



Entrar

Cadastrar

Esta é uma cópia de uma conversa entre ChatGPT e Anônimo.  
Denunciar conteúdo

Arquivo carregado

Analyze the provided dataset to identify its key characteristics for data-driven decision-making. Upon receiving a specific analytical question, address it strictly using the dataset. Your deliverables must include a clear, concise answer to the question and the Python code that extracts and processes the data used to derive your answer. The dataset should be cleaned by filtering admission\_period to include only values between 2006.1 and 2019.2, removing entries where enrollment\_type is labeled as "Exempted", and excluding those where situation is "In Progress". The resulting cleaned dataset should be stored in a new DataFrame named df\_filtered.

I'll start by loading the dataset and inspecting its structure to understand its contents. Then, I'll clean the data according to your specified conditions. Let's begin.

The dataset contains 30,130 entries and 34 columns. Now, I'll clean the data by:

1. Filtering admission\_period to include only values between 2006.1 and 2019.2.
2. Removing entries where enrollment\_type is labeled as "Exempted".
3. Excluding records where situation is "In Progress".

I'll create a new DataFrame named df\_filtered with the cleaned data. [~]

The cleaned dataset (df\_filtered) contains 16,563 entries and retains all 34 columns. You can now provide a specific analytical question, and I'll address it strictly using this filtered dataset. [~]

Determine the skewness of the distribution of students' general averages and discuss its implications for evaluating academic performance.

I'll calculate the skewness of the distribution of students' general\_average in the cleaned dataset and discuss its implications for evaluating academic performance. Let's compute it.

The skewness of the general\_average distribution is approximately -4.10, indicating a strong left-skewed (negative) distribution.

#### Implications for Evaluating Academic Performance:

1. **Most Students Have High Scores** – Since the distribution is left-skewed, a majority of students have relatively high general averages, with fewer students scoring significantly lower.
2. **Potential Outliers** – The strong negative skew suggests that a small group of students has much lower grades than the majority, pulling the distribution leftward.
3. **Evaluation Adjustments** – When assessing performance, decision-makers should be cautious about using the mean alone, as it may underestimate most students' performance due to the influence of a few low scores. Instead, the median may provide a better central tendency measure.

Would you like further visualization, such as a histogram, to better illustrate the distribution? [~]

Envie uma mensagem para o ChatGPT

@ Anexar

Voz

© ChatGPT pode cometer erros. Considere verificar informações importantes.

?