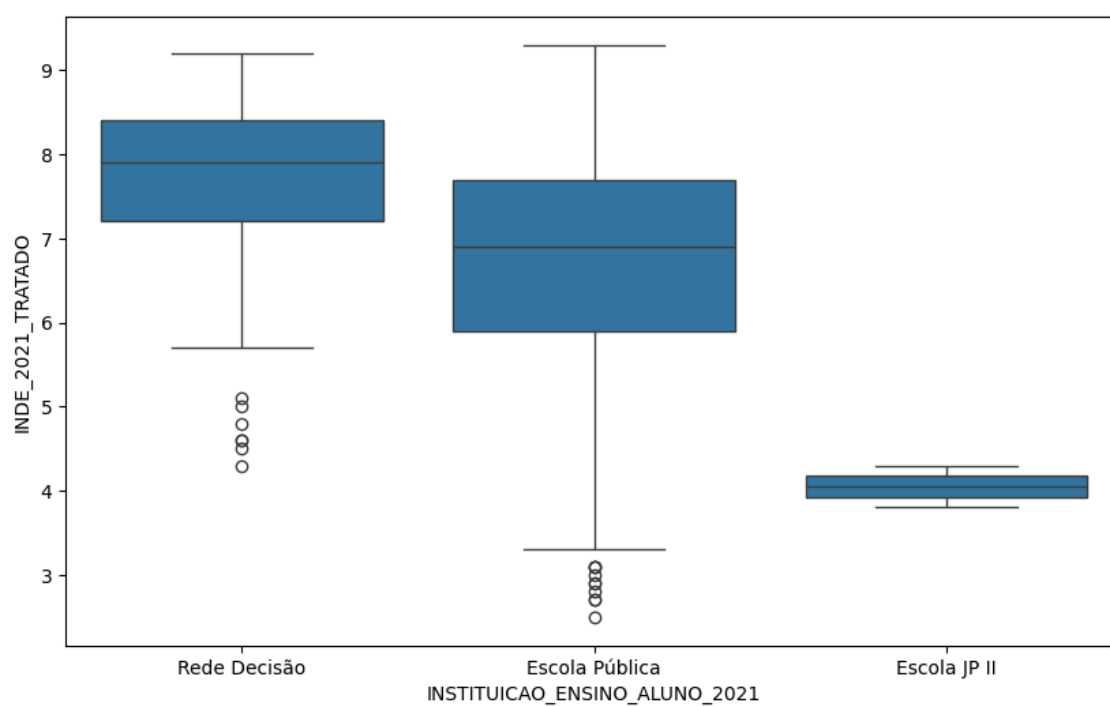
INDE por alunos ingressantes e veteranos

Há 435 veteranos e 249 alunos ingressantes no ano de 2021. A média do INDE entre eles ficou entre 6.8 e 6.9 respectivamente.



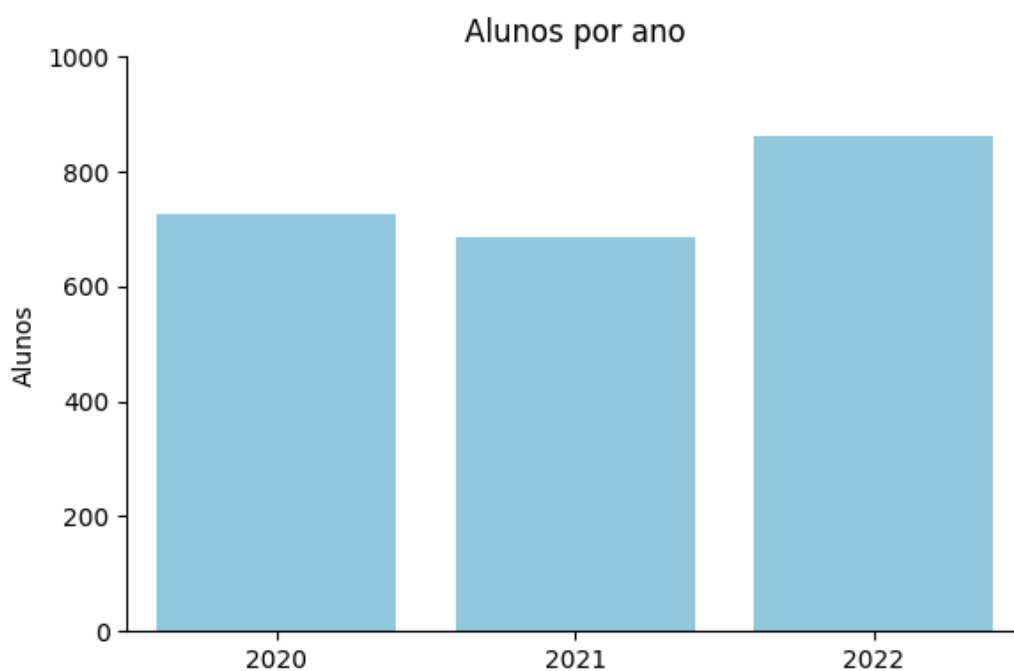
INDE por instituição de ensino

A Escola João Paulo II tem somente 2 alunos, pouca significância em comparar e analisar suas estatísticas de performance. Porém a média do INDE entre a Escola Pública e bolsistas de instituição privada é de 6.7 e 7.6 respectivamente.

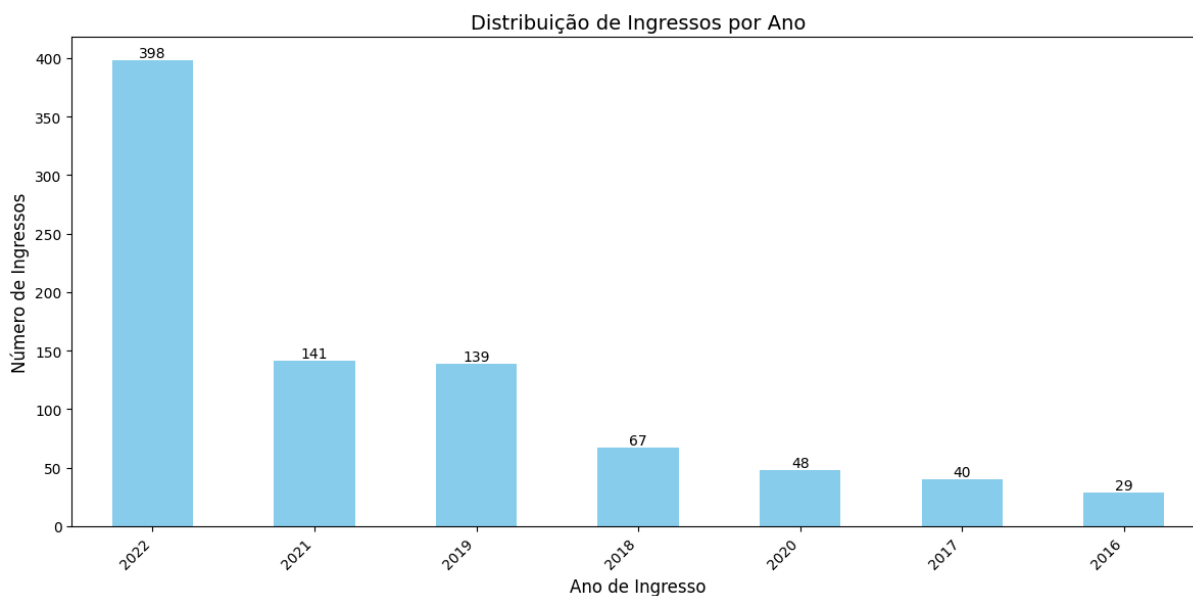


Análise dos dados de 2022

No ano de 2022 temos o total de 862 alunos matriculados na Associação. Um aumento de 25% comparado com o ano anterior. Comparado com os dois anos anteriores, o ano de 2022 demonstra um aumento significativo no total de alunos:



