

3.3 - desc

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 lendo o arquivo que você enviou para entender a estrutura dos dados. Em seguida, prosseguirei com a limpeza conforme as suas instruções.

O arquivo carregado possui várias colunas. Agora vou aplicar os filtros conforme suas instruções:

1. **período_ingresso**: Limitar de 2006.1 até 2019.2
2. **tipo_matricula**: Remover linhas onde o valor é "Dispensa"
3. **situacao**: Remover linhas onde o valor é "Em Curso"

Depois disso, eu criei o novo dataframe chamado `df_filtrado`. Vamos proceder com a limpeza dos dados. [↗]

A limpeza dos dados foi concluída com sucesso. O novo dataframe `df_filtrado` contém 13.831 registros que seguem as condições especificadas:

- **Período de ingresso** entre 2006.1 e 2019.2.
- **Tipo de matrícula** diferente de "Dispensa".
- **Situação** diferente de "Em Curso".

Se você precisar de mais análises ou manipulações neste dataframe, ou se desejar exportá-lo, me avise! [↗]

Considere que a média geral do aluno é definida como a média de todas as notas não nulas nas disciplinas que o aluno cursou. Qual é o grau de assimetria na distribuição das médias gerais dos alunos e como isso afeta o desempenho acadêmico geral?



Para responder à sua pergunta, primeiramente calcularei a média geral dos alunos, excluindo as notas nulas, para cada aluno no dataframe filtrado. Em seguida, examinarei a assimetria (skewness) da distribuição dessas médias para entender a tendência de como as notas estão distribuídas – se inclinam mais para notas mais altas, mais baixas ou são relativamente simétricas.

Vamos calcular a média geral dos alunos e a assimetria dessa distribuição.

A distribuição das médias gerais dos alunos possui as seguintes características:

- **Média**: 6.13
- **Desvio padrão**: 2.19
- **Mínimo**: 0.00
- **Máximo**: 9.90
- **Mediana (50%)**: 6.60

A assimetria (skewness) da distribuição é de aproximadamente -0.90. Isso indica uma assimetria negativa, significando que a distribuição das médias tem uma cauda mais pesada para o lado das notas mais baixas e a maioria das notas está inclinada para valores mais altos. Isto sugere que, enquanto a maioria dos alunos tem um desempenho acadêmico relativamente bom (médias mais altas), há um número significativo de alunos com desempenho bastante baixo que puxa a média para baixo.

Essa distribuição pode afetar o desempenho acadêmico geral em várias maneiras:

- **Educação e recursos:** Pode indicar uma necessidade de mais suporte educacional ou de recursos para os alunos que estão tendo desempenho abaixo da média.
- **Políticas acadêmicas:** Pode ser útil revisar as políticas acadêmicas para entender se há barreiras sistêmicas que estão dificultando o desempenho de certos grupos de alunos.
- **Intervenções direcionadas:** Programas de tutoria ou cursos de reforço podem ser necessários para ajudar os alunos que estão lutando para alcançar médias mais altas.

Se precisar de mais análises ou gráficos para visualizar esses dados, estou à disposição para ajudar!

[\[-\]](#)

Comece com Data Analyst

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

?