

UNIVERSIDADE FEDERAL DO PARANÁ

AVALIAÇÃO DO DESEMPENHO DOS ESTUDANDES DA UFPR

CURITIBA

2023

ANA CAROLINA CANOSSA BECKER

AVALIAÇÃO DO DESEMPENHO DOS ESTUDANDES DA UFPR

Atividade submetida ao curso de Especialização em Data Science & Big Data, Universidade Federal do Paraná.

CURITIBA

2023

ATIVIDADE 1

Dataset Notas

Vamos obter informações sobre o rendimento de alunos de um curso em algumas disciplinas ao longo dos anos.

O primeiro arquivo contido no Dataset Notas, `historico-alg1_SIGA_ANONIMIZADO.csv`, refere-se ao aproveitamento de estudantes na disciplina ALGORITMOS 1 entre os anos de 2011 e 2022.

A primeira coluna ("matricula") é composta por **números inteiros**, onde cada número representa um indivíduo. Assim, repetições nessa coluna indicam que o estudante fez mais de uma vez a mesma matéria.

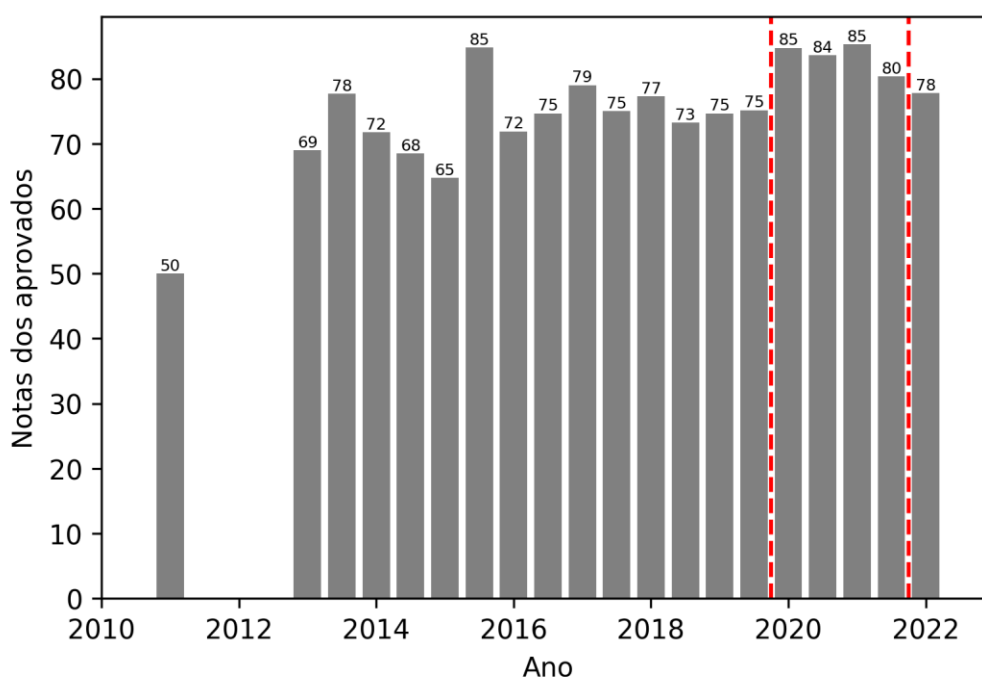
Atenção: R-nota indica REPROVAÇÃO POR NOTA e R-freq REPROVAÇÃO POR FALTA. Se houver outro "status" para representar reprovação, este deve ser trocado para o rótulo adequado (R-nota ou R-freq). Frequências < 75 causam reprovação por falta; Médias abaixo de 50 causam reprovação por nota.

Analise o dataset do referido arquivo para responder as seguintes perguntas:

1. Qual é a média de nota dos aprovados (no período total e por ano)?

A média geral das notas dos alunos aprovados é 78. A média ao longo do tempo está representada na Figura 1. Vale destacar que, para a obtenção dos resultados, foram desconsideradas as aprovações obtidas por meio de equivalência.

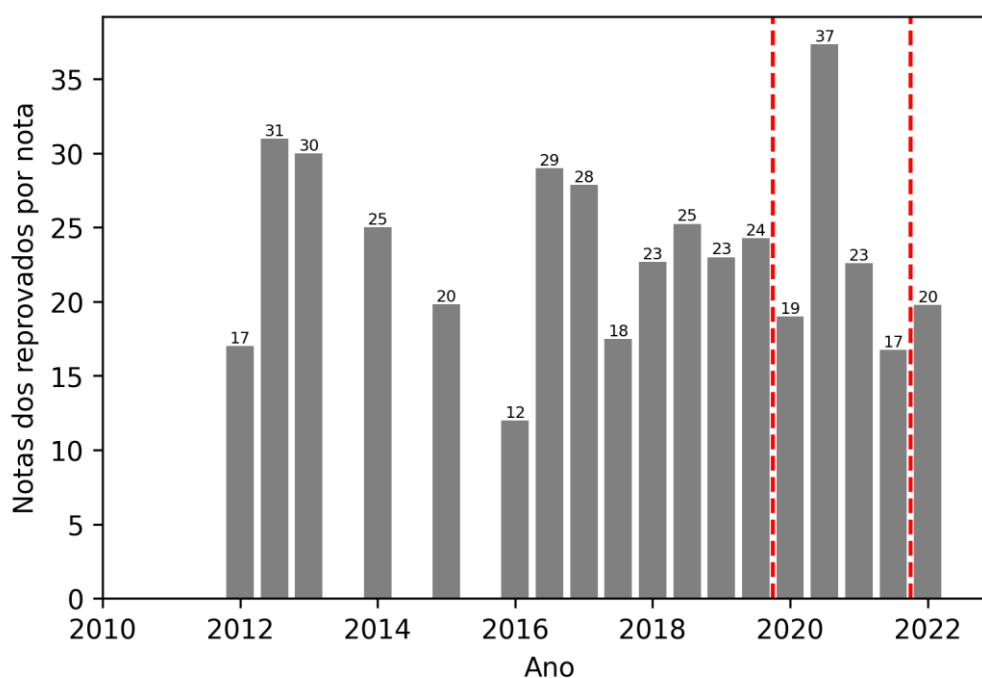
FIGURA 1 – MÉDIA DOS ALUNOS APROVADOS



2. Qual é a média de nota dos reprovados por nota (período total e ano)?

A média geral das notas dos alunos reprovados por nota resultou em 22. A média ao longo do tempo está representada na Figura 2.

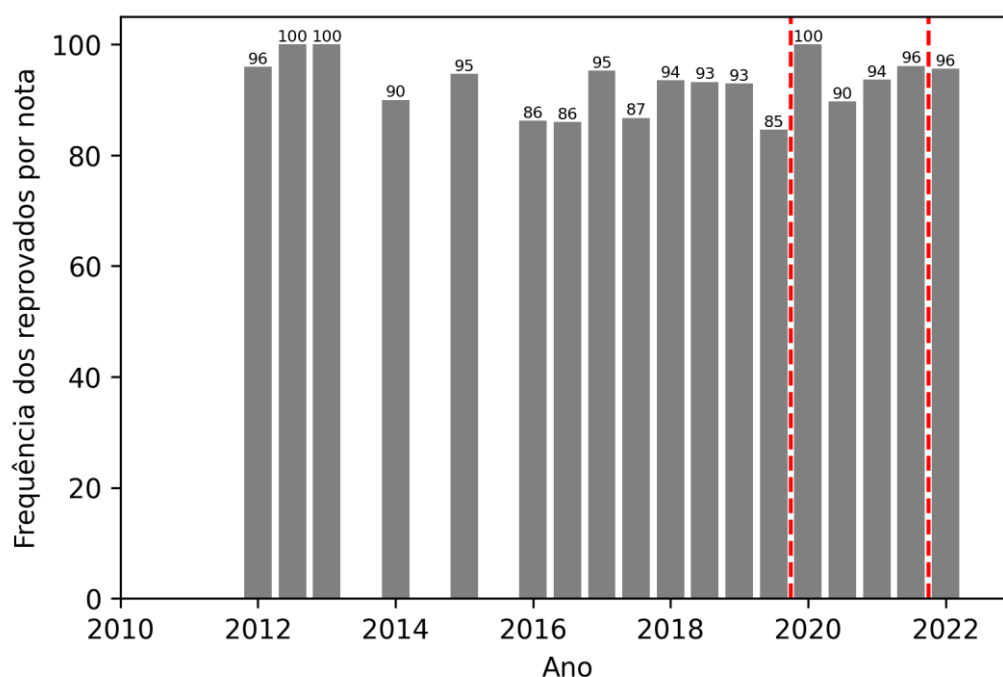
FIGURA 2 – MÉDIA DOS ALUNOS REPROVADOS POR NOTA



3. Qual é a frequência dos reprovados por nota (período total e por ano)?

A frequência geral das notas dos alunos reprovados por nota resultou em 93%. A frequência ao longo do tempo está representada na Figura 3.

FIGURA 3 – FREQUÊNCIA DOS ALUNOS REPROVADOS POR NOTA



4. Qual a porcentagem de evasões (total e anual)?

O percentual geral de evasões foi de 5%. Esta avaliação também foi realizada ao longo do tempo, por meio da proporção do número de evasões (Figura 4) em relação ao número total de alunos (Figura 5). O resultado do percentual de evasões foi apresentado na Figura 6.

FIGURA 4 – NÚMERO DE EVASÕES

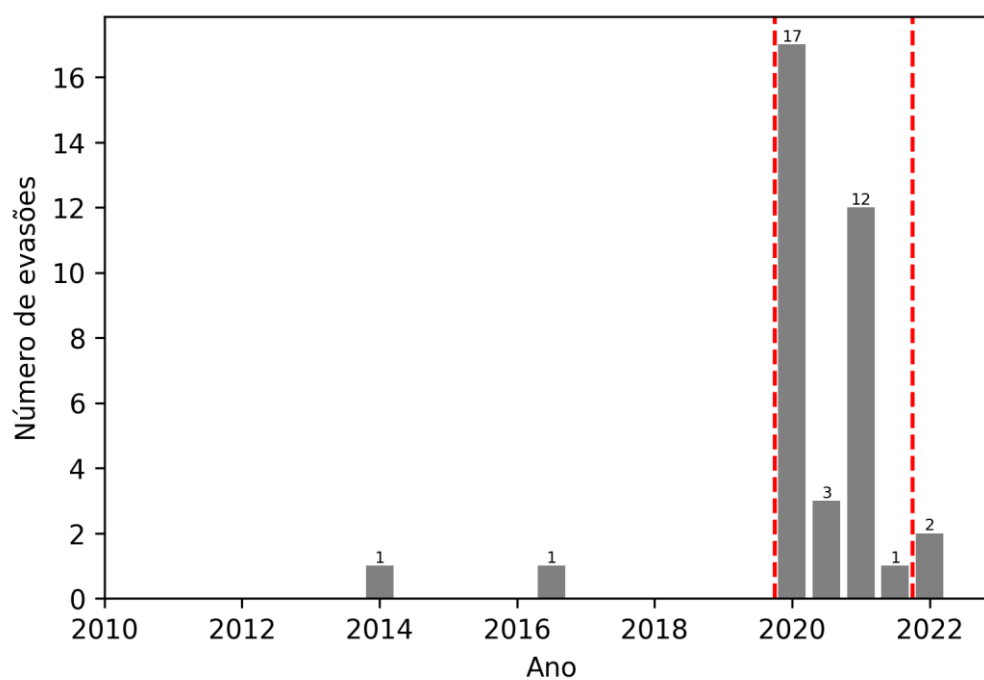


FIGURA 5 – NÚMERO TOTAL DE ALUNOS

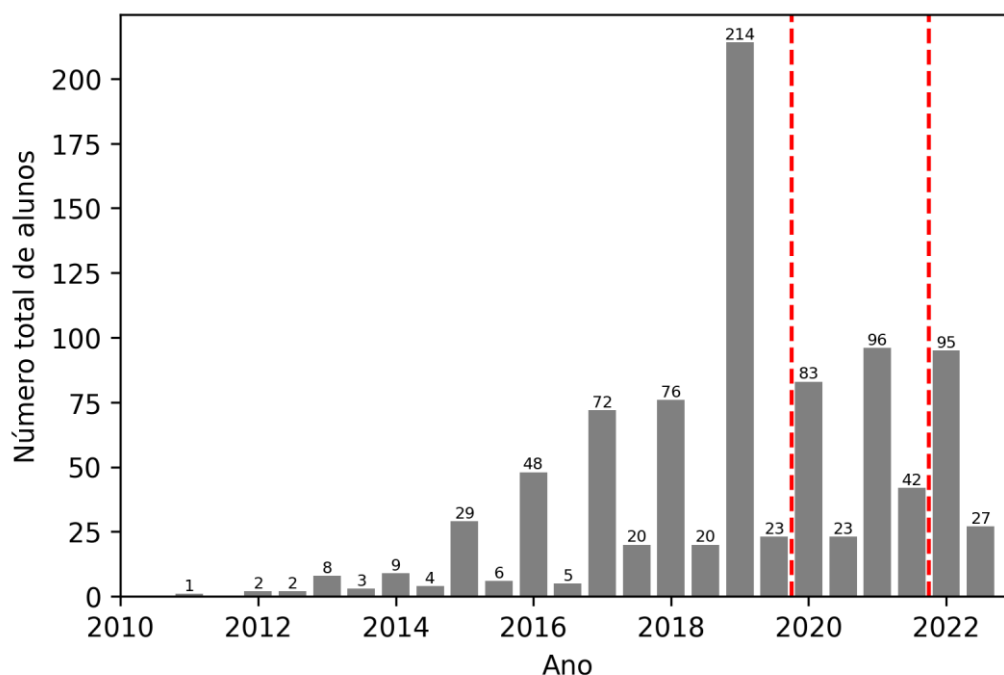
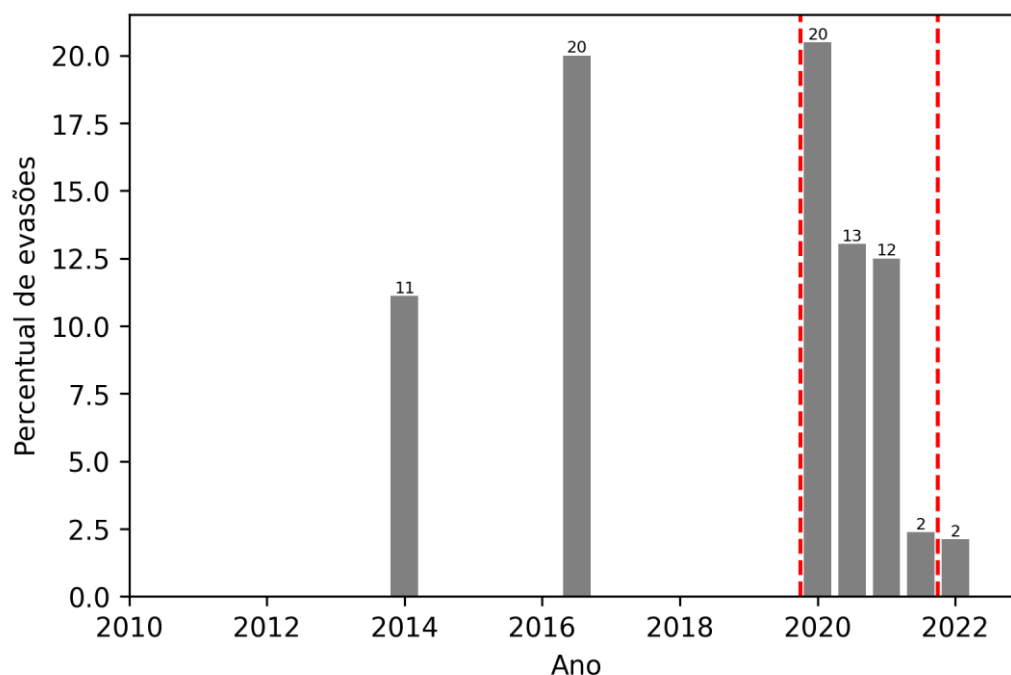


FIGURA 6 – PERCENTUAL DE EVASÕES



5. Como os anos de pandemia impactaram no rendimento dos estudantes em relação aos anos anteriores, considerando o rendimento dos aprovados, a taxa de cancelamento e as reprovações? Considere como anos de pandemia os anos de 2020 e 2021.

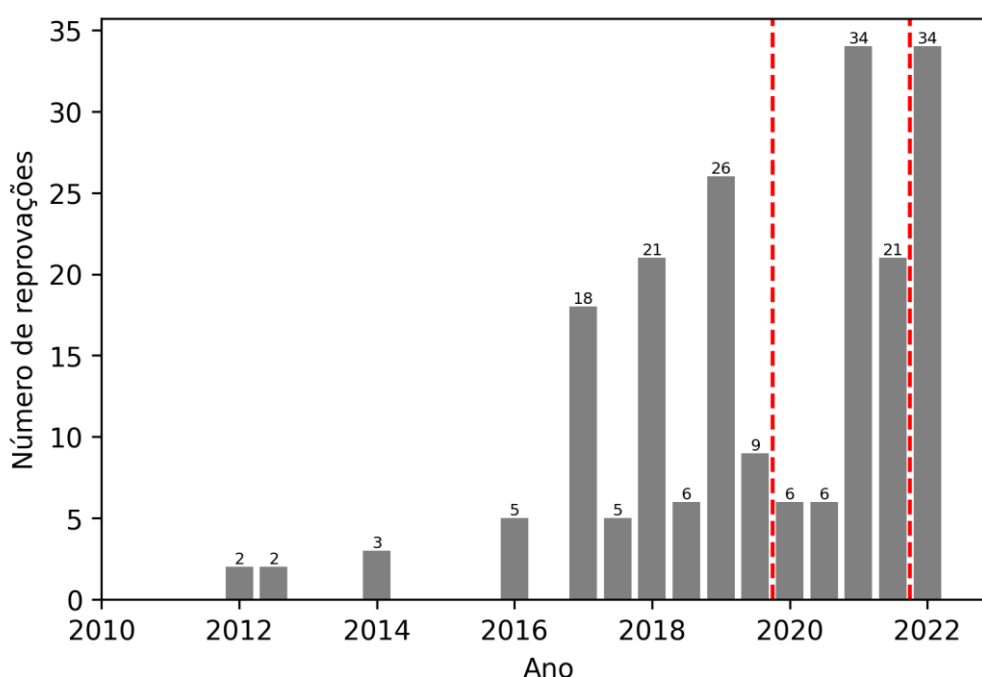
Durante os anos de pandemia, destacados por meio da linha tracejada vermelha nas figuras apresentadas anteriormente, foi verificado que o desempenho dos alunos aprovados

(Figura 1) foi superior à média geral de todos os anos (78), variando entre os valores de 80 a 85 nesse período.

A média dos alunos reprovados por nota (Figura 2) apresentou, no período de pandemia, o maior valor da série histórica. A frequência dos alunos reprovados por nota (Figura 3) não apresentou diferença no período de pandemia.

Conforme pode ser visto nas figuras, a pandemia contribuiu para o aumento do número de evasões (Figura 4), especialmente no primeiro semestre de 2020 e de 2021, nos quais ocorreram 17 e 12 evasões, respectivamente. Adicionalmente, foi verificado aumento do número de reprovações no início de 2021, conforme Figura 7.

FIGURA 7 – NÚMERO DE REPROVAÇÕES



6. Compare a volta às aulas híbrida (2022 período 1) com os anos de pandemia e os anos anteriores.

Conforme Figura 7, as reprovações no primeiro período de 2022 foram na mesma ordem de grandeza do primeiro semestre de 2021, indicando que o retorno às aulas híbridas também causaram alterações nos rendimentos dos alunos. Entretanto, o percentual de evasão diminuiu em relação aos anos de 2020 e 2021.

7. Compare a volta às aulas presencial (2022 período 2) com a volta híbrida do item anterior.

No dataset fornecido não há dados relativos ao rendimento dos alunos no segundo período de 2022. Entretanto, podemos notar que o número total de alunos seguiu o mesmo padrão dos anos anteriores, com diminuição das matrículas no segundo semestre, conforme Figura 5.

ATIVIDADE 2

Dataset notas

Agora vamos obter informações sobre o desempenho geral de estudantes do curso de computação. O dataset do arquivo historico-alg1_SIE_ANONIMIZADO.csv contém a situação de 424 estudantes entre 2011 e 2020 em várias disciplinas. Cada estudante é representado por um número inteiro na primeira coluna (MATR_ALUNO).

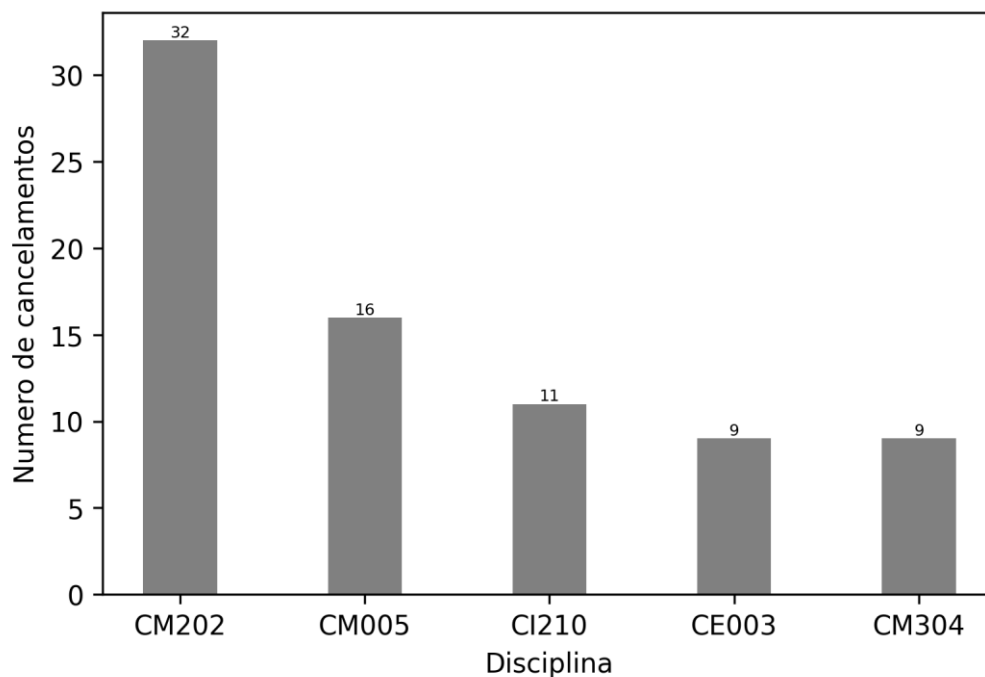
Atenção: algumas matérias (códigos) "deixam" de existir, enquanto outras "surtem", pois houve modificação curricular no período abrangido pelo dataset.

Analise o dataset e responda as seguintes perguntas:

1. Quais as top 5 disciplinas que mais geram cancelamentos por parte dos estudantes ao longo dos anos (plote a resposta em um gráfico);

As disciplinas com maior número de cancelamentos foram mostradas na Figura 8.

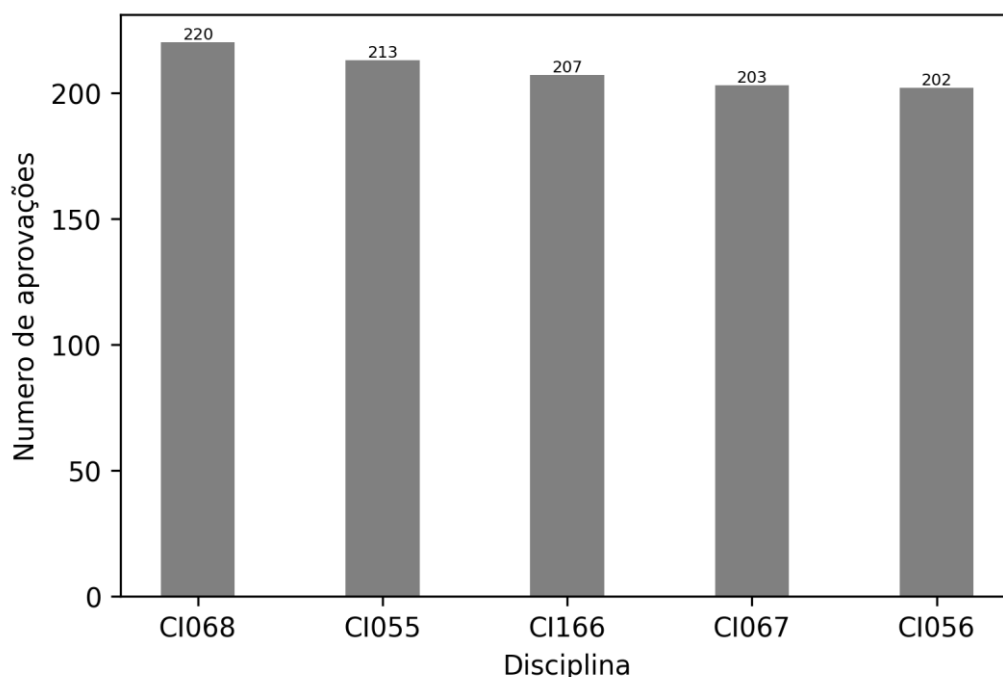
FIGURA 8 – TOP 5 NÚMERO DE CANCELAMENTOS



2. Quais as top 5 disciplinas que possuem aprovações em geral?

As disciplinas com maior número de aprovações foram mostradas na Figura 9.

FIGURA 9 – TOP 5 NÚMERO DE APROVAÇÕES



3. Quais disciplinas possuem taxa de aprovação maior que 70% durante o todo o período?

As disciplinas com taxa de aprovação maior do que 70% foram listadas a seguir:

BC056, CI098, CI1020, CI1086, CI1339, CI1355, BG054, CI318, CI396, CM106, CMA212, CMM102, CMM202, DS210, DS221, HE1020, HE1027, HE872, HL116, HPAC0590, TE042, CI099, TPC225, CF085, CF358, CI094, CI091, BQ054, CI089, CI087, CI071, CI072, CI076, CI074, CEG225, CI082, CI1007, CM100, CI220, PC002, CI302, CI062, PC001, CI057, CI309, CI364, CI061, CI205, CI056, CI316, CI305, CI166, CI301, CI162, CI163, CI204, CI339, CI218, CI058, CI351, CI258, CI256, CI1163, CI068, CI212, CI097

4. Quais disciplinas possuem taxa de reprovação maior que 70% a cada vez que são oferecidas (ex.: DiscXYZ teve 75% de reprovação em 2012; DiscABC teve 99% de reprovação em 2019, etc.)?

As disciplinas com taxa de aprovação maior do que 70% a cada vez que foram oferecidas foram listadas a seguir:

BF075, CI337, CI212, CI251, CI251, CI251, CI256, CI259, CI260, CI303, CI312, CI355, CI169, CM005, CM005, CM043, CM045, CM046, CM046, CMA112, CMM022, EM200, BQ031, CI209, CI164, CI075, CF092, CI055, CI055, CI056, CI057, CI067, CI068, CI068, CI075, CI077, CI083, CI086, CI092, CI1083, CI1251, CI1315, CI1394, TE145, CI210, CI251, CM201, CI081

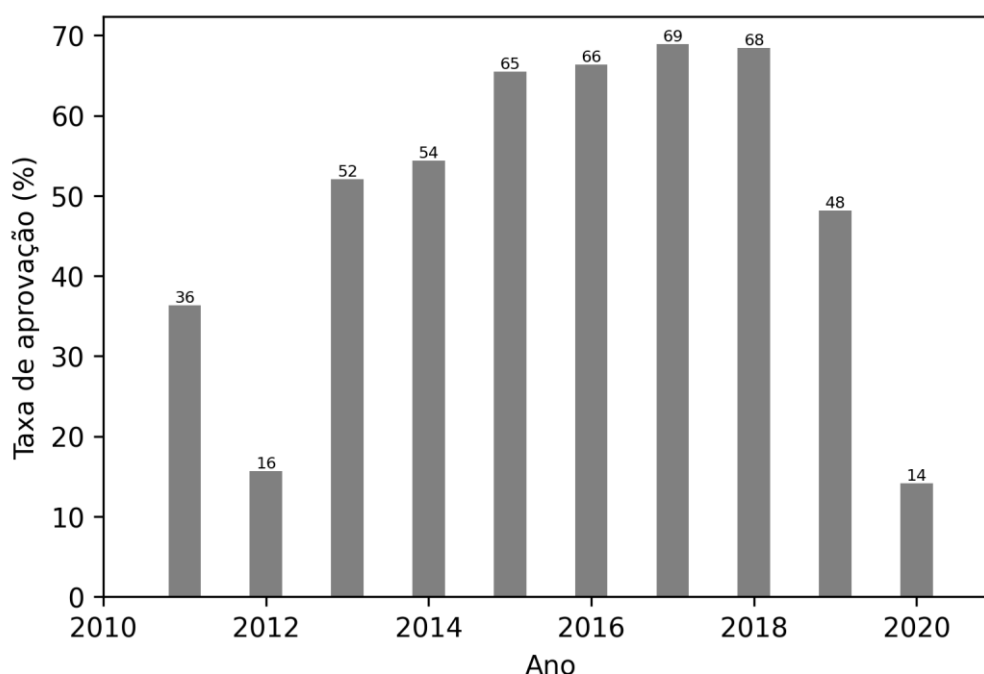
5. Das disciplinas com reprovações, qual o máximo de vezes que um estudante teve que se matricular, excluindo cancelamentos, para ser aprovado?

O número máximo de matrículas do mesmo aluno na mesma disciplina até a ocorrência de um evento de aprovação foi de 6 vezes.

6. Como evoluem as taxas aprovados/reprovados nas disciplinas ao longo do tempo?

A taxa de aprovação (Figura 10) apresentou média geral de 49% e apresentou valores superiores a 60% nos anos de 2015 a 2018. Em 2019, apresentou uma queda para 48%. Vale destacar que o ano de 2020 não está com o dado completo no dataset, com vários alunos ainda com situação "Matriculado".

FIGURA 10 – TAXA DE APROVAÇÃO



A taxa de reprovação (Figura 11) apresentou média geral de 32% e apresenta valores inferiores a 40% desde 2013. Vale destacar que o ano de 2020 não está com o dado completo no dataset, com vários alunos ainda com situação "Matriculado".

FIGURA 11 – TAXA DE REPROVAÇÃO