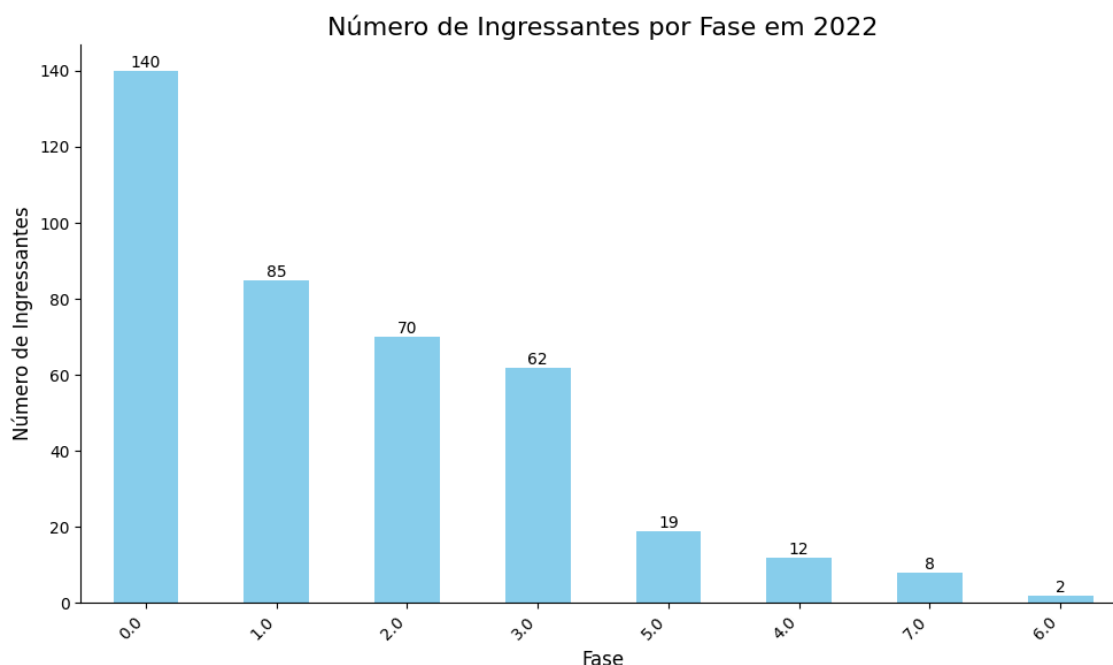
Na base de 2022 temos maiores detalhes acerca do ano que o aluno ingressou na associação, desde 2017 (um ano após a ONG ter começado a ampliar seu projeto de transformação da educação aos jovens) até 2022.



Isso mostra como o ingresso de novos alunos tem aumentado significativamente ao decorrer dos anos desde a iniciação do projeto. Apesar da baixa em 2020 (pode ser devido a pandemia de COVID-19) logo se recuperou em 2021 e teve um “boom” em 2022, denotando a popularidade que a iniciativa tem ganhado entre os cidadãos de Embu Guaçu. Considerando essa distribuição, é importante manter e/ou implementar estratégias para atrair um bom número de ingressos a cada ano e manter o projeto cada vez mais célebre.

Ingressantes por fase

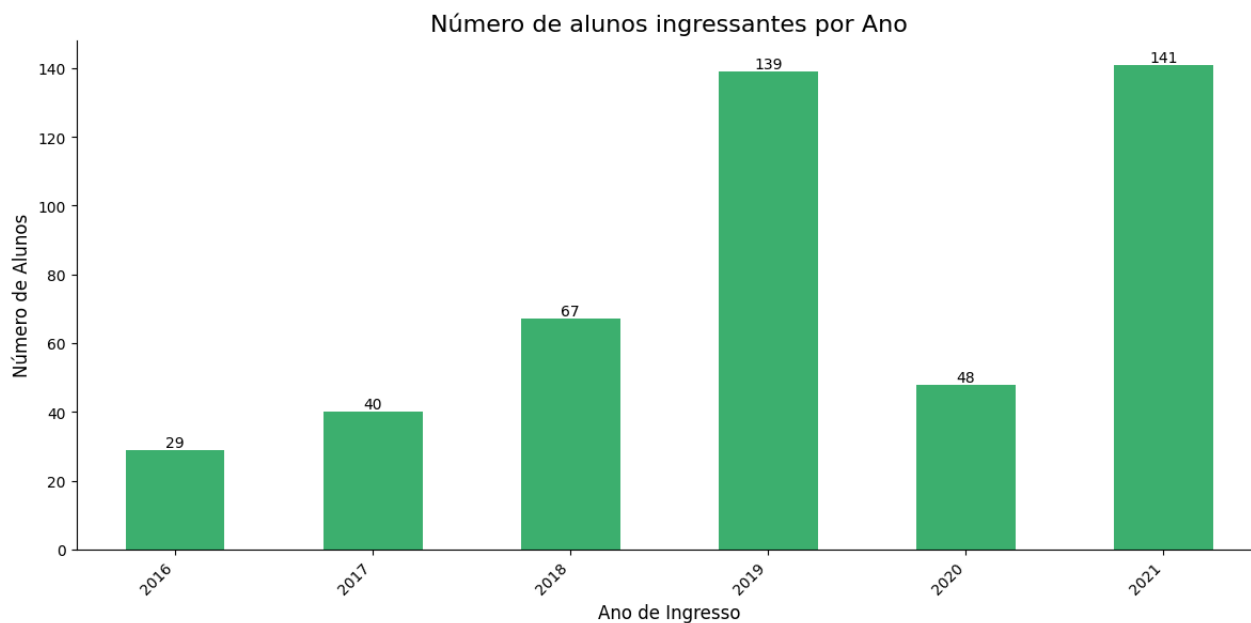
Quando analisamos os ingressantes no ano de 2022 por fase, temos a seguinte distribuição:



Como já analisado em anos anteriores, as fases iniciais são as que mantêm mais alunos por turma e, são as fases que também mais recebem alunos ingressantes na Associação.

Quantidade de alunos veteranos no ano de 2022

Do total de 862 alunos, com 398 ingressantes no ano de 2022, logo teremos o montante de 464 alunos veteranos. Destes alunos veteranos, 141 são ingressados de 2021, 48 do ano de 2020, 139 do ano de 2019, 67 do ano de 2018, 40 do ano de 2017 e 29 do ano de 2016.



No ano de 2019 (pré-pandemia) houve números expressivos de alunos ingressantes e que se mantém em 2022. Apesar da queda em 2020 devido a pandemia, logo se recuperou com o ingresso de alunos em 2021 que mantiveram-se associados ao Passos Mágicos no ano de 2022.

Como só temos os dados para análise dos anos de 2020 e 2021, podemos analisar o total de alunos com os alunos que se mantiveram no ano de 2022, utilizando o cálculo da taxa de retenção, que seria:

*(Número de alunos remanescentes em 2022 / Número de alunos no ano de ingresso) * 100*

Cálculo das Taxas de Retenção (2020 e 2021):

- 2020: $(48 / 727) * 100 = 6.60\%$
- 2021: $(141 / 684) * 100 = 20.61\%$

Conclui-se que apenas 6.60% dos alunos que ingressaram em 2020 permaneceram na escola até 2022. Isso sugere uma alta taxa de evasão por conta da pandemia. No entanto, 20.61% dos alunos que ingressaram em 2021 permaneceram na escola até 2022. Embora melhor que 2020, ainda indica uma taxa de retenção relativamente baixa.

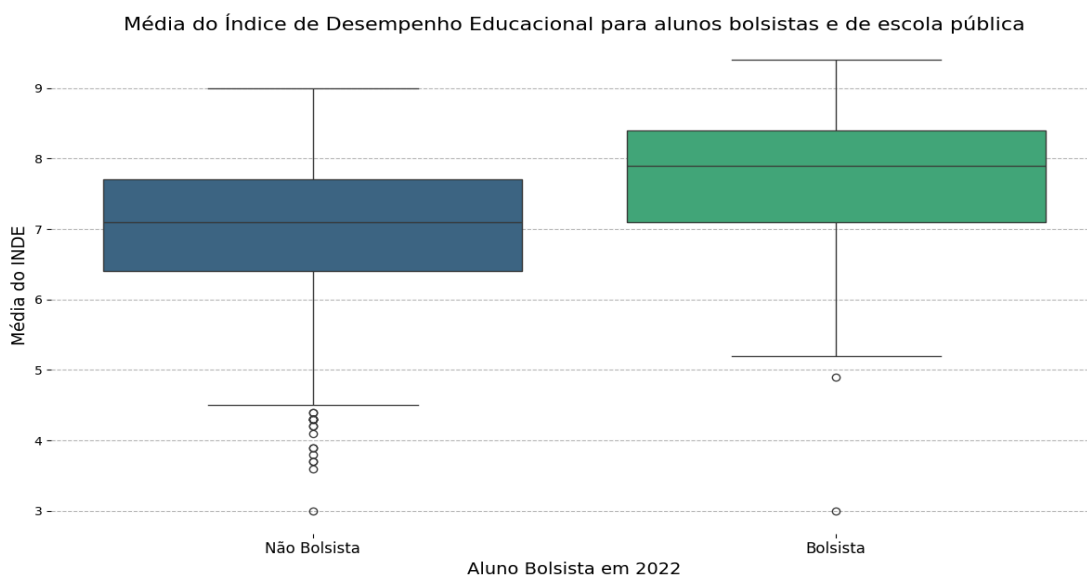
Quantidade de alunos por fase

Em 2022, a distribuição de alunos por fase apresentou uma predominância nas fases iniciais. A fase 0 e a fase 1 concentraram, cada uma, cerca de 22% dos alunos (190 e 192, respectivamente). As fases 2 e 3 também tiveram um número expressivo de alunos, com 17,98% (155 alunos) e 17,16% (148 alunos), respectivamente. As fases seguintes mostraram uma diminuição na quantidade de alunos, com 8,8% (76 alunos) na fase 4, 6,96% (60 alunos) na fase 5, 2,08% (18 alunos) na fase 6 e 2,66% (23 alunos) na fase 7. Mais uma vez, não há alunos para a fase 8 de universitários.

Alunos Bolsistas

Em 2022, foi levantado alunos que são ou não bolsistas, diferente dos anos anteriores que foi por Instituição de ensino - público, ou bolsista de instituição privada. Do total de 862 alunos, 754 não eram bolsistas (87,47%) e 108 eram bolsistas (12.52%).

Resultado médio do INDE para Alunos bolsistas x Escola pública



Os alunos não bolsistas de escola pública tiveram a média de 6.93 enquanto alunos bolsistas tiveram 7.69 de média do INDE.

Alunos Indicados a Bolsa

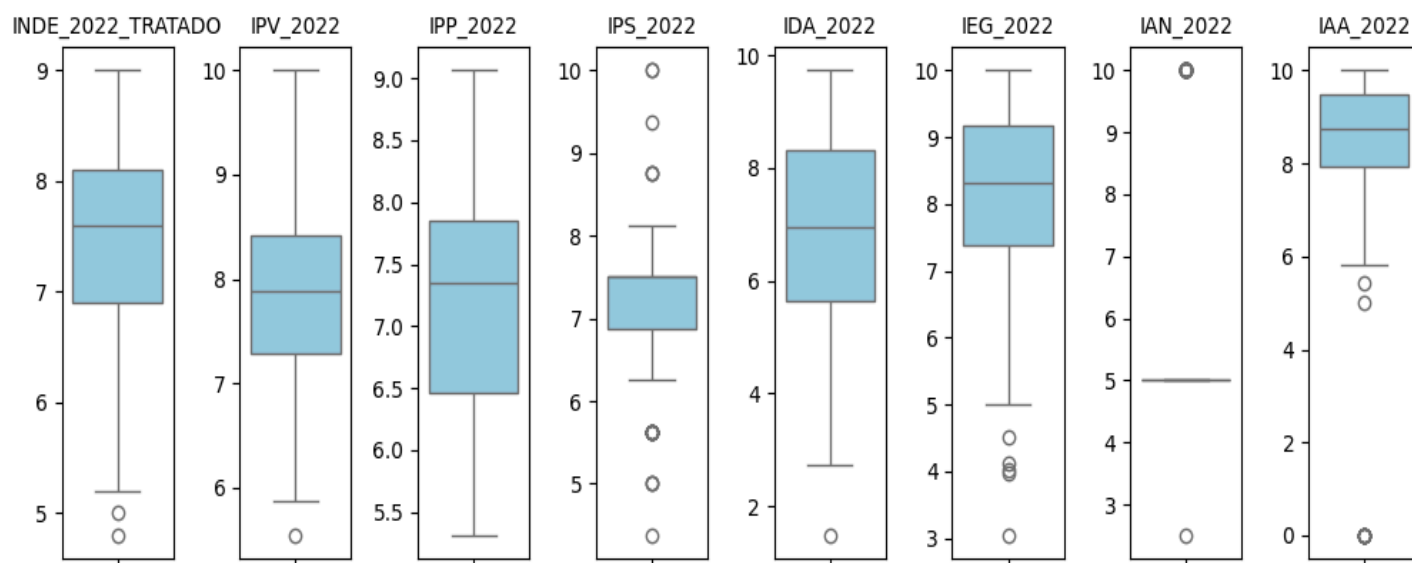
Ademais, no ano de 2022, como resultado da análise de desempenho acadêmico e avaliações pedagógicas realizadas, 132 alunos não bolsistas foram indicados a bolsas de estudo.

Resultado médio do INDE para Alunos indicados à bolsa

Certamente, a média do INDE dos alunos indicados a uma bolsa foi elevada, correspondendo a média de 7.41 para o total de 132 alunos.

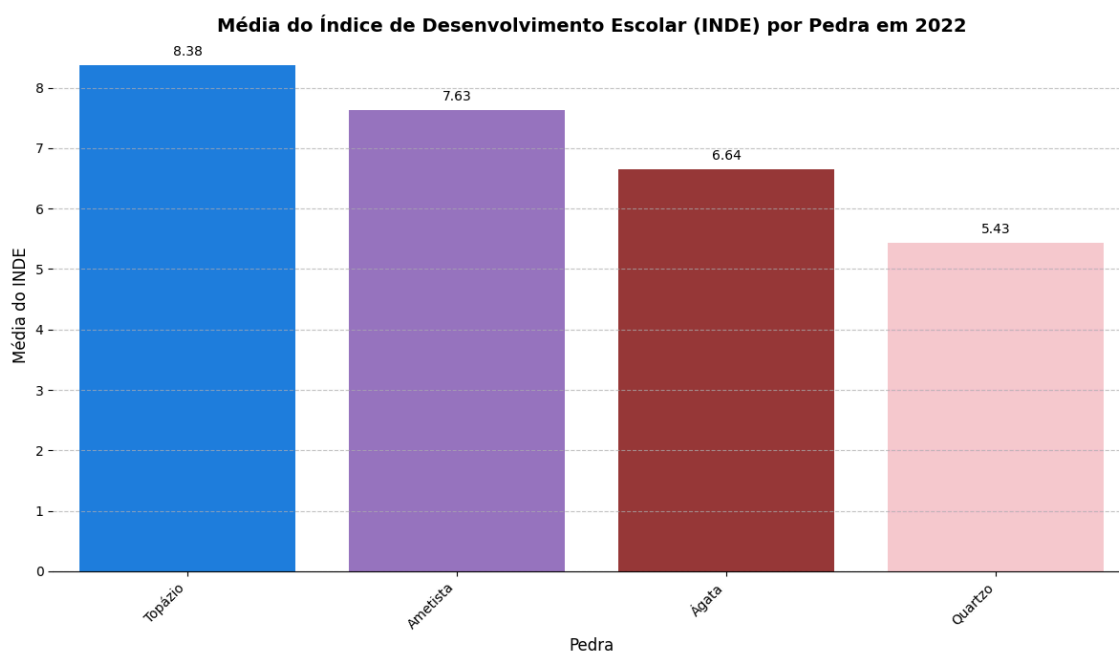
Ao destrinchar os resultados por cada indicador, temos:

Indicadores



Uma média elevada não só no INDE, mas no IPV, próximo a 8, que destaca alunos bem avaliados pelos seus pedagogos e/ou instrutores. E os indicadores IEG e IAA passando a marca de 8. O que indica um forte engajamento com o aprendizado e entrega de atividades com uma disposição ativa para participar de atividades extracurriculares, proporcionando grande peso na nota do INDE.

Ao classificá-los pelas pedras, teremos a seguinte distribuição:



Os resultados demonstram uma notável maioria de alunos que se destacaram entre Ágata, um pouco a média do esperado, e Ametista e Topázio que se sobressaem nos resultados. Cabe destacar a importância que a organização Passos Mágicos dá aos seus alunos com melhores desempenhos, sendo um papel fundamental nesse progresso para melhorarem ainda mais, oferecendo oportunidades de crescimento por meio de bolsas de estudo.

Permanência entre 2020 a 2022 conforme base do PEDE (2020-2022)

Entre os anos de 2020 a 2022, 314 alunos que estiveram no ano de 2020 permaneceram entre 2021 e 2022. Podemos então considerar essa taxa de Retenção em 3 anos:

$$(314 / 727) * 100 = 43.19\%$$

Isso significa que 43.19% dos alunos que ingressaram em 2020 permaneceram na instituição por 3 anos, chegando até 2022.

Análise dos dados de 2023

Os dados de 2023 foram extraídos em 29/01/2025, utilizando uma nova base do PEDE que abrange os anos de 2022, 2023 e 2024. No período em que analisamos essa nova base, identificamos algumas inconsistências em comparação com a base antiga, que ia de 2020 a 2022. No entanto, sintetizamos informações relevantes da nova base para apresentar uma análise abrangente do projeto e do desempenho dos alunos nos últimos quatro anos.

A nova base contém 1014 alunos, de certa forma com dados novos e faltantes no que se refere às bases antigas. Em destaque está o gênero dos alunos (F - Feminino / M - Masculino) que consta na nova base que iremos utilizar para levantar alguns indicadores.

Em 2023, com base nos dados do PEDE 2024, foram registrados 369 novos alunos. A distribuição por fase, desses novos alunos, teve as respectivas porcentagens, foi a seguinte: 154 na fase Alfa (41,7%), 57 na fase 1 (15,4%), 71 na fase 2 (19,2%), 29 na fase 3 (7,9%), 23 na fase 4 (6,2%), 18 na fase 5 (4,9%), 7 na fase 6 (1,9%), 9 na fase 9 (2,4%) e 1 na fase 8 (0,3%). As fases iniciais, especialmente Alfa, 2 e 3, concentraram o maior número de ingressantes, assim como nos anos anteriores.

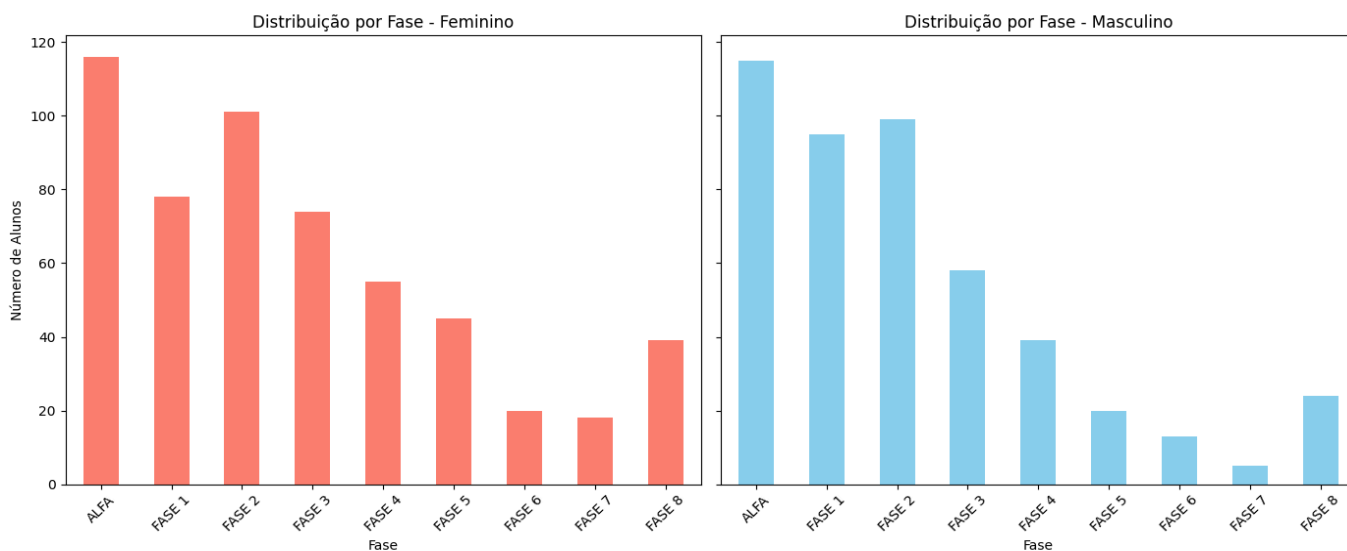
Referente aos alunos que entraram em 2022 e continuam em 2023, correspondem a 268 alunos, do total anterior de 398 alunos ingressantes no ano de 2022, com uma taxa de retenção de 1 ano para 67.34%.

Alunos por Gênero

Do total de 1014 alunos na base em 2023, 546 são do sexo Feminino e 468 do masculino. Entre os alunos novos na Associação, 199 alunos são do sexo Feminino e 170 Masculino.

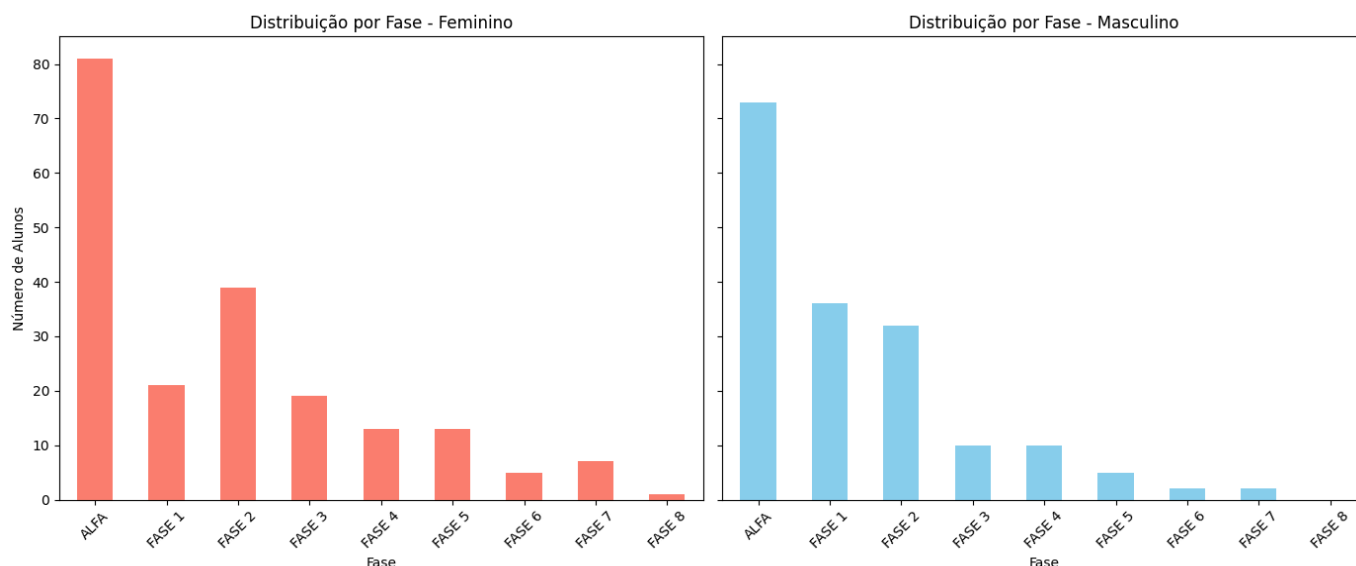
Distribuição de sexo por fase

Total de alunos



A fase Alfa continua apresentando o maior número de alunos. No entanto, o número de alunas superam em quase todas as fases os meninos, com exceção da fase 1 que corresponde a 78 do sexo feminino e 95 masculino, já para as demais, o que temos é uma diferenciação.

Entre os alunos ingressantes



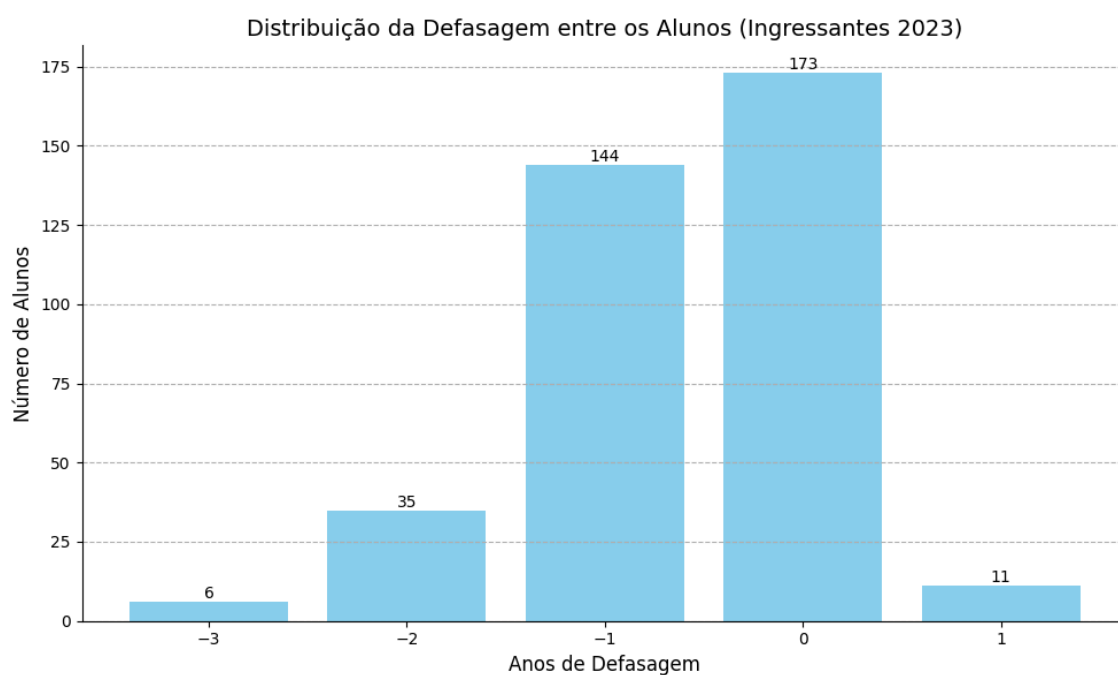
Alunos do sexo feminino estão na frente do sexo masculino em praticamente todas as fases. Vale destacar também que, o único aluno fase 8 (última para os universitários) para esse ano é do sexo feminino.

Instituição de ensino dos alunos em 2023

No total da base de alunos, 809 alunos são de escola pública, enquanto 196 são de instituição privada. Já para os alunos ingressantes, a maioria são de escola pública 354 (95,9%), enquanto 6 (1,6%) são bolsistas de instituições privadas.

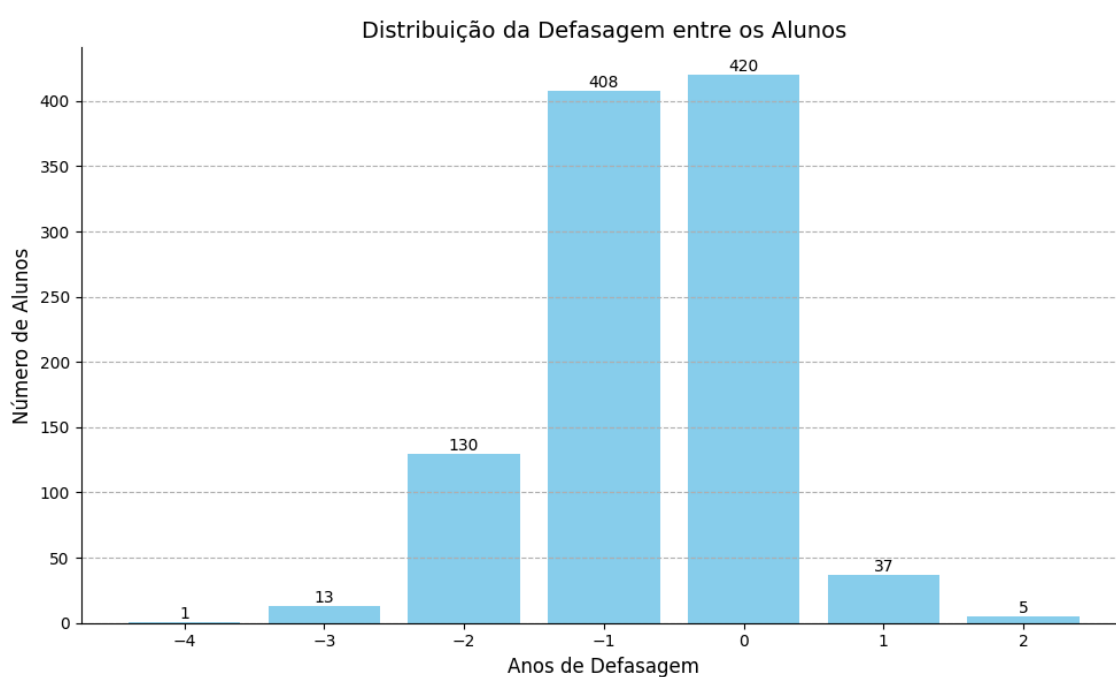
Defasagem entre os alunos em 2023

Alunos que não se encontram na fase ideal para sua idade são determinados com defasagem. Dos ingressantes em 2023, 173 alunos estão em suas fases certas. Já 144 estão atrasados em 1 ano; 35 alunos com defasagem de 2 anos e 6 alunos com 3 anos de defasagem. Somente 11 alunos estão acima do nível padrão de ensino.



Do total de alunos com defasagem de -1, -2 e -3 anos (185), 177 são alunos de escola pública.

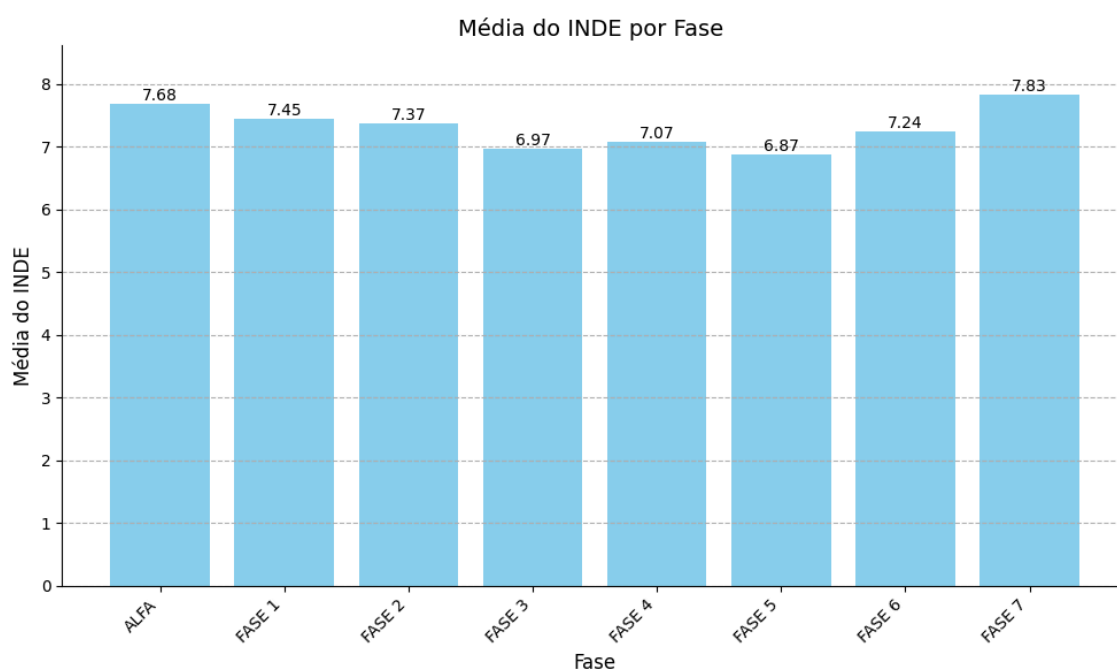
Se olharmos a defasagem como um todo, teremos um grande panorama com muitas informações:



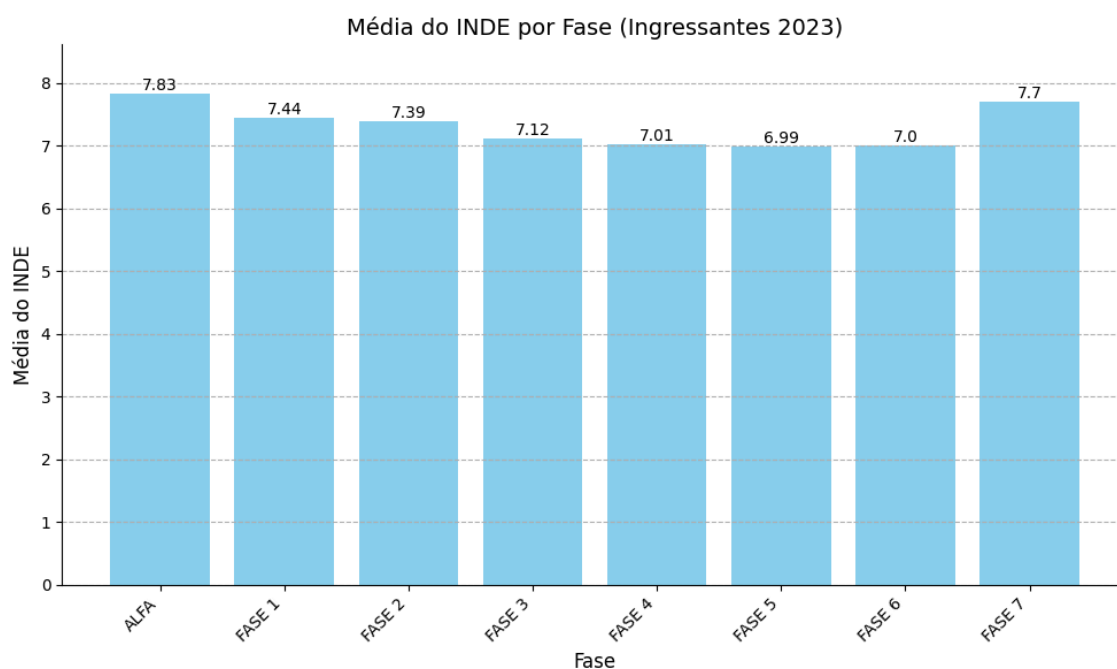
Com 367 alunos atrasados em pelo menos 1 ano que são da escola pública. 113 alunos com 2 anos de defasagem e os 13 alunos apresentados com 3 anos e o único aluno com 4 anos de defasagem são de escola pública também.

INDE médio dos alunos por fase

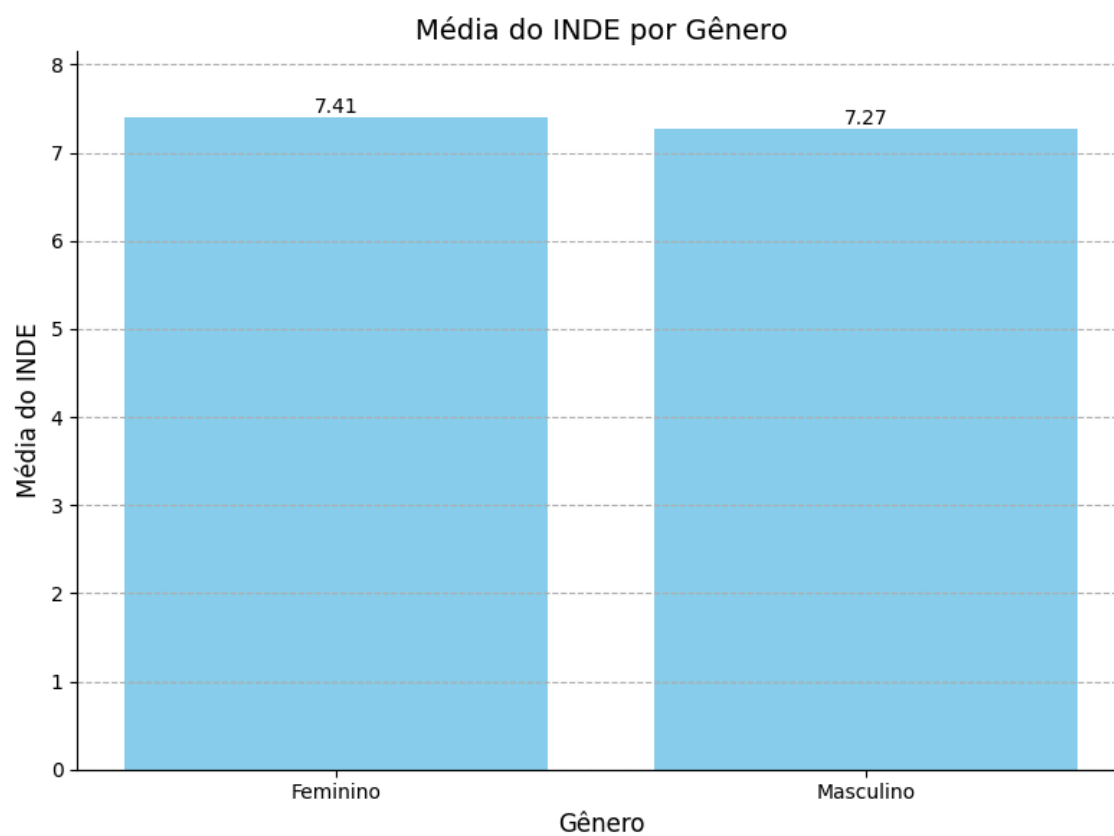
Por total de alunos



Por ingressantes



INDE médio por sexo



Análise dos dados de 2024

Finalmente, analisaremos os dados mais recentes da pesquisa PEDE. Com o total de 1156 alunos registrados na base, a maior já construída até então. Para esse ano, 349 alunos ingressaram e 268 são ingressantes do ano de 2023 que permaneceram em 2024, com a taxa de retenção de aproximadamente 72,63%.

A distribuição por fase de ensino demonstra a seguinte composição: 196 alunos na fase Alfa, 185 na fase 1, 185 na fase 2, 211 na fase 3, 115 na fase 4, 100 na fase 5, 25 na fase 6, 37 na fase 7 e 64 na fase 8.

Em 2024 teve-se um aumento dos alunos na fase 8 que estão cursando o ensino superior, o que promove a motivação que a Associação tem dado através de bolsas de ensino superior aos seus alunos e que tem aumentado esse incentivo.

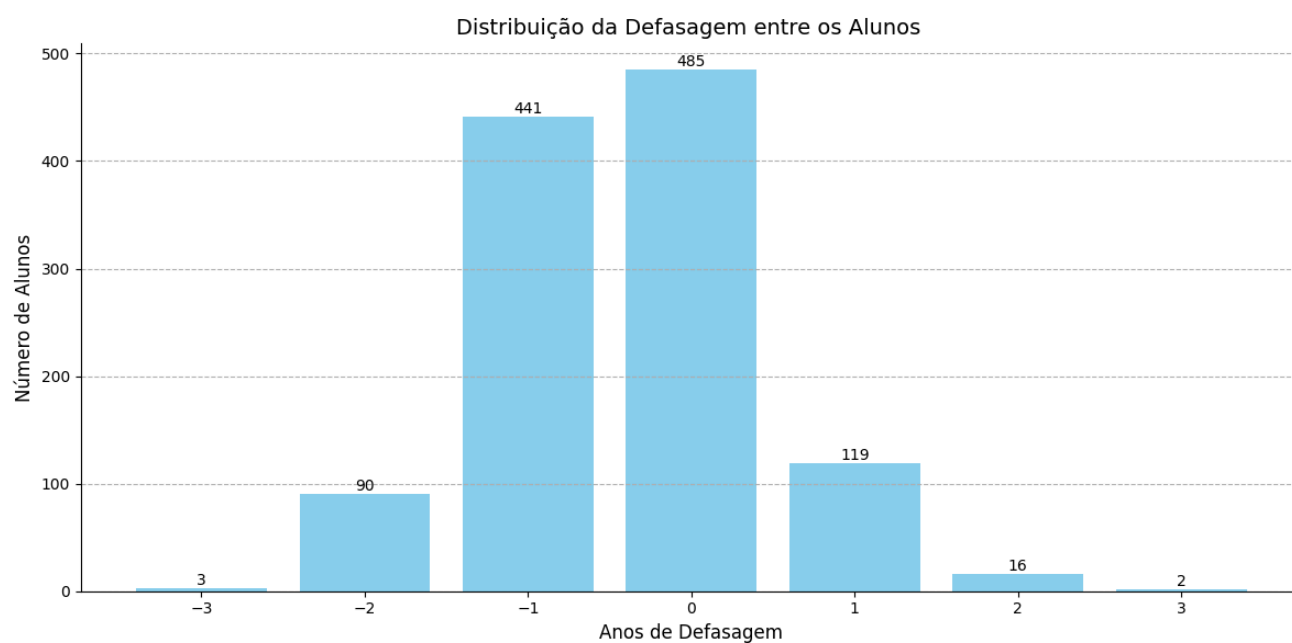
Distribuição por instituição de ensino

No que concerne à procedência escolar, observa-se que 913 alunos cursaram o ensino fundamental e/ou médio em instituições públicas, ao passo que 222 o fizeram em instituições privadas.

Distribuição por gênero

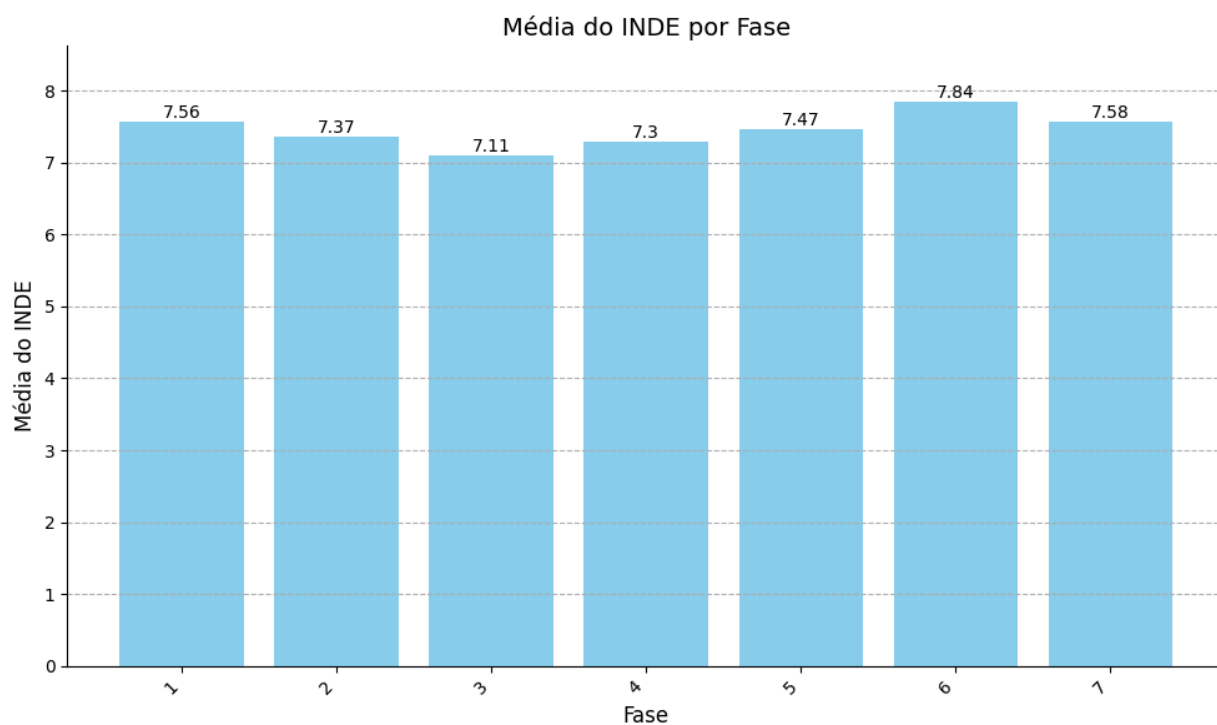
A análise da distribuição por gênero revela a prevalência do sexo feminino em mais um ano, com 623 alunas matriculadas, superando o número de alunos do sexo masculino, que totalizam 533.

Defasagem



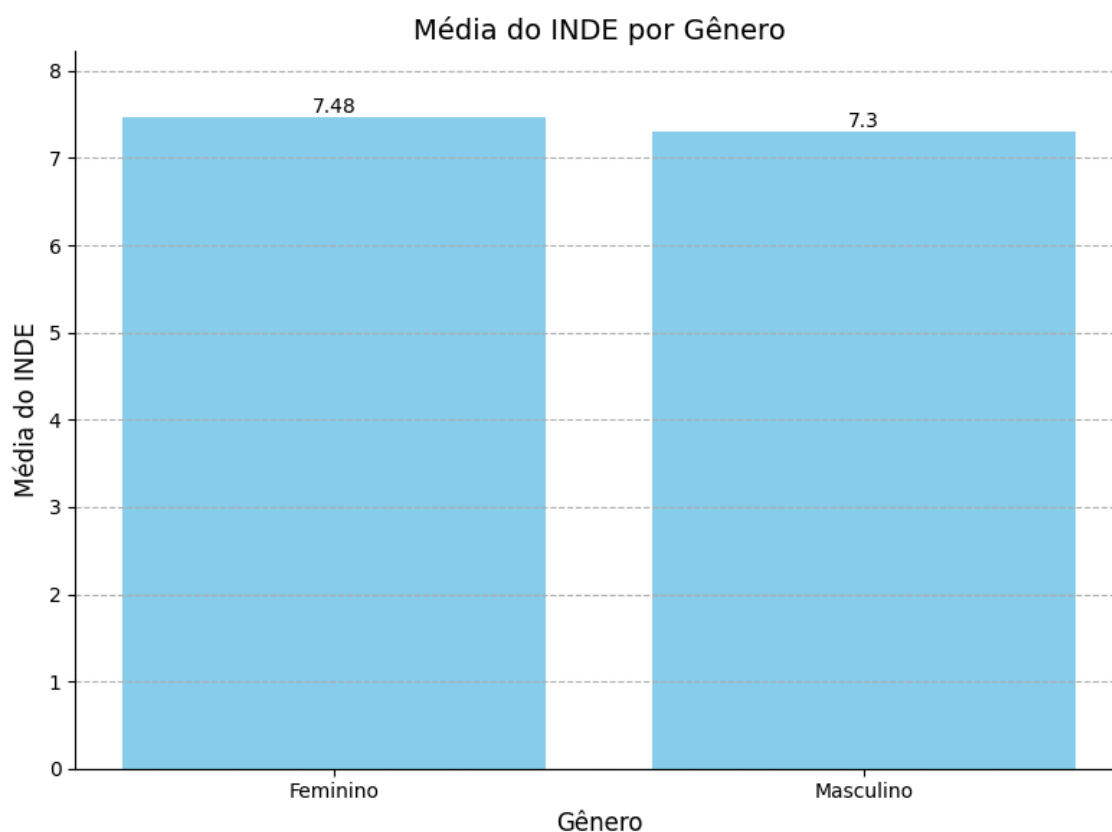
420 alunos de escola pública estão com defasagem de 1 ano. 85 estão com 2 anos de defasagem e os 3 alunos com defasagem de 3 anos também são de escola pública.

INDE médio dos alunos por fase



Para as fases Alfa, a média ficou entre 7,5. Pode-se notar que continua uma variação entre as fases 2 e 3 que tem as maiores dificuldades

INDE médio dos alunos por sexo



Com isso, concluímos a análise de dados referente aos 4 anos de registro pela PEDE feita pela associação Passos Mágicos. Contudo, gostaríamos de, através do que tudo que foi levantado, pesquisado e analisado entre os dados, seja estrategicamente utilizado em insights e tomadas de decisões estratégicas pontuais a fim de aprimorar no que se refere à atuação entre os jovens de maior vulnerabilidade social, com a acessibilidade ao ensino de qualidade, ao acompanhamento profissional e psicopedagógico e a inserção do jovem como protagonista no meio social.

Análise preditiva dos dados

Para a modelagem preditiva, empregamos o algoritmo XGBoost, uma solução de aprendizado de máquina amplamente utilizada baseada em árvores de decisão que se beneficia do método de boosting. No boosting, múltiplas árvores de decisão simples (modelos fracos) são treinadas sequencialmente, com cada árvore subsequente focada em corrigir os erros da(s) anterior(es), resultando em um modelo de alta eficiência e robustez.

Nosso intuito foi prever qual a Pedra, classificação dada ao aluno pelo resultado da sua nota no INDE, através de sua nota no INDE, sua fase e idade.

Utilizamos o Streamlit que é uma biblioteca de código aberto que permite a criação de aplicativos web de maneira prática e de forma gratuita. Com ela é permitido criar interfaces de usuário interativos, gráficos dinâmicos e dashboards completos, com dados em tempo real para os usuários finais, e por meio dela disponibilizamos nossa análise preditiva.

Ao combinar a capacidade preditiva do XGBoost com a praticidade do Streamlit, desenvolvemos e implementamos uma aplicação web para prever a classificação "Pedra" de um aluno.

A aplicação está disponível em: <https://fase5fiapgrupo44.streamlit.app/>.

Insights, Decisões e KPIs: Análises Estratégicas

A análise abrangente dos dados de desempenho educacional e socioeconômicos dos alunos da Associação Passos Mágicos (2020-2024), que considerou diversos fatores como indicadores de performance, revelou insights valiosos. Destacam-se:

Disparidade no desempenho de alunos de instituições públicas e privadas, e necessidade de ações específicas: Observou-se uma disparidade significativa no desempenho entre alunos de diferentes instituições de ensino. Alunos provenientes de escolas públicas apresentam maiores defasagens, indicando a necessidade urgente de atenção e alocação de recursos específicos. Recomenda-se a implementação de acompanhamento pedagógico individualizado e estratégias de reforço focadas nas necessidades específicas desse grupo.

Defasagem e continuidade: A análise revela que a maioria dos alunos com defasagem de pelo menos um ano em relação à sua fase ideal provém de escolas públicas. Essa defasagem impacta negativamente na continuidade no programa de ensino da Passos Mágicos. Propõe-se o desenvolvimento de estratégias focadas nesses alunos, visando motivá-los a permanecer no programa, permitindo que, ao longo do tempo e das fases, recuperem o aprendizado e melhorem seu desempenho. Adicionalmente, sugere-se a implementação de uma avaliação pedagógica contínua e abrangente, que considere o comportamento e o estilo de aprendizagem de cada aluno. Essa avaliação sistemática permitirá identificar o momento ideal para elevar o nível do aluno e confirmar a superação da defasagem, além de validar a eficácia das práticas de ensino implementadas.

Ponto de Virada e Excelência nas Fases Avançadas: A análise do Índice de Desenvolvimento (INDE) aponta que os alunos com resultados acima da média concentram-se nas fases mais avançadas (6 a 8). Esse achado sugere que muitos alunos atingem um "ponto de virada" e demonstram excelência em outros indicadores de desempenho acadêmico com o amadurecimento proporcionado pelo programa. Este "ponto de virada" pode ser melhor compreendido e replicado nas fases iniciais através de um estudo aprofundado das práticas e metodologias utilizadas nas fases 6 a 8.

Isso também é um fator indicador para que o programa traga mais oportunidades de bolsas a esse grupo que estão próximos de ingressar no ensino superior, avaliando pelo mérito desses alunos uma bolsa para as faculdades parceiras.

Fluxo Contínuo nas Fases Iniciais e Impacto no Desempenho: O fluxo contínuo de novos alunos nas fases iniciais (ALFA, Fase 1 e Fase 2) desde 2019, aliado ao baixo INDE médio dessas fases em comparação com as demais, sinaliza uma preocupação crucial. O desempenho inferior desses alunos pode impactar negativamente sua continuidade no programa. Em resposta a esse desafio, recomenda-se a abertura contínua de novas turmas nessas fases, com foco no desenvolvimento de habilidades de comunicação e participação. Dado que muitos alunos apresentam baixos indicadores relacionados ao engajamento, estratégias direcionadas para estimular a participação ativa e o envolvimento no programa são essenciais para melhorar o desempenho geral e a retenção.

Referências

Link do github:

<https://github.com/DevzsJhonny/Datathon-passos-magicos>

Aplicação Streamlit:

<https://fase5fiapgrupo44.streamlit.app/>

Passos Mágicos. Impacto e Transparência - Indicadores de Impacto.

Disponível em:

<https://passosmagicos.org.br/impacto-e-transparencia/#indicadores-de-impacto>.

Acesso em: 4 fev. 2025.

Universidade Federal de Minas Gerais. Qualidade de vida e desempenho:

estudo mostra diferenças entre alunos de escolas públicas e privadas.

Disponível em:

<https://ufmg.br/comunicacao/noticias/qualidade-de-vida-e-desempenho-estudo-mostra-diferencas-entre-alunos-de-escolas-publicas-e-privadas#:~:text=Estudantes%20da%20rede%20privada%20t%C3%A7%C3%A3o,desenvolvida%20na%20Faculdade%20de%20Medicina>. Acesso em: 4 fev. 2025.

PEDE. Relatório PEDE 2020.

PEDE. Relatório PEDE 2021.

PEDE. Relatório PEDE 2022.

PEDE. Dicionário de Dados PEDE.

PEDE. PEDE 2024 DATATHON.