

[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 first inspect the dataset to understand its structure and contents. Then, I'll clean it based on your specified criteria. Let me start by loading and previewing the data.

The dataset consists of 30,130 entries with 34 columns. Key columns relevant to filtering are:

- `admission_period` (float64): Must be between 2006.1 and 2019.2.
- `enrollment_type` (object): Entries labeled "Exempted" must be removed.
- `situation` (object): Entries labeled "In Progress" must be removed.

Now, I'll apply these filters and create the cleaned DataFrame, `df_filtered`. [↗](#)

After filtering, the dataset has been reduced from 30,130 entries to 16,563 entries. The cleaned dataset is now stored in `df_filtered`. Let me know your specific analytical question, and I'll proceed with the analysis. [↗](#)

What is the proportion of student dropouts for each year of admission, considering dropout rate as a proportion of total admissions in that year?

To determine the dropout proportion for each admission year:

1. Count the number of students admitted per year.
2. Count how many of those students dropped out.
3. Calculate the dropout rate as:

$$\text{Dropout Rate} = \frac{\text{Dropouts in Year X}}{\text{Total Admissions in Year X}}$$

I'll now compute these proportions using the cleaned dataset (`df_filtered`).

Here are the dropout rates (%) for each admission year:

- **2011–2013** had the highest dropout rates, peaking at **34.96% in 2011**.
- The dropout rate declined after **2014**, reaching **3.92% in 2019**.
- **2006** had the **lowest dropout rate (9.34%)** before the rise in later years.

Would you like any additional insights or visualizations? [↗](#)

Envie uma mensagem para o ChatGPT

@ Anexar

🗣️ Voz

© ChatGPT pode cometer erros. Considere verificar informações importantes.

?