

1.2 - 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



Vou começar analisando o arquivo de dados para entender sua estrutura e, em seguida, fazer as limpezas conforme suas instruções. Vamos lá!

Agora que já visualizamos a estrutura do dataset, vamos prosseguir com a limpeza dos dados conforme suas instruções:

1. Filtrar a coluna ``período_ingresso`` para incluir apenas os períodos de 2006.1 até 2019.2.
2. Remover entradas na coluna ``tipo_matricula`` que contêm o valor "Dispensa".
3. Remover entradas na coluna ``situacao`` que contêm o valor "Em Curso".

Vou aplicar esses filtros agora. [↗]

A limpeza dos dados foi realizada com sucesso. O novo dataframe, ``df_filtrado``, contém entradas apenas no período de ingresso entre 2006.1 e 2019.2, sem tipos de matrícula "Dispensa" e sem situações "Em Curso". O dataframe resultante tem 13,831 entradas.

Se precisar de mais alguma análise ou modificação nos dados, é só avisar! [↗]

Como a frequência de evasão mudou ao longo dos períodos?



Para analisar como a frequência de evasão mudou ao longo dos períodos, vamos começar por definir claramente o que consideramos "evasão". Se evasão for definida pela coluna ``forma_saida`` ou qualquer outra indicação específica de evasão, usaremos isso como base para contar os casos em cada período de ingresso.

Vou verificar as categorias na coluna ``forma_saida`` para identificar como a evasão é registrada nos dados e depois calcular a frequência de evasão por período de ingresso.

A coluna ``forma_saida`` contém as categorias "Graduado", "Evadido", "Reingresso", "Regular" e "Outros". Para análise de evasão, utilizaremos a categoria "Evadido".

Agora, vou calcular a frequência de evasão por período de ingresso. [↗]

Aqui estão as frequências de evasão por período de ingresso. Parece que houve uma tendência geral de aumento na evasão a partir de 2007, alcançando picos entre 2011 e 2013, e então um declínio acentuado após 2016.

Se você quiser visualizar esses dados em um gráfico para uma compreensão mais clara das tendências ao longo do tempo, posso gerar um gráfico para você. Gostaria de ver isso? [↗]

Comece com Data Analyst

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

?