

1.1 - diag

17 de julho de 2024



historico_estratificado.csv
Planilha

Analise os dados e limpe as colunas da seguinte forma:
período_ingresso: de 2006.1 até 2019.2
tipo_matricula: remover "Dispensa"
situacao: remover "Em Curso"
o novo dataframe após a limpeza deve se chamar df_filtrado



Para começar, vou carregar os dados do arquivo que você forneceu, dar uma olhada nas primeiras linhas para entender a estrutura, e então aplicar os filtros que você solicitou. Vamos lá!

Agora que temos uma visão geral do dataframe, vamos aplicar os filtros especificados:

1. Filtrar o ``período_ingresso`` para que esteja entre 2006.1 e 2019.2.
2. Remover as linhas onde ``tipo_matricula`` é "Dispensa".
3. Remover as linhas onde ``situacao`` é "Em Curso".

Vou proceder com esses filtros agora. [\[↗\]](#)

O dataframe foi filtrado de acordo com as suas instruções. Agora, ele contém apenas os registros que estão dentro do período de ingresso de 2006.1 a 2019.2, que não têm o tipo de matrícula como "Dispensa", e que não estão na situação "Em Curso". Se precisar de mais alguma análise ou modificação, estou à disposição! [\[↗\]](#)

Qual é a taxa de aprovação por disciplina?



Para calcular a taxa de aprovação por disciplina, vamos considerar as situações em que os alunos foram "Aprovados". Vou fazer isso agora, computando a taxa de aprovação como a proporção de aprovações em relação ao total de registros para cada disciplina.

Aqui estão as taxas de aprovação por disciplina, com algumas delas apresentando uma taxa de aprovação de 100%. Parece que há também algumas disciplinas sem registros suficientes para calcular a taxa de aprovação (resultando em NaN).

Se você quiser mais detalhes sobre alguma disciplina específica ou outra análise relacionada, sinta-se à vontade para pedir! [\[↗\]](#)

Comece com Data Analyst

[Denunciar conteúdo](#) | [Termos de uso](#) | [Política de privacidade](#) |

?